



# CHORUS

This is the accepted manuscript made available via CHORUS. The article has been published as:

## Prospects for Resolving the Hubble Constant Tension with Standard Sirens

Stephen M. Feeney, Hiranya V. Peiris, Andrew R. Williamson, Samaya M. Nissanke, Daniel J. Mortlock, Justin Alsing, and Dan Scolnic

Phys. Rev. Lett. **122**, 061105 — Published 14 February 2019

DOI: [10.1103/PhysRevLett.122.061105](https://doi.org/10.1103/PhysRevLett.122.061105)

# Prospects for resolving the Hubble constant tension with standard sirens

Stephen M. Feeney,<sup>1,\*</sup> Hiranya V. Peiris,<sup>2,3</sup> Andrew R. Williamson,<sup>4</sup>  
Samaya M. Nissanke,<sup>4</sup> Daniel J. Mortlock,<sup>5,6,7</sup> Justin Alsing,<sup>1</sup> and Dan Scolnic<sup>8</sup>

<sup>1</sup>*Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA*

<sup>2</sup>*Department of Physics & Astronomy, University College London, Gower Street, London WC1E 6BT, UK*

<sup>3</sup>*Oskar Klein Centre for Cosmoparticle Physics, Stockholm University, AlbaNova, Stockholm SE-106 91, Sweden*

<sup>4</sup>*Institute of Mathematics, Astrophysics and Particle Physics,*

*Radboud University, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands*

<sup>5</sup>*Astrophysics Group, Imperial College London, Blackett Laboratory, Prince Consort Road, London SW7 2AZ, UK*

<sup>6</sup>*Department of Mathematics, Imperial College London, London SW7 2AZ, UK*

<sup>7</sup>*Department of Astronomy, Stockholm University, AlbaNova, SE-10691 Stockholm, Sweden*

<sup>8</sup>*Kavli Institute for Cosmological Physics, The University of Chicago, Chicago, IL 60637, USA*

The Hubble constant ( $H_0$ ) estimated from the local Cepheid-supernova (SN) distance ladder is in  $3\text{-}\sigma$  tension with the value extrapolated from cosmic microwave background (CMB) data assuming the standard cosmological model. Whether this tension represents new physics or systematic effects is the subject of intense debate. Here, we investigate how new, independent  $H_0$  estimates can arbitrate this tension, assessing whether the measurements are consistent with being derived from the same model using the posterior predictive distribution (PPD). We show that, with existing data, the inverse distance ladder formed from BOSS baryon acoustic oscillation measurements and the Pantheon SN sample yields an  $H_0$  posterior near-identical to the *Planck* CMB measurement. The observed local distance ladder value is a very unlikely draw from the resulting PPD. Turning to the future, we find that a sample of  $\sim 50$  binary neutron star “standard sirens” (detectable within the next decade) will be able to adjudicate between the local and CMB estimates.

## INTRODUCTION

The Hubble constant ( $H_0$ )—the current expansion rate of the Universe [1]—is one of few cosmological parameters that can be estimated locally, using a minimal physical model. Such measurements are invaluable in breaking degeneracies between  $H_0$  and other cosmological parameters (e.g., the spatial curvature of the Universe or number/mass of neutrinos). A plethora of methods exist to estimate  $H_0$ , using Cepheid variables, red-giant stars, SNe, gravitational lenses, galaxies, the CMB and neutron-star mergers [most recently 2–12]. The best cosmology-independent constraints come from the SH0ES Cepheid-SN distance ladder [4]; the tightest constraints come from the *Planck* CMB data, assuming a standard  $\Lambda$ CDM cosmology [3]. [These estimates are discrepant at the  \$3\text{-}\sigma\$  level, suggesting the possibility that the measurements contain unmodeled systematics or that  \$\Lambda\$ CDM is not the true cosmology \[13\].](#)

Numerous attempts have been made to reconcile the two results through new physics [14] or improved astrophysical, experimental and statistical modeling [13, 15], yielding no compelling explanation. Here, we look to the inverse distance ladder and gravitational wave (GW) standard sirens [16] to provide the independent information needed to arbitrate this tension, which we frame in a new, intuitive way using the posterior predictive distribution (PPD). [Unlike existing tension metrics based on the “ \$n\text{-}\sigma\$ ” discrepancy \[e.g., 4\], Kullback-Leibler divergence \[e.g., 17, 18\] or Bayesian evidence ratio \[e.g., 13, 19\], the PPD is simple to interpret and cheap to calcu-](#)

[late for non-Gaussian distributions, and does not require the specification of a \(potentially arbitrary\) alternative model.](#)

## QUANTIFYING TENSION

**Inverse Distance Ladder.** We first demonstrate the PPD’s utility as a tension metric using the inverse distance ladder constructed from existing baryon acoustic oscillation (BAO) and Type Ia SN observations [2]. Galaxy redshift surveys measure the BAO scale parallel and perpendicular to the line of sight,  $\alpha^{\parallel}$  and  $\alpha^{\perp}$ . These are linked to the sound horizon at radiation drag,  $r_d$ , the Hubble parameter  $H(z)$  at the redshift  $z$  of the observations, and the transverse comoving distance [20]  $d_M(z) = \int_0^z c dz'/H(z')$  (assuming a flat universe), by [21, 22]

$$\alpha^{\parallel} = \frac{[H(z)r_d]_{\text{fid}}}{H(z)r_d} \quad \text{and} \quad \alpha^{\perp} = \frac{d_M(z)}{r_d} \left[ \frac{r_d}{d_M(z)} \right]_{\text{fid}}, \quad (1)$$

where the comparison is to a fiducial cosmology. Given a CMB measurement of  $r_d$ , a BAO survey at redshift  $z$  therefore constrains both  $H(z)$  and  $d_M(z)$ .

By adopting a model for  $H(z)$ , the BAO measurements can be extrapolated to redshift zero and hence converted to estimates of  $H_0$ ; however, additional data are required to constrain flexible models. Modern SN surveys are ideal for this task, providing  $\mathcal{O}(10^3)$  relative distance measurements over the relevant redshift range. The apparent magnitude  $m$  of a SN of absolute magnitude  $M$  probes

the luminosity distance,  $d_L(z) = (1+z)d_M(z)$ , via

$$m = 5 \log_{10} \left( \frac{d_L(z)}{\text{pc}} \right) + M - 5. \quad (2)$$

The absolute distance scale of a pure-SN dataset is completely degenerate with the unknown value of  $M$ , but combining with BAO data (transverse measurements in particular) breaks this degeneracy, allowing precise determination of the distance-redshift relation well into the linear regime. The resulting inverse distance ladder prefers [2, 7, 23] values of  $H_0$  in close agreement with that of the *Planck* flat  $\Lambda$ CDM analysis, and is thus in tension with the SH0ES Cepheid distance ladder estimate. The recent release of the Pantheon SNe sample [24]—with 50% greater statistical power than the previous gold standard [25] and a full recalibration of all subsamples used [26]—strongly motivates revisiting this analysis.

In order not to restrict ourselves to a particular physical model, we assume only that the expansion is smooth, adopting the third-order Taylor expansion of the luminosity distance (used by SH0ES):

$$d_L(z) = \frac{cz}{H_0} \left[ 1 + \frac{z}{2}(1 - q_0) - \frac{z^2}{6}(1 - q_0 - 3q_0^2 + j_0) \right], \quad (3)$$

where  $q_0$  and  $j_0$  are the deceleration and jerk parameters.<sup>1</sup> Our inverse distance ladder therefore depends on only five parameters,  $\theta = \{H_0, q_0, j_0, r_d, M\}$ . Given  $n_s$  observed SN apparent magnitudes  $\hat{m}$  and  $n_b$  BAO observations  $\hat{\alpha}$ , the joint posterior of these parameters is

$$\Pr(\theta | \hat{m}, \hat{\alpha}, I) \propto \Pr(\theta | I) N(\hat{m}; \mathbf{m}, \Sigma_s) N(\hat{\alpha}; \boldsymbol{\alpha}, \Sigma_b), \quad (4)$$

where the theoretical SN magnitudes  $\mathbf{m}$  and BAO measurements  $\boldsymbol{\alpha}$  are given by Eqs. 1 to 3,  $N(\mathbf{x}; \boldsymbol{\mu}, \Sigma)$  is a multivariate normal distribution with mean  $\boldsymbol{\mu}$  and covariance  $\Sigma$ , and  $\Sigma_s$  and  $\Sigma_b$  are the SN and BAO covariance matrices. We adopt uniform priors on all parameters apart from  $r_d$ , for which we assume a Gaussian prior **derived from CMB observations**. We sample the joint posterior distribution (Eq. 4) using *emcee* [28].

Combining the BOSS DR12 BAO measurements [22] with the Pantheon SN sample [24] and *Planck*'s “TT+lowP+lensing”  $r_d$  posterior [3] restricts the expansion history to lie within the blue contours on the main panel of Fig. 1, yielding the posterior on  $H_0$  plotted in the left panel. The corresponding contours and posterior from *Planck*'s  $\Lambda$ CDM analysis are overlaid in grey. The inverse distance ladder  $H_0$  constraint ( $68.57 \pm 0.93$

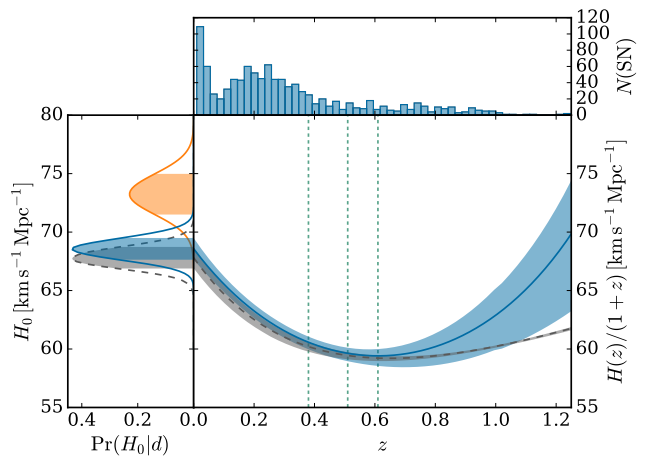


FIG. 1. Main panel: expansion history for BOSS BAO, Pantheon SNe and *Planck*  $r_d$  assuming smooth expansion and early-time physics only (blue), and for *Planck* assuming  $\Lambda$ CDM (grey). BAO redshifts are shown as short-dashed lines. Left panel: corresponding  $H_0$  posteriors and Cepheid distance ladder measurement (orange). Top panel: redshift distribution of Pantheon SNe.

$\text{km s}^{-1} \text{Mpc}^{-1})^2$  is as precise as *Planck*'s  $\Lambda$ CDM constraint ( $67.81 \pm 0.92 \text{ km s}^{-1} \text{Mpc}^{-1}$ )—the flexibility of the model is offset by the extra data—and agrees to well within the 68% credible intervals.

One potential concern here is that using the *Planck*  $r_d$  posterior introduces model inconsistency, as this assumes  $\Lambda$ CDM and includes late-time information from lensing and the integrated Sachs-Wolfe effect. However,  $r_d$  is much less sensitive to late-time physics than  $H_0$ :  $r_d$  constraints do not change significantly when the observational effects of late-time physical processes on the CMB are either removed or marginalized over [29]. For example, removing the lensing likelihood from the *Planck*  $r_d$  posterior shifts our  $H_0$  posterior by less than  $0.2\sigma$ . We conclude that, to a good approximation, the  $r_d$  posterior employed here depends only on the assumption of standard pre-recombination physics.<sup>3</sup> Using the WMAP9  $r_d$  posterior [30] in place of *Planck*'s also does not change the conclusions, yielding  $H_0 = 68.2 \pm 1.1 \text{ km s}^{-1} \text{Mpc}^{-1}$ .

**Posterior Predictive Distribution.** With multiple discrepant  $H_0$  estimates in hand, the task now is to define an intuitive measure of tension, which we will base on the PPD. The PPD is the sampling distribution for new data ( $\mathbf{d}'$ ) given existing data ( $\mathbf{d}$ ) and a model ( $I$ ) [31] and so

<sup>1</sup> While the Taylor expansion is designed for  $z < 1$ , it performs well in our redshift range: modeling the expansion history as a Gaussian Process (following Ref. [27]) instead yields near-identical  $H_0$  posteriors.

<sup>2</sup> We find  $q_0 = -0.50 \pm 0.08$  and  $j_0 = 0.53 \pm 0.38$ , consistent with both SH0ES (who fix  $(q_0, j_0) = (-0.55, 1)$ ) and the Planck Collaboration (who find  $q_0 = -0.54 \pm 0.02$  and fix  $j_0 = 1$ ), and we thus discuss tension in terms of  $H_0$  only.

<sup>3</sup> Changes to the pre-recombination Universe do, however, have an impact: for example, allowing  $N_{\text{eff}}$  to vary yields a significantly broader posterior:  $H_0 = 69.0 \pm 1.6 \text{ km s}^{-1} \text{Mpc}^{-1}$ .

is given by averaging the likelihood of the new data over the posterior of the parameters ( $\theta$ ) describing the model:

$$\Pr(\mathbf{d}'|\mathbf{d}, I) = \int \Pr(\mathbf{d}'|\theta, I) \Pr(\theta|\mathbf{d}, I) d\theta. \quad (5)$$

With new data in hand, the PPD allows discrepancies between the data and model to be assessed: if the new data are not consistent with being drawn from the PPD, the model is not capable of fitting the data, and an alternative should be sought.

Although the PPD is typically employed to check the consistency of a replication of an experiment under the assumed model, there is no requirement for the two datasets to be derived from the same experimental process. Here, its utility in addressing tension between datasets becomes clear: given a dataset and a preferred model, the PPD can be used to simulate different measurements and hence assess whether the two datasets are consistent with being drawn from the same model.

To demonstrate the PPD's utility, we use it to predict the SH0ES data given our inverse distance ladder data. For clarity, rather than predicting the full Cepheid distance ladder dataset<sup>4</sup>, we predict the value of the resultant maximum likelihood estimate of the Hubble constant,  $\hat{H}_0^{\text{CDL}}$ . Converting the inverse distance ladder  $H_0$  posterior into a PPD for  $\hat{H}_0^{\text{CDL}}$ , i.e.,  $\Pr(\hat{H}_0^{\text{CDL}}|\hat{\mathbf{m}}, \hat{\alpha}, I)$ , is done by drawing one sample from the “likelihood”  $\Pr(\hat{H}_0^{\text{CDL}}|H_0, I)$  for each sample from the inverse distance ladder posterior. Taking the likelihood to be a Gaussian with standard deviation  $1.74 \text{ km s}^{-1} \text{ Mpc}^{-1}$ , we obtain the PPD plotted in solid dark blue in Fig. 2; the posterior from which it is derived is plotted in dashed dark blue. The actual  $\hat{H}_0^{\text{CDL}}$  measured by SH0ES, overlaid as a solid orange line, is well into the tails of the PPDs: it is an unlikely draw from this sampling distribution.

In order to quantify the tension we calculate a simple statistic – the “PPD ratio” – defined as the ratio of the PPD at the observed  $\hat{H}_0^{\text{CDL}}$  to its maximum:

$$\rho = \frac{\Pr(\hat{H}_0^{\text{CDL,obs}}|\hat{\mathbf{m}}, \hat{\alpha}, I)}{\max[\Pr(\hat{H}_0^{\text{CDL}}|\hat{\mathbf{m}}, \hat{\alpha}, I)]}. \quad (6)$$

The PPD ratio has a number of advantages over other tension metrics [e.g., 4, 13, 17–19]: it can be generated at the cost of a single likelihood draw per posterior sample; it is simple to calculate even when the posterior is not convex or unimodal; and it is meaningful even in these general settings, unlike other summary statistics (e.g.,

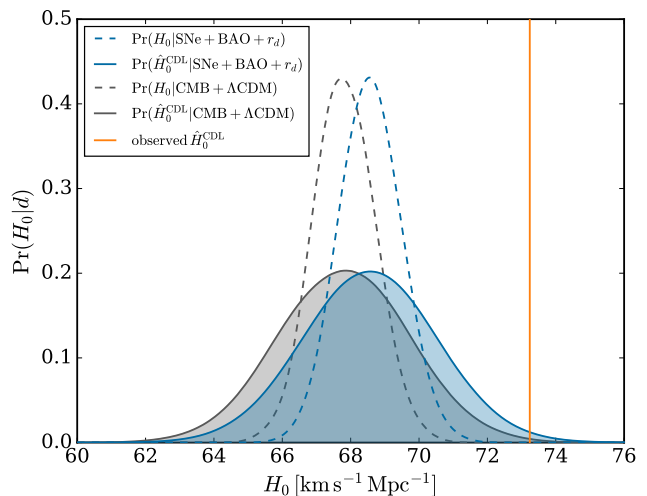


FIG. 2. PPDs (shaded) for the Cepheid distance ladder  $\hat{H}_0$ , conditioned on inverse distance ladder data assuming a smooth expansion history (blue) or CMB data assuming  $\Lambda$ CDM (grey). The SH0ES measurement is plotted as an orange solid line, and the  $H_0$  posteriors from which the PPDs derive are plotted as dashed lines.

$n$ - $\sigma$  discrepancies or  $p$ -values). Finally, unlike model-comparison techniques, there is no need to specify an alternative model, nor is there strong dependence on the prior: informative data will make strong predictions even if the prior is improper.

The PPD ratio is also the likelihood ratio that would result from comparing a null model (that the SH0ES  $\hat{H}_0^{\text{CDL}}$  is a random draw from its PPD) to an alternative “just-so” model in which the true  $H_0$  is fixed to the SH0ES value. As such, the PPD ratio can be interpreted as a lower bound on the posterior probability of the hypothesis that the two experiments measure the same  $H_0$  without systematics. In this instance, the PPD ratio<sup>5</sup> is  $1/17 \simeq 0.06$  at the SH0ES  $\hat{H}_0^{\text{CDL}}$ , so the probability that the distance ladders are unaffected by systematics, and that the apparent discrepancy is simply random, is at least 6%. The PPD constructed from the *Planck*  $\Lambda$ CDM posterior is shifted toward lower  $H_0$  than the inverse distance ladder and so yields a lower ratio of  $1/45$ . For comparison, the  $3$ - $\sigma$  threshold commonly used in the Gaussian setting corresponds to a ratio of  $1/90$ .

<sup>4</sup> Predicting a less processed version of the dataset, e.g., the Cepheid magnitudes, could potentially yield greater insight into undiagnosed systematics at the cost of increasing the dimensionality of the PPDs.

<sup>5</sup> Replacing the approximate Gaussian likelihood with the full asymmetric likelihood of Ref. [13] reduces the quoted values of our PPD ratio less than 10%.

## ARBITRATING TENSION WITH STANDARD SIRENS

Observations of binary neutron star (BNS) mergers offer a method of measuring  $H_0$  [9, 16, 32, 33] that is completely independent of the Cepheid distance ladder and CMB. Fitting a merger’s GW signal yields constraints on the luminosity distance ( $d$ ) to the binary. Where a unique electromagnetic (EM) counterpart can be identified, a spectroscopic redshift for the host may be obtained, allowing a direct estimate of  $H_0$  via

$$cz = v^{\text{P}} + H_0 d. \quad (7)$$

The peculiar velocity ( $v^{\text{P}}$ ) can be left as a nuisance parameter [34] or estimated [9] from ancillary data.

By simulating BNS data we can **investigate** the number of mergers needed to arbitrate the tension between the Cepheids and CMB using the PPD. Consider a set of  $n$  mergers with GW observations  $\{\mathbf{x}\}$ , peculiar velocity estimates  $\{\hat{v}^{\text{P}}\}$  and perfectly observed redshifts  $\{\hat{z}\}$ . Assuming Gaussian  $\hat{v}^{\text{P}}$  likelihoods (with uncertainties  $\sigma_i$ ) and a Gaussian  $v^{\text{P}}$  prior (of width  $\sigma$ ), the marginal  $H_0$  posterior becomes

$$\Pr(H_0|\{\mathbf{x}\}, \{\hat{v}^{\text{P}}\}, \{\hat{z}\}, I) \propto \Pr(H_0|I) \prod_{i=1}^n \int dd_i \Pr(d_i|\mathbf{x}_i, I) \text{N}\left(H_0 d_i; c \hat{z}_i - \frac{\sigma^2 \hat{v}_i^{\text{P}}}{\sigma^2 + \sigma_i^2}, \frac{\sigma^2 \sigma_i^2}{\sigma^2 + \sigma_i^2}\right) \quad (8)$$

(see supplemental material for more detail) if the events are selected by their GW signal-to-noise ratio (SNR) [c.f. 9, 35]. Converting this posterior into a PPD for the CMB or Cepheid distance ladder measurements is a straightforward integral with the relevant “likelihood”  $\Pr(\hat{H}_0|H_0, I)$ .

We **simulate a sample of BNS mergers and process it using the same Bayesian parameter-estimation pipeline as employed on real data, including the effects of amplitude and phase calibration uncertainties.** We simulate BNS detections during the next three LIGO-Virgo (LV) observing runs assuming an underlying rate of  $3000 \text{ Gpc}^{-3} \text{ yr}^{-1}$  (consistent with the bounds from GW170817 [36] at 90% confidence), and a three-detector duty cycle of 40%. Events are **assumed to be independently** distributed uniformly in co-moving volume, with NS masses drawn from the Gaussian  $\text{N}(m_i; 1.4 M_{\odot}, (0.2 M_{\odot})^2)$  restricted to the range 1–3  $M_{\odot}$ . Binary orientations and NS spins are isotropically oriented, with spin magnitudes  $\leq 0.05$  [36]. Each simulated waveform is generated using a time-domain post-Newtonian approximation [37, 38] and embedded in colored Gaussian noise realizations with power spectral densities [39, Fig. 1] appropriate to the detection date:  $\sim 2019$  (1 year);  $\sim 2021$  (1 year); and 2022+ (Design, 2

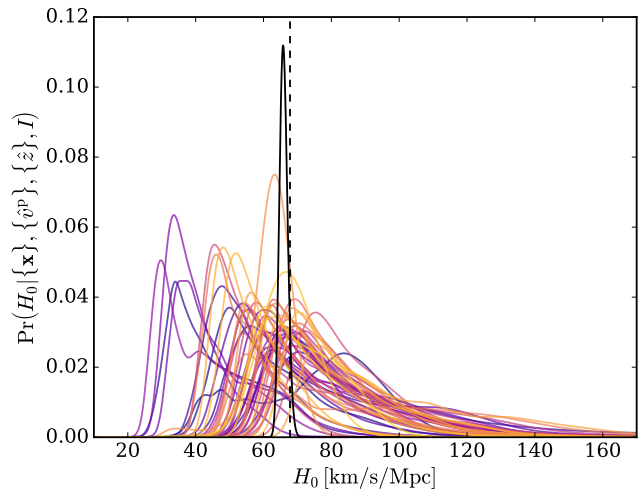


FIG. 3.  $H_0$  posteriors for individual BNS mergers (purple to yellow, sorted by signal-to-noise) and the full sample (black solid; scaled by a factor of 1/3), assuming a true  $H_0$  of  $67.81 \text{ km s}^{-1} \text{ Mpc}^{-1}$  (black dashed).

years). We deem BNS events “GW detectable” when two or more detectors have  $\text{SNRs} \geq 6$ , and the network has  $\text{SNR} \geq 12$ . This yields 51 detections. Fixing the sky position by assuming known host galaxies, we sample the parameter posteriors for each detection using a complete Bayesian MCMC analysis [40] with a frequency-domain post-Newtonian waveform model [37, 38] spanning the range 30–2048 Hz.<sup>6</sup> For estimating  $H_0$ , we retain each event’s distance posterior, marginalizing over all other parameters.

To complete the simulated dataset we need  $\hat{v}^{\text{P}}$  estimates and hence a true  $H_0$ . For illustrative purposes, we use two true  $H_0$  values, assuming either *Planck* or SH0ES is correct. We generate Gaussian measurement errors for each source’s  $\hat{v}^{\text{P}}$  with standard deviation  $200 \text{ km s}^{-1}$  [24, 41]. The  $H_0$  posterior for the resulting simulated BNS dataset (assuming a true  $H_0$  of  $67.81 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ) is plotted in Fig. 3, along with posteriors for each individual event, colored by SNR. Our 1.8%  $H_0$  constraint from 51 mergers is in good agreement with the recent analysis of Ref. [35]. This complementary study uses an approximate 3D localization of GW sources [42] to rapidly average over samples of mergers between compact objects of a single mass, with or without EM counterparts. Ref. [35] finds that  $\sim 60$  mergers between 1.4  $M_{\odot}$  BNSs will, on average, constrain  $H_0$  to 2% assuming unique EM counterparts can be identified.

To convert the BNS  $H_0$  posteriors to PPDs for the CMB and Cepheid distance ladder measurements, we take Gaussian likelihoods  $\Pr(\hat{H}_0|H_0, I)$  with standard

<sup>6</sup> This takes a few CPU weeks per BNS posterior.

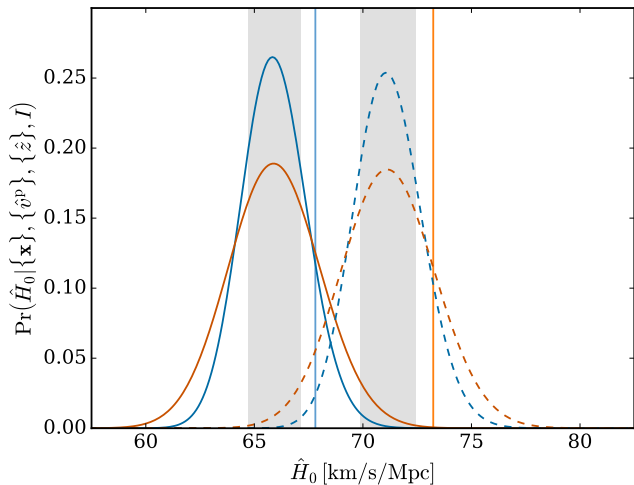


FIG. 4. PPDs for the CMB (dark blue) and Cepheid distance ladder (dark orange)  $\hat{H}_0$  measurements, given the simulated BNS data. Solid/dashed curves assume the true  $H_0$  to be the *Planck*/SH0ES measured value, indicated by the light blue/orange solid line. The  $1\text{-}\sigma$  variations in PPD means due to sample variance are shaded grey.

deviations of  $0.92$  and  $1.74 \text{ km s}^{-1} \text{ Mpc}^{-1}$ , respectively. The results are plotted in Fig. 4. The solid curves, for which we assume the *Planck*  $H_0$  is correct, demonstrate the ability of this BNS sample to arbitrate the tension. The observed SH0ES  $\hat{H}_0$  (solid light orange) would be an extremely unlikely draw from its sampling distribution (solid dark orange): the PPD ratio is  $\sim 1/300$ , much lower than the  $3\text{-}\sigma$  equivalent ratio of  $1/90$ . The *Planck* observation (solid light blue) would, as expected, be consistent with its PPD (solid dark blue). The BNS and CMB observations would decisively favor the underlying value of  $H_0$  used in the simulations.

The *dashed* curves in Fig. 4, in which we assume the SH0ES  $H_0$  is correct, demonstrate another important aspect of this analysis: sample variance due to the limited number of detectable events. The posterior for our simulated sample happens to be scattered to low  $H_0$ .<sup>7</sup> Even though the BNS data strongly constrain  $H_0$ —the posterior uncertainty is  $1.2 \text{ km s}^{-1} \text{ Mpc}^{-1}$ , less than a quarter of the tension—sample variance means we could not arbitrate in favor of one dataset. Indeed, we should expect to see realization-dependent variations in the PPD means on the scale of the posterior standard deviation. We confirm this using 1000 bootstrapped resamples of our dataset, shading the range of PPD means in grey in Fig. 4. As such, while samples of  $\sim 50$  BNS mergers are certainly sufficient to arbitrate the tension, realization noise plays a role in determining whether it is possible

for a given dataset. If the SH0ES  $H_0$  measurement is correct, samples of  $\sim 80$  events will arbitrate the tension even if the BNS  $H_0$  posterior is shifted by  $1\text{-}\sigma$  towards the *Planck* estimate by realization noise. If *Planck* is correct, significantly larger samples ( $\sim 3000$ ) are needed, as the PPD width is dominated by the SH0ES likelihood, which is independent of the BNS sample size.

## CONCLUSIONS

We have demonstrated how existing and upcoming datasets can arbitrate the tension between estimates of  $H_0$  from the CMB and local distance ladder. Throughout, we adopt the minimal cosmological model: a smooth expansion history and standard pre-recombination physics. We find that the inverse distance ladder formed from BOSS BAO measurements and the Pantheon SN sample yields an  $H_0$  posterior near-identical to *Planck* and inconsistent with the observed local distance ladder value. We quantify this tension using a model-testing framework based on the posterior predictive distribution, which relies only on the sampling distribution for one dataset conditional on another, finding that the probability that the two distance ladders measure  $H_0$  without systematics is at least 6%. We then demonstrate how a typical sample of  $\sim 50$  BNS standard sirens, detectable by the LIGO and Virgo experiments within a decade, can independently arbitrate this tension.

## ACKNOWLEDGMENTS

We are very grateful for useful discussions with Lauren Anderson, George Efstathiou, Archisman Ghosh, David W. Hogg, Daniel Holz, Benjamin Joachimi and Benjamin Wandelt. HVP is grateful for the hospitality of the Centro de Ciencias de Benasque Pedro Pascual, where elements of this work were conceived. HVP was partially supported by the European Research Council (ERC) under the European Community’s Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement number 306478-CosmicDawn, and the research environment grant “Gravitational Radiation and Electromagnetic Astrophysical Transients (GREAT)” funded by the Swedish Research council (VR) under Dnr 2016-06012. ARW and SMN acknowledge the generous financial support of the Netherlands Organization for Scientific Research through the NWO VIDI No. 639.042.612 (PI:Nissanke) and NWO TOP Grants No. 62002444 (PI:Nissanke). DS is supported by NASA through Hubble Fellowship grant HST-HF2-51383.001 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS 5-26555. The Flatiron Institute is sup-

<sup>7</sup> We have verified in further work that the  $H_0$  posterior is statistically unbiased in this setting [43].

ported by the Simons Foundation. This work was partially enabled by funding from the UCL Cosmoparticle Initiative.

---

\* sfeeney@flatironinstitute.org

- [1] E. Hubble, PNAS **15**, 168 (1929).
- [2] BOSS Collaboration, Phys. Rev. D **92**, 123516 (2015), arXiv:1411.1074.
- [3] Planck Collaboration, A&A **594**, A13 (2016), arXiv:1502.01589.
- [4] A. G. Riess *et al.*, Astrophys. J. **826**, 56 (2016), arXiv:1604.01424.
- [5] I. S. Jang and M. G. Lee, ArXiv e-prints (2017), arXiv:1702.01118.
- [6] V. Bonvin *et al.*, MNRAS **465**, 4914 (2017), arXiv:1607.01790.
- [7] G. E. Addison, D. J. Watts, C. L. Bennett, M. Halpern, G. Hinshaw, and J. L. Weiland, ArXiv e-prints (2017), arXiv:1707.06547.
- [8] J. W. Henning *et al.*, Astrophys. J. **852**, 97 (2018), arXiv:1707.09353.
- [9] B. P. Abbott *et al.*, Nature (London) **551**, 85 (2017), arXiv:1710.05835.
- [10] DES Collaboration, ArXiv e-prints (2017), arXiv:1711.00403.
- [11] J. Vega-Ferrero, J. M. Diego, V. Miranda, and G. M. Bernstein, ArXiv e-prints (2017), arXiv:1712.05800.
- [12] A. Gómez-Valent and L. Amendola, ArXiv e-prints (2018), arXiv:1802.01505.
- [13] S. M. Feeney, D. J. Mortlock, and N. Dalmasso, ArXiv e-prints (2017), arXiv:1707.00007.
- [14] M. Wyman, D. H. Rudd, R. A. Vanderveld, and W. Hu, Physical Review Letters **112**, 051302 (2014), arXiv:1307.7715 [astro-ph.CO]; A. Pourtsidou and T. Tram, Phys. Rev. D **94**, 043518 (2016), arXiv:1604.04222; E. Di Valentino, A. Melchiorri, and J. Silk, Physics Letters B **761**, 242 (2016), arXiv:1606.00634; Q.-G. Huang and K. Wang, EPJC **76**, 506 (2016), arXiv:1606.05965; J. L. Bernal, L. Verde, and A. G. Riess, JCAP **10**, 019 (2016), arXiv:1607.05617; P. Ko and Y. Tang, Physics Letters B **762**, 462 (2016), arXiv:1608.01083 [hep-ph]; T. Karwal and M. Kamionkowski, Phys. Rev. D **94**, 103523 (2016), arXiv:1608.01309; S. Kumar and R. C. Nunes, Phys. Rev. D **94**, 123511 (2016), arXiv:1608.02454; B. Santos, A. A. Coley, N. Chandrachani Devi, and J. S. Alcaniz, JCAP **2**, 047 (2017), arXiv:1611.01885; V. Prilepina and Y. Tsai, JHEP **9**, 33 (2017), arXiv:1611.05879 [hep-ph]; G.-B. Zhao *et al.*, Nature Astronomy **1**, 627 (2017), arXiv:1701.08165; M.-M. Zhao, D.-Z. He, J.-F. Zhang, and X. Zhang, Phys. Rev. D **96**, 043520 (2017), arXiv:1703.08456; J. Solà, A. Gómez-Valent, and J. de Cruz Pérez, Physics Letters B **774**, 317 (2017), arXiv:1705.06723; E. Di Valentino, E. Linder, and A. Melchiorri, ArXiv e-prints (2017), arXiv:1710.02153; E. Di Valentino, C. Bøehm, E. Hivon, and F. R. Bouchet, *ibid.* (2017), arXiv:1710.02559.
- [15] G. Efstathiou, MNRAS **440**, 1138 (2014); D. N. Spergel, R. Flauger, and R. Hložek, Phys. Rev. D **91**, 023518 (2015), arXiv:1312.3313; M. Rigault *et al.*, Astrophys. J. **802**, 20 (2015), arXiv:1412.6501; D. O. Jones, A. G. Riess, and D. M. Scolnic, Astrophys. J. **812**, 31 (2015), arXiv:1506.02637; G. E. Addison, Y. Huang, D. J. Watts, C. L. Bennett, M. Halpern, G. Hinshaw, and J. L. Weiland, Astrophys. J. **818**, 132 (2016), arXiv:1511.00055; Planck Collaboration, A&A **596**, A107 (2016), arXiv:1605.02985; W. Cardona, M. Kunz, and V. Pettorino, ArXiv e-prints (2016), arXiv:1611.06088; B. R. Zhang, M. J. Childress, T. M. Davis, N. V. Karpenka, C. Lidman, B. P. Schmidt, and M. Smith, MNRAS **471**, 2254 (2017), arXiv:1706.07573; H.-Y. Wu and D. Huterer, MNRAS **471**, 4946 (2017), arXiv:1706.09723; B. Follin and L. Knox, *ibid.* (2017), arXiv:1707.01175; S. Dhawan, S. W. Jha, and B. Leibundgut, *ibid.* (2017), arXiv:1707.00715.
- [16] B. F. Schutz, Nature (London) **323**, 310 (1986).
- [17] S. Seehars, S. Grandis, A. Amara, and A. Refregier, Phys. Rev. D **93**, 103507 (2016), arXiv:1510.08483.
- [18] Y. Wang, L. Xu, and G.-B. Zhao, Astrophys. J. **849**, 84 (2017), arXiv:1706.09149.
- [19] P. Marshall, N. Rajguru, and A. Slosar, Phys. Rev. D **73**, 067302 (2006), astro-ph/0412535.
- [20] D. W. Hogg, ArXiv Astrophysics e-prints (1999), astro-ph/9905116.
- [21] N. Padmanabhan and M. White, Phys. Rev. D **77**, 123540 (2008), arXiv:0804.0799.
- [22] BOSS Collaboration, MNRAS **470**, 2617 (2017), arXiv:1607.03155.
- [23] A. J. Cuesta, L. Verde, A. Riess, and R. Jimenez, MNRAS **448**, 3463 (2015), arXiv:1411.1094.
- [24] D. M. Scolnic *et al.*, ArXiv e-prints (2017), arXiv:1710.00845.
- [25] M. Betoule *et al.*, A&A **568**, A22 (2014), arXiv:1401.4064.
- [26] D. Scolnic, S. Casertano, A. Riess, A. Rest, E. Schlafly, R. J. Foley, D. Finkbeiner, C. Tang, W. S. Burgett, K. C. Chambers, P. W. Draper, H. Flewelling, K. W. Hodapp, M. E. Huber, N. Kaiser, R. P. Kudritzki, E. A. Magnier, N. Metcalfe, and C. W. Stubbs, Astrophys. J. **815**, 117 (2015), arXiv:1508.05361 [astro-ph.IM].
- [27] A. Shafieloo, A. G. Kim, and E. V. Linder, Phys. Rev. D **85**, 123530 (2012), arXiv:1204.2272 [astro-ph.CO].
- [28] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, PASP **125**, 306 (2013), arXiv:1202.3665 [astro-ph.IM].
- [29] L. Verde, E. Bellini, C. Pigozzo, A. F. Heavens, and R. Jimenez, JCAP **4**, 023 (2017), arXiv:1611.00376.
- [30] G. Hinshaw *et al.*, ApJS **208**, 19 (2013), arXiv:1212.5226.
- [31] A. Gelman, J. Carlin, H. Stern, D. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian Data Analysis, Third Edition*, Chapman & Hall/CRC Texts in Statistical Science (Taylor & Francis, 2013).
- [32] S. Nissanke, D. E. Holz, S. A. Hughes, N. Dalal, and J. L. Sievers, Astrophys. J. **725**, 496 (2010), arXiv:0904.1017 [astro-ph.CO].
- [33] S. Nissanke, D. E. Holz, N. Dalal, S. A. Hughes, J. L. Sievers, and C. M. Hirata, ArXiv e-prints (2013), arXiv:1307.2638 [astro-ph.CO].
- [34] C. Guidorzi *et al.*, ArXiv e-prints (2017), arXiv:1710.06426.
- [35] H.-Y. Chen, M. Fishbach, and D. E. Holz, ArXiv e-prints (2017), arXiv:1712.06531.
- [36] B. Abbott *et al.* (Virgo, LIGO Scientific), Phys. Rev. Lett. **119**, 161101 (2017), arXiv:1710.05832 [gr-qc].

- [37] L. Blanchet, *Living Rev. Rel.* **17**, 2 (2014), arXiv:1310.1528 [gr-qc].
- [38] A. Buonanno, B. Iyer, E. Ochsner, Y. Pan, and B. S. Sathyaprakash, *Phys. Rev.* **D80**, 084043 (2009), arXiv:0907.0700 [gr-qc].
- [39] B. P. Abbott *et al.* (VIRGO, LIGO Scientific), (2013), 10.1007/lrr-2016-1, [*Living Rev. Rel.*19,1(2016)], arXiv:1304.0670v4 [gr-qc].
- [40] J. Veitch *et al.*, *Phys. Rev.* **D91**, 042003 (2015), arXiv:1409.7215 [gr-qc].
- [41] J. Carrick, S. J. Turnbull, G. Lavaux, and M. J. Hudson, *MNRAS* **450**, 317 (2015), arXiv:1504.04627.
- [42] H.-Y. Chen and D. E. Holz, *ArXiv e-prints* (2016), arXiv:1612.01471 [astro-ph.HE].
- [43] D. J. Mortlock, S. M. Feeney, and H. V. Peiris, *ArXiv e-prints*, arXiv:1811.11723 (2018), arXiv:1811.11723 [astro-ph.CO].