Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **546,** 1–24 (2026) https://doi.org/10.1093/mnras/staf2260
Advance Access publication 2025 December 23

1.8 per cent measurement of H_0 from Cepheids alone

Richard Stiskalek ,[1★] Harry Desmond ,[2★] Eleni Tsaprazi ,[3] Alan Heavens ,[3] Guilhem Lavaux ,[4] Stuart McAlpine [5] and Jens Jasche [5]

[1]*Astrophysics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK*
[2]*Institute of Cosmology & Gravitation, University of Portsmouth, Dennis Sciama Building, Portsmouth PO1 3FX, UK*
[3]*Imperial Centre for Inference and Cosmology (ICIC) & Imperial Astrophysics, Department of Physics, Imperial College, Blackett Laboratory, Prince Consort Road, London SW7 2AZ, UK*
[4]*CNRS & Sorbonne Université, Institut d'Astrophysique de Paris (IAP), UMR 7095, 98 bis bd Arago, F-75014 Paris, France*
[5]*The Oskar Klein Centre, Department of Physics, Stockholm University, Albanova University Center, SE-106 91 Stockholm, Sweden*

Accepted 2025 December 18. Received 2025 December 18; in original form 2025 September 15

ABSTRACT

One of the most pressing problems in current cosmology is the cause of the *Hubble* tension. We revisit a *two-rung distance ladder*, composed only of Cepheid periods and magnitudes, anchor distances in the Milky Way, Large Magellanic Cloud, NGC 4258, and host galaxy redshifts. We adopt the SH0ES (Supernovae and H_0 for the Equation of State of dark energy) data for the most up-to-date and carefully vetted measurements, where the Cepheid hosts were selected to harbour also Type Ia supernovae. We introduce two important improvements: a rigorous selection modelling and a state-of-the-art density and peculiar velocity model using `Manticore-Local`, based on the *Bayesian Origin Reconstruction from Galaxies* (BORG) algorithm. We infer $H_0 = 71.7 \pm 1.3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, assuming the Cepheid host sample was selected by supernova magnitudes. However, the actual selection criteria are not clear, and other assumptions can increase H_0 by up to one statistical standard deviation. The posterior has a lower central value and a 45 per cent smaller uncertainty than a previous study using the same distance-ladder data. The result is also slightly lower than the supernova-based SH0ES inferred value of $H_0 = 73.2 \pm 0.9\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, and is in 3.3σ tension with the latest cosmic microwave background results in the standard cosmological model. These results demonstrate that a measurement of H_0 of sufficient precision to weigh in on the *Hubble* tension is achievable using second-rung data alone, underscoring the importance of robust and accurate statistical and velocity-field modelling.

Key words: galaxies: distances and redshifts – cosmological parameters – distance scale.

1 INTRODUCTION

The *Hubble* tension is potentially the most serious challenge faced by the concordance model of cosmology, Λ-cold dark matter (Λ-CDM). This is a claimed 5.8σ discrepancy (L. Breuval et al. 2024) between the present expansion rate of the Universe – the *Hubble* parameter H_0 – inferred from the cosmic microwave background (CMB) by the *Planck* satellite (Planck Collaboration VI 2020; M. Tristram et al. 2024) versus the local distance ladder constructed from Cepheids and Type Ia supernova (SNe) in the *Supernovae and H_0 for the Equation of State of dark energy* programme (SH0ES; A. G. Riess et al. 2022a). A vast array of models have been proposed in response to this, ranging from the addition of a dark energy-like component before recombination to modified gravity in Cepheid stars, yet none is able to resolve the tension while retaining consistency with other observations (for reviews, see E. Di Valentino et al. 2021; N. Schöneberg et al. 2022; E. Di Valentino et al. 2025).

A persistent concern is that the *Hubble* tension may arise not from new physics, but rather from unknown systematics or other modelling deficiencies in the H_0 inference pipelines. This has led to a battery of cross-checks on both the CMB and distance ladder sides, which has not however yielded any definitive conclusion. Alternative CMB data and analysis methods corroborate the *Planck* measurement to high precision (C. L. Bennett et al. 2013; Planck Collaboration LI 2017; S. Aiola et al. 2020; L. Balkenhol et al. 2023; G. Efstathiou, E. Rosenberg & V. Poulin 2024; E. Calabrese et al. 2025; E. Camphuis et al. 2025), a range of systematics have been investigated and found not to impact the inferred H_0 significantly (A. G. Riess et al. 2022a, b, 2023, 2024; A. Bhardwaj et al. 2023; B. Carreres et al. 2025; J. A. Nájera & H. Desmond 2025; E. Tsaprazi & A. F. Heavens 2025), and multiple alternative reconstructions of the local distance ladder – and other low-redshift methods – prefer a higher H_0, albeit with lower precision than SH0ES (e.g. C. R. Burns et al. 2018; D. W. Pesce et al. 2020; J. Schombert, S. McGaugh & F. Lelli 2020; J. P. Blakeslee et al. 2021; S. Dhawan et al. 2022; T. Jaeger et al. 2022; C. Vogl et al. 2025; J. B. Jensen et al. 2025). These studies have consolidated the belief that the *Hubble* tension is real and demands an explanation invoking new physics rather than deficiencies in the data analysis.

★ E-mail: richard.stiskalek@physics.ox.ac.uk (RS); harry.desmond@port.ac.uk (HD)

A key approach for assessing the robustness of the SH0ES inference – and that of the local distance ladder more generally – is to drop or swap out various of the star types used. In particular the SNe, which in SH0ES extend the *Hubble* diagram from $z = 0.0233$ to 0.15, are sometimes considered a weak link: they require standardization based on light curve and environmental properties, complex dust corrections, absolute-magnitude calibration from lower rung distance data and complex simulation-based selection modelling (e.g. R. Wojtak & J. Hjorth 2022, 2024, 2025; C. Gall et al. 2024; G. Efstathiou 2025; T. J. Hoyt et al. 2025). This raises the question of whether H_0 may be inferred (albeit to lower precision) without them, which was in fact the original method of E. Hubble (1929), and also employed by J. A. Willick & P. Batra (2001), who measured H_0 from geometry- and Cepheid-based distances – just as we do here. A discrepancy could then imply systematics in the full SH0ES analysis; as this is the only H_0 analysis that exhibits a statistically significant discrepancy with *Planck*, this has the potential to pull the rug out from under the *Hubble* tension itself.

This approach was taken earlier by W. D. Kenworthy et al. (2022, hereafter K22). This study took 35 galaxies with both Cepheid and SN measurements from the SH0ES project. In SH0ES, the Cepheids are used merely to calibrate the SN absolute magnitude (i.e. the redshifts of the hosts are unused); here, by contrast, the SNe are discarded and H_0 is inferred from the Cepheid distances and host galaxy redshifts. The Cepheid period–luminosity relation (CPLR; H. S. Leavitt & E. C. Pickering 1912) is calibrated from the first, geometric rung of the distance ladder, namely parallax using *Gaia* within the Milky Way (MW; A. G. Riess et al. 2018, 2021), detached eclipsing binaries within the Large Magellanic Cloud (LMC; G. Pietrzyński et al. 2019), and the water maser in NGC 4258 (M. J. Reid, D. W. Pesce & A. G. Riess 2019). At these distances, peculiar velocities are a significant perturbation to the cosmological redshifts and must be modelled. K22 develop and deploy models for these peculiar velocities, as well as possible sample selection effects, to find an overall result $H_0 = 72.9^{+2.4}_{-2.2}\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. This is in agreement with the SH0ES constraint and in mild, 2.3σ tension with *Planck*, leading the authors to conclude that the SNe in SH0ES are not appreciably altering the best-fitting H_0, and therefore are not responsible for the *Hubble* tension.

The present study is a re-analysis of that data and a reassessment of that conclusion. We construct a Bayesian forward model of the distance ladder for inferring H_0 (and a slew of nuisance parameters) from the Cepheid properties, geometric anchor data, and host galaxy redshifts. We present a principled Bayesian framework to account for sample selection and demonstrate its impact on the inferred H_0. In addition, we employ state-of-the-art peculiar velocity modelling with the `Manticore-Local` reconstruction. In comparison with K22, we find a lower central value of H_0, and a much reduced error bar. To understand the differences, we highlight several crucial improvements over their analysis.

The structure of the paper is as follows. Section 2 describes the SH0ES data that we use, which is identical to that of K22. Section 3 presents our method: the inference framework (Section 3.1), peculiar velocity models (Section 3.2), and selection function modelling (Section 3.3). Section 4 details our results: we set the stage with a distance-only analysis not involving redshifts (Section 4.1), before moving onto the H_0 inference and its robustness to peculiar velocity modelling and selection effects (Section 4.2). Section 5.1 discusses the importance of selection function modelling, Section 5.2 provides a detailed comparison with the method of K22, Section 5.3 discusses the broader implications of our results and suggests further work, and Section 6 concludes. Appendix D lists the properties of the Cepheid host galaxies used in our analysis, Appendix A details the mock tests that we develop to demonstrate that our model is unbiased, Appendix B describes the `Manticore-Local` and J. Carrick et al. (2015) peculiar velocity reconstructions and Appendix C summarizes the calculation of the Λ-CDM peculiar velocity covariance.

All logarithms are base-10 unless otherwise stated. We use the notation $\mathcal{N}(x;\mu,\sigma)$ to denote a 1D normal distribution with mean μ and standard deviation σ, evaluated at x; in higher dimensions μ is a vector and σ is replaced by a covariance matrix. We define $h \equiv H_0\left(100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}\right)$.

2 DATA

We use data from the SH0ES programme (A. G. Riess et al. 2022a), which aims to measure the *Hubble* constant H_0 using the local distance ladder. The catalogue combines observations of Cepheids in the MW and nearby galaxies, including hosts with both Cepheids and Type Ia SNe.

The CPLR is anchored by MW parallaxes (A. G. Riess et al. 2018, 2021), detached eclipsing binaries in the LMC (G. Pietrzyński et al. 2019), and megamasers in NGC 4258 (M. J. Reid et al. 2019). M. J. Reid et al. (2019) measure a distance modulus to NGC 4258 of 29.398 ± 0.032 mag, while G. Pietrzyński et al. (2019) determine a distance modulus to the LMC of 18.477 ± 0.026 mag; both are adopted by SH0ES as geometric anchors for the Cepheid calibration.[1] SH0ES also incorporates 55 Cepheids in M31 observed with the *Hubble Space Telescope* (*HST*) using the same three-filter system, as presented by S. Li et al. (2021). This sample provides precise measurements that constrain the slope of the period–luminosity relation.

Beyond the MW, LMC, M31, and NGC 4258, the SH0ES sample includes 35 galaxies hosting both Cepheids and Type Ia SNe. An additional 254 Type Ia SNe, observed in galaxies without Cepheids, extend the measurement into the *Hubble* flow, with their calibration tied to the Cepheid-SN hosts. However, in this work, we do not use any SH0ES SN data, but instead focus on a two-rung distance ladder involving only geometric anchors and Cepheids. This has the crucial advantage of affording an SN-independent determination of H_0, helping to check for potential systematics in the SN analysis. We now provide a brief overview of the geometric anchor and Cepheid distance ladder in SH0ES.

The jth Cepheid in the ith host has apparent magnitude

$$m_{W,ij} = \mu_i + M_W + b_W \log\left(P_{ij}/\hat{P}\right) + Z_W[\mathrm{O/H}]_{ij}, \tag{1}$$

where μ_i is the host distance modulus, M_W is the fiducial absolute magnitude of a Cepheid with a period of 10 d and solar metallicity (zero-point), b_W is the period–luminosity slope, P_{ij} is the Cepheid period, $\hat{P} = 10$ d, Z_W is the coefficient of the metallicity correction, and $[\mathrm{O/H}]_{ij}$ is the metallicity at the Cepheid's galactocentric position.

To account for systematic uncertainty in the metallicity, SH0ES propagates the difference between the mean of nine calibrations of strong-line abundance measurements and the scale of M. Pettini & B. E. J. Pagel (2004), which matches well with direct extragalactic stellar abundances (F. Bresolin et al. 2016).

This dependence is encoded in the covariance matrix of the apparent magnitudes, which depends on Z_W, whose fiducial value in A. G. Riess et al. (2022a) is determined from the joint optimization

---

[1] Recently, the Small Magellanic Cloud (SMC) has also been used as a geometric anchor by L. Breuval et al. (2024).

Table 1. External observational constraints entering the Bayesian hierarchical model. The *HST* and *Gaia* MW calibrations constrain the Cepheid absolute magnitude zero-point, while the geometric anchors constrain the distances to the anchor galaxies LMC and NGC 4258. We use these measurements as they have been reported by A. G. Riess et al. (2022a).

Constraint	Value	Reference
HST Milky Way Cepheid zero-point	-5.804 ± 0.082 mag	A. G. Riess et al. (2018)
Gaia Milky Way Cepheid zero-point	-5.903 ± 0.025 mag	A. G. Riess et al. (2021)
LMC distance modulus	18.477 ± 0.0263 mag	G. Pietrzyński et al. (2019)
NGC 4258 distance modulus	29.398 ± 0.032 mag	M. J. Reid et al. (2019)

of the geometric anchor, Cepheid, and SN components of the SH0ES distance ladder. We adopt the same approach and neglect the dependence of the covariance on Z_W during inference, assuming the same covariance matrix with a fiducial Z_W as A. G. Riess et al. (2022a).

An additional contribution to the Cepheid magnitude covariance arises from uncertainties in the sky background estimation due to crowding. This is quantified by injecting artificial stars near each Cepheid at expected magnitudes from trial period–luminosity fits. The discrepancy between input and recovered magnitudes informs a correction to the photometry and yields a background-induced covariance term for Cepheids in the same host galaxy. The resulting uncertainty, which reflects correlated photometric biases among Cepheids in the same host, contributes an error floor of 0.03–0.06 mag per host. We denote the final Cepheid magnitude covariance matrix as $\mathbf{\Sigma}_{\rm Ceph}$, which combines the contributions from the systematic uncertainties in the metallicity scale, the background-induced photometric biases, and the intrinsic Cepheid uncertainties. For more details, see section 2.1 of A. G. Riess et al. (2022a).

The SH0ES analysis incorporates an external calibration of the fiducial Cepheid absolute magnitude using MW Cepheids with trigonometric parallax distances. Two samples are used: eight Cepheids with high-precision *HST* spatial-scan parallaxes (A. G. Riess et al. 2018) and 75 Cepheids with *Gaia* EDR3 parallaxes (A. G. Riess et al. 2021), both with fluxes measured on the same *HST* photometric system and with direct spectroscopic metallicities. These samples provide independent constraints on M_W, denoted $M_W^{\rm HST} = -5.804 \pm 0.082$ and $M_W^{\rm Gaia} = -5.903 \pm 0.025$, derived from *HST* and *Gaia* parallaxes, respectively. In addition to the MW calibration, SH0ES uses Cepheid observations in the LMC to further constrain the luminosity scale. A systematic offset $\Delta_{\rm ZP}$ with uncertainty $\sigma_{\rm grnd} = 0.10$ mag is introduced to account for differences between ground- and space-based photometry in the LMC calibration. The anchor constraints are summarized in Table 1.

Unlike the SH0ES analysis, which does not use the redshifts of Cepheid host galaxies but instead infers H_0 from more distant SNe (A. G. Riess et al. 2022a), we infer H_0 directly from the redshifts of the Cepheid hosts, without relying on SNe. Specifically, we use the observed redshifts of the 37 galaxies that host both Cepheids and Type Ia SNe, excluding the geometric anchors (LMC, M31, and NGC 4258). These redshifts, converted to the CMB frame, are taken from the PANTHEON + sample, which includes these galaxies as part of its SN compilation (D. Brout et al. 2022). Similarly to K22, we exclude the two most distant host galaxies out of the 37, as they were targeted with a different *HST* programme and are therefore subject to a distinct selection function. Our final sample therefore comprises 35 galaxies, whose coordinates are also listed in Appendix D.

All Cepheid host galaxies are confined to within approximately 40 Mpc (or redshift less than 0.011), indicating the presence of some selection effects in constructing the sample. This selection does not arise from incompleteness in SN detection at these distances. Rather, the selection reflects the fact that only a (random) subset of detected SNe with $m_{\rm SN} \lesssim 14$ mag or $cz \lesssim 3300\,{\rm km\,s^{-1}}$ were selected for *HST* Cepheid follow-up observations. We verify that it is a random subset by comparing the SN magnitudes and redshifts of the Cepheid host galaxies to the full PANTHEON + sample. It is not clear whether this selection is in the SN apparent magnitude, Cepheid apparent magnitude, the host galaxy redshift or their combination, given that Cepheid observations were assembled over many years from independent proposals without a unified selection strategy (K22). We discuss our approach to modelling the selection function, which differs from that of K22, in Section 3.3. To model selection in SN apparent magnitude, we use the observed apparent magnitudes of the SNe in the 35 host galaxies, adopting the bias-corrected apparent magnitudes from the SH0ES sample together with the corresponding covariance matrix $\mathbf{\Sigma}_{\rm SN}$.

In Fig. 1, we show the distribution of the SN apparent magnitudes for the 35 host galaxies and the distribution of their host observed redshifts. The left panel compares the apparent magnitudes of the SH0ES SNe to those of the full PANTHEON + sample. We assume that PANTHEON + is complete at such small redshifts ($z < 0.011$) down to $m_{\rm SN} = 14$ mag. The distribution of the 35 Cepheid hosts' SN magnitudes, which are below $m_{\rm SN} = 14$ mag, is consistent with PANTHEON + ($p = 0.94$ from a Kolmogorov–Smirnov test), which implies that it is plausible that the Cepheid host sample has been selected with a hard limiting estimated magnitude of $m_{\rm SN} = 14$ mag. The same trend can be seen in the right-hand side panel, where we compare the redshift distributions between the PANTHEON + and Cepheid host samples (mutually consistent with $p = 0.60$).

Although the real selection criteria may be complicated and not universal, further support for modelling the selection on the basis of apparent SN magnitude comes from the distributions of magnitude and redshift. Had the sample been selected on either the basis of magnitudes or redshifts but not both, a tail would be visible in the other distribution. However, as shown in Fig. 1, the subset of PANTHEON + with SN magnitudes below 14 leads to a sharp truncation in the observed redshift distribution at $3300\,{\rm km\,s^{-1}}$, whereas the subset of PANTHEON + hosts with redshift below $3300\,{\rm km\,s^{-1}}$ shows a small tail in the SN apparent magnitude above 14. We therefore conclude on the basis of current evidence that SN selection is the most likely model. It was also a primary model considered by K22. Both the magnitude and redshift distributions show a small deficit relative to PANTHEON + near the limiting values. However, these deficiencies are not statistically significant if accounting for the expected Poisson uncertainties. Nevertheless, we account for non-trivial selection near

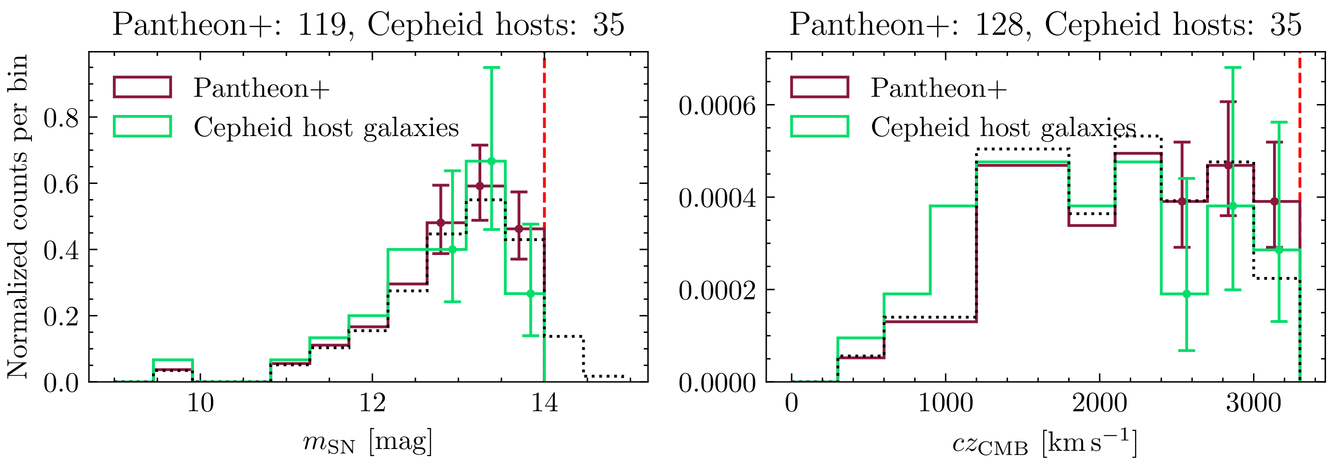

Figure 1. Distribution of SN apparent magnitudes for the 35 Cepheid host galaxies and their observed host redshifts either below a limiting SN magnitude of 14 mag or host observed redshift of $3300\,\mathrm{km\,s^{-1}}$ (left and right panels, respectively). Following a KolmogorSmirnov test, the two samples are mutually consistent with $p = 0.94$ and 0.60. Assuming that the PANTHEON+ sample is complete within $\sim$40 Mpc, this implies random selection of Cepheid hosts below a limiting SN apparent magnitude of 14 and a host observed redshift of $3300\,\mathrm{km\,s^{-1}}$. The error bars are 1σ Poisson counting errors. In addition, the dotted black lines show the magnitude distribution of PANTHEON+ hosts with observed redshifts below $3300\,\mathrm{km\,s^{-1}}$ in the left panel, and the redshift distribution of hosts with magnitudes below 14 in the right panel. Within PANTHEON+, a truncation in magnitude produces a sharply truncated redshift distribution, whereas a truncation in redshift yields a small tail toward higher magnitudes. We therefore regard SN magnitude selection as marginally more likely and adopt it as our fiducial scenario.

the edge by modelling a smooth detection probability function rather than assuming a sharp selection threshold.

3 METHODOLOGY

3.1 Bayesian forward modelling framework

We now describe our Bayesian hierarchical model. Our approach differs from that of K22 in that, rather than using pre-computed distances from the fiducial SH0ES analysis (A. G. Riess et al. 2022a) to predict the observed redshifts of the Cepheid host galaxies in a two-step inference, we forward-model the Cepheid observables directly in a self-contained single-step inference. The data consist of Cepheid magnitudes, periods, and metallicities, and the estimated redshifts of their host galaxies. We also include as observational inputs the MW calibration of the Cepheid absolute magnitude, and the geometric distance moduli of the LMC and NGC 4258. Under the assumption of the SN magnitude selection of the host galaxy sample, we also include these magnitude as inputs to our model, as detailed in Section 3.3.4. The model parameters for this two-rung distance ladder are the Cepheid absolute magnitude M_W, the period–luminosity slope b_W, the metallicity dependence Z_W, and the *Hubble* constant H_0 as well as the standardized SN magnitude M_B in the case that we model SN selection. Each host galaxy is assigned a latent distance modulus μ_i, providing another 35 parameters to infer. Additional free parameters relate to the peculiar velocity field, for which we consider several models as described in Section 3.2. In Section 3.3, we describe our approach to modelling selection effects in either SN magnitude, Cepheid magnitude, or host galaxy redshift. We illustrate the model with a directed acyclic graph in Fig. 2 and summarize the prior distributions of the model's free parameters in Table 2.

We sample M_W, b_W, and Z_W from broad uniform priors. Similarly, H_0 is sampled uniformly from 10 to $100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. The MW Cepheid calibration is imposed through two Gaussian likelihood terms, corresponding to the *HST* and *Gaia* geometric calibration of the fiducial absolute magnitude:

$$\mathcal{L}(M_W^{\rm HST} \mid M_W) = \mathcal{N}(M_W^{\rm HST}; M_W, \sigma_{\rm HST}), \tag{2}$$
$$\mathcal{L}(M_W^{\rm Gaia} \mid M_W) = \mathcal{N}(M_W^{\rm Gaia}; M_W, \sigma_{\rm Gaia}), \tag{3}$$

where $\sigma_{\rm HST}$ and $\sigma_{\rm Gaia}$ are the uncertainties associated with the *HST* and *Gaia* calibrations, respectively. We impose priors on the distance moduli μ_i, which describe the distances to the Cepheid host galaxies, and are related to luminosity distance $D_{{\rm L},i}$ of the ith host as

$$\mu_i = 5 \log \frac{D_{{\rm L},i}}{\rm Mpc} + 25, \tag{4}$$

which is related to comoving distance r as $D_L = (1 + z_{\rm cosmo})r$, where $z_{\rm cosmo}$ is the cosmological redshift. For distances, we choose a prior that is uniform in volume: along a given line of sight (LOS), the probability of finding a galaxy increases as the square of the distance due to the volume element. Sampling the distance moduli μ_i implies the prior

$$p(\mu_i) = p(r_i) \left| \frac{{\rm d}r}{{\rm d}\mu} \right|_{\mu_i}, \tag{5}$$

where r_i is the physical distance, $p(r_i) \propto r_i^2$ and $|{\rm d}r/{\rm d}\mu|$ is the Jacobian for the transformation from distance to distance modulus. (We later introduce further modification of it, the 'inhomogeneous Malmquist bias', to account for the tendency of galaxies to reside in dense environments.) This prior is important, as we discuss further in Section 4.1. For the LMC and NGC 4258, we treat the geometric distance calibrations as additional observational constraints with Gaussian likelihoods:

$$\mathcal{L}(\tilde{\mu}_k \mid \mu_k) = \mathcal{N}(\tilde{\mu}_k; \mu_k, \sigma_k), \tag{6}$$

where k represents either LMC or NGC 4258, μ_k is the sampled distance modulus, $\tilde{\mu}_k$ is the reported measurement from either G. Pietrzyński et al. (2019) or M. J. Reid et al. (2019), respectively, and σ_k is the corresponding measurement uncertainty. The Cepheid apparent magnitudes are predicted according to equation (1), leading

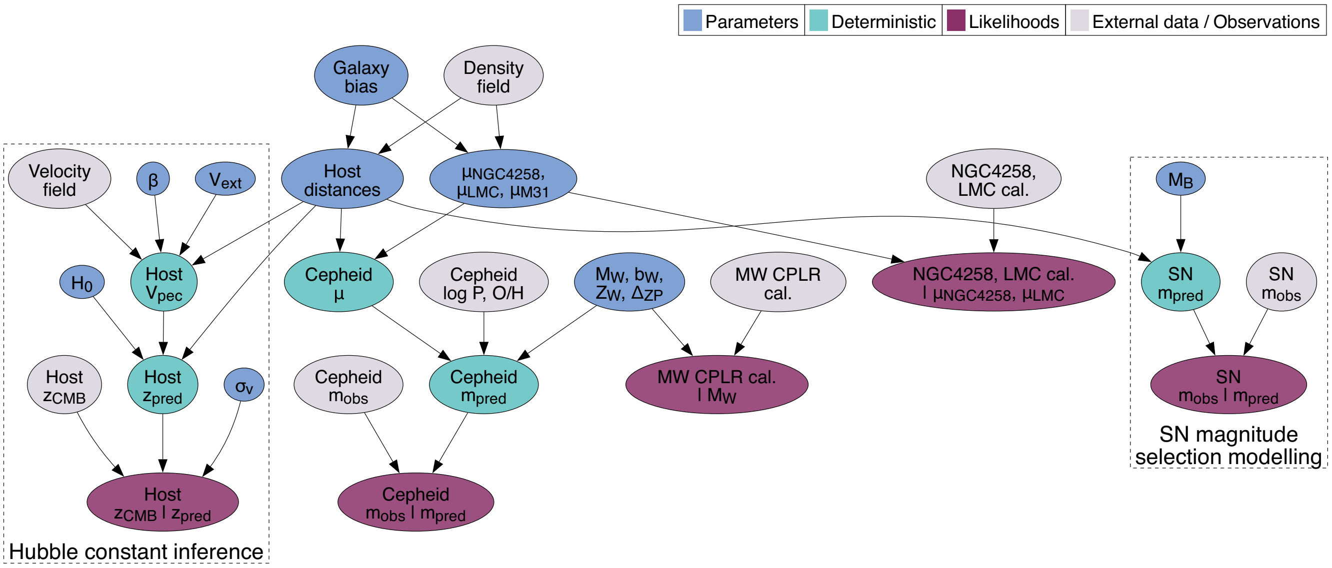

Figure 2. Directed acyclic graph of the probabilistic model used to forward model the Cepheid magnitudes and the host galaxy redshifts. The left-hand dashed black box delineates the portion of the model used to constrain H_0. Excluding it yields the analysis in Section 4.1, where we infer Cepheid host galaxy distances using only the CPLR and the two geometric anchors (LMC and NGC 4258). See Section 3.3 for details on our modelling of sample selection. The dashed black box on the right marks the SN magnitude selection, for which we sample M_B and forward-model the SN apparent magnitudes to model the selection. When assuming redshift selection for the host sample, this step is omitted.

Table 2. Free parameters of the Bayesian hierarchical model described in Section 3, with their corresponding priors. The inclusion of specific peculiar velocity parameters depends on the chosen velocity modelling approach; see Section 3.2. $\mathcal{U}(a, b)$ denotes a uniform prior over the interval $[a, b]$, while $\mathcal{N}(\mu, \sigma)$ denotes a normal prior with mean μ and standard deviation σ. The width of the Δ_{ZP} prior follows A. G. Riess et al. (2022a).

Parameter	Description	Prior
H_0	*Hubble* constant	$\pi(H_0) = \mathcal{U}(10, 100)\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$
Cepheid period–luminosity relation		
M_W	Absolute magnitude zero-point	$\pi(M_W) = \mathcal{U}(-7,\ -5)$
b_W	Period–luminosity slope	$\pi(b_W) = \mathcal{U}(-6,\ 0)$
Z_W	Metallicity correction	$\pi(Z_W) = \mathcal{U}(-2,\ 2)$
Δ_{ZP}	Zero-point offset between ground- and space-based LMC photometry	$\pi(\Delta_{ZP}) = \mathcal{N}(0,\ 0.1)$
Physical distances		
μ_i	Distance moduli of 35 Cepheid host galaxies, LMC, M31 and NGC 4258	$\pi(\mu_i) = \pi(r_i) \left\lvert \frac{\mathrm{d}r}{\mathrm{d}\mu} \right\rvert_{\mu_i}$ with $\pi(r_i) \propto r_i^2$
$\alpha_{low}, \alpha_{high}, \ln \rho_t$	`Manticore-Local` galaxy bias parameters	$\pi(\alpha_{low}) = \mathcal{N}(1, 1)$, $\pi(\alpha_{high}) = \mathcal{N}(0.5, 1)$ (both truncated below at zero) and $\pi(\ln \rho_t) = \mathcal{N}(0, 2)$
Peculiar velocity modelling		
σ_v	Small-scale velocity dispersion	$\pi(\sigma_v) \propto 1/\sigma_v$
$\sigma_{v,\mathrm{low}},\ \sigma_{v,\mathrm{high}},\ \ln \rho_{\sigma_v},\ k_{\sigma_v}$	Density-dependent dispersion parameters of equation (B4) (optional)	$\pi(\sigma_{v,\mathrm{low}}) = \pi(\sigma_{v,\mathrm{high}}) = \mathrm{Maxwell}(\mathrm{scale} = 200\,\mathrm{km\,s^{-1}})$, $\pi(\ln \rho_{\sigma_v}) = \mathcal{N}(1, 5)$, $\pi(k_{\sigma_v}) = \mathcal{N}(1, 1)$ truncated at $k_{\sigma_v} \geq 0$
$\boldsymbol{V}_{ext}$	Constant (external) flow vector	$\pi(\lvert \boldsymbol{V}_{ext} \rvert) = \mathcal{U}(0,\ 1000)\,\mathrm{km\,s^{-1}}$ and uniform in sky direction
A	Scaling of the Λ-CDM velocity covariance matrix	$\pi(A) = \mathcal{N}(1, 0.5)$ truncated at $A \geq 0$
β	J. Carrick et al. (2015) velocity field scaling parameter	$\pi(\beta) = \mathcal{N}(0.43, 0.02)$
Selection function modelling		
M_B	Standardized SN absolute magnitude	$\pi(M_B) = \mathcal{U}(-22, -18)$

to the likelihood

$$\mathcal{L}\left(\boldsymbol{m}_W \mid \boldsymbol{m}_W^{\mathrm{pred}}\right) = \mathcal{N}\left(\boldsymbol{m}_W; \boldsymbol{m}_W^{\mathrm{pred}},\ \boldsymbol{\Sigma}_{\mathrm{Cep}}\right), \tag{7}$$

where the bold font denotes a vector over all 3130 Cepheids.

Up to this point, we have used only the geometric anchors and the CPLR relation to constrain the physical distances to the Cepheid host galaxies. Constraining H_0 requires also including redshift information. The predicted host redshifts, boosted into the CMB frame, are given by

$$1 + z_{\mathrm{CMB}}^{\mathrm{pred}} = (1 + z_{\mathrm{cosmo}})(1 + z_{\mathrm{pec}}), \tag{8}$$

where z_{cosmo} is the cosmological redshift derived from the host distance modulus. We compute the cosmological redshift assuming a flat Λ-CDM cosmology with $\Omega_{\mathrm{m}} = 0.3$ and the sampled value of H_0 (though we note that the dependence on Ω_{m} is negligible because of the small redshift range of our sample). $z_{\mathrm{pec}} = V_{\mathrm{pec}}/c$ is the contribution from the LOS peculiar velocity $V_{\rm pec}$. Assuming Gaussian uncertainty covariance $\mathbf{\Sigma}_{\rm cz}$ of the observed redshifts produces the likelihood term

$$\mathcal{L}(z_{\rm CMB} \mid z_{\rm CMB}^{\rm pred}) = \mathcal{N}(cz_{\rm CMB};\, cz_{\rm CMB}^{\rm pred},\, \mathbf{\Sigma}_{cz}). \tag{9}$$

To sample the posterior distribution we use the NUMPYRO[2] package (E. Bingham et al. 2019; D. Phan, N. Pradhan & M. Jankowiak 2019), specifically the No-U-Turn Sampler method of Hamiltonian Monte Carlo sampling (M. D. Hoffman & A. Gelman 2011). We run 12 independent chains of 6000 samples each, discarding the first 1000 as burn-in. Convergence is ensured by requiring the Gelman–Rubin statistic $\hat{R} - 1 \leq 0.01$ for all parameters.

3.2 Peculiar velocity modelling

Since the Cepheid host galaxies are relatively nearby (the great majority have $cz < 3000\,\mathrm{km\,s^{-1}}$), peculiar velocities can be a significant contributor to their redshifts. It is therefore important that they are modelled reliably, and that the systematic uncertainties they contribute are reliably assessed. We adopt a series of models to achieve this.

(i) Our baseline, least realistic model assumes no coherent flows so that $z_{\rm pec} = 0$ on average. This makes the host galaxy redshifts independent so that the covariance matrix $\mathbf{\Sigma}_{\rm cz}$ is diagonal. We assume that the recession velocity of a given galaxy is Gaussian distributed with a width σ_v, a free parameter to be inferred. The velocity variances are therefore $\sigma_v^2 + \sigma_{cz}^2$, where σ_{cz} is the redshift measurement uncertainty (typically subdominant since $\sigma_v \approx 250\,\mathrm{km\,s^{-1}}$). We sample σ_v from a reference (scale-invariant) prior, such that

$$\pi(\sigma_v) \propto 1/\sigma_v. \tag{10}$$

(ii) This model retains the assumption of a diagonal covariance between predicted and observed redshifts, but explicitly accounts for coherent bulk velocity by modelling the local velocity field as a constant vector $\boldsymbol{V}_{\rm ext}$. We assign a uniform prior on the magnitude of $\boldsymbol{V}_{\rm ext}$ and a uniform prior on its direction over the sky. The LOS peculiar velocity of the ith host is then

$$V_{{\rm pec},i} = \boldsymbol{V}_{\rm ext} \cdot \hat{\boldsymbol{r}}_i, \tag{11}$$

where $\hat{\boldsymbol{r}}_i$ is the LOS unit vector to the ith galaxy. We again include σ_v sampled from a reference prior.

(iii) Rather than explicitly modelling the flow, we use the Λ-CDM peculiar velocity covariance matrix $\mathbf{\Sigma}_{\Lambda\rm CDM}$, thereby marginalizing over all possible Λ -CDM realizations of the local velocity field. The total velocity covariance is then

$$\mathbf{\Sigma}_{cz} = \mathbf{\Sigma}_{\Lambda\rm CDM} + \boldsymbol{I}\ \left(\sigma_v^2 + \sigma_{cz}^2\right), \tag{12}$$

where σ_v captures residual small-scale velocity dispersion and is sampled from the same reference prior as in previous models and $\boldsymbol{I}$ is the unit matrix. The construction of $\mathbf{\Sigma}_{\Lambda\rm CDM}$ is discussed later in this section. We show in Fig. 3 the correlation coefficients between the peculiar velocities of the 35 Cepheid host galaxies. This is the most conservative model.

(iv) Our fourth model extends the previous one by introducing a global scaling parameter A that multiplies $\mathbf{\Sigma}_{\Lambda\rm CDM}$, following D. Huterer, D. L. Shafer & F. Schmidt (2015). This is designed to approximate the effect of deviations from the fiducial cosmology used to compute the covariance matrix. By default, we adopt a truncated Gaussian prior on A, bounded below at zero, with mean unity and standard deviation 0.5. We also test a uniform prior on A over the range [0, 5], verifying that this choice does not appreciably affect the results. The limiting case $A = 1$ recovers the unscaled Λ-CDM model, while $A = 0$ reduces to the σ_v-only model with diagonal covariance.

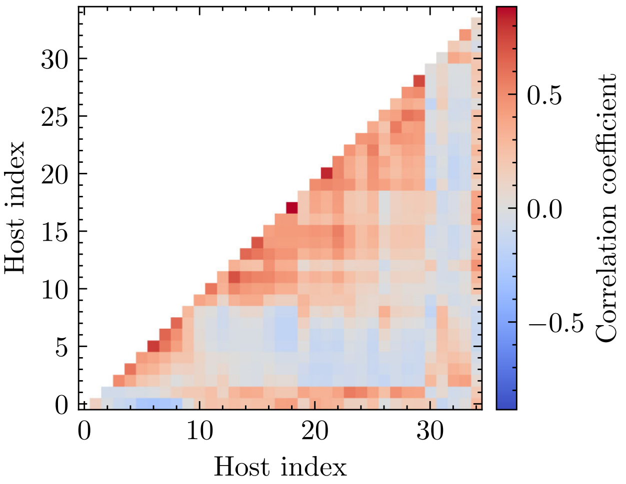

Figure 3. The expected peculiar velocity correlation coefficients computed from the Λ -CDM peculiar velocity covariance matrix (see Appendix C) for the 35 Cepheid host galaxies. A large fraction of the peculiar velocities are strongly correlated, reducing the effective sample size and highlighting the need for a local Universe reconstruction such as J. Carrick et al. (2015).

(v) Our fifth model uses the reconstruction of J. Carrick et al. (2015, hereafter C15), which provides a single realization of both the density and velocity fields of the local Universe as a function of sky position and radial distance in h^{-1} Mpc. The peculiar velocity of the ith host is obtained by evaluating the reconstructed velocity field at its sampled position. In this case, we also extract the local density from the same reconstruction to model inhomogeneous Malmquist bias. The reconstructed peculiar velocities are scaled by a factor β, treated as a free parameter with either a uniform prior or a Gaussian prior based on the measurement $\beta = 0.43 \pm 0.02$ from C15. It is defined as $\beta \equiv f/b_1$, where $f \approx \Omega_{\rm m}^{0.55}$ is the dimensionless growth rate in Λ-CDM (F. R. Bouchet et al. 1995; L. Wang & P. J. Steinhardt 1998) and b_1 is the linear galaxy bias. We further include $\boldsymbol{V}_{\rm ext}$, with uniform prior in both magnitude and direction. Since the reconstruction captures the large-scale velocity field, we assume that residual small-scale motions are described by a diagonal covariance matrix with constant variance $\sigma_v^2 + \sigma_{cz}^2$ as previously. Note that in this case, the explicit modelling of the field-level velocity field renders $\mathbf{\Sigma}_{\Lambda\rm CDM}$ unnecessary; we investigate the effect of a possible residual covariance below the resolution of C15 in Section 5.2.3. The reconstruction of C15 was also employed by K22. Further details of the reconstruction are provided in Appendix B2.

(vi) Our final, fiducial and most sophisticated (and most realistic) model is based on `Manticore-Local` (S. McAlpine et al. 2025), a density and velocity field reconstruction of the local Universe derived from the *Bayesian Origin Reconstruction from Galaxies* (BORG) algorithm (J. Jasche & B. D. Wandelt 2013; G. Lavaux & J. Jasche 2016; J. Jasche & G. Lavaux 2019; S. Stopyra et al. 2024). It is shown in S. McAlpine et al. (2025) and R. Stiskalek et al. (2025) that this is the most accurate among all velocity fields currently in the scientific literature. Unlike C15, `Manticore-Local` provides posterior samples of the initial conditions, thereby quantifying the

---

[2] https://num.pyro.ai/en/latest/

reconstruction uncertainty of the local large-scale structure. We use 30 independent posterior samples, each resimulated at higher resolution with an N-body code. We describe `Manticore-Local` further in Appendix B1. Importantly, we work in real-space rather than redshift-space, thereby avoiding the triple-valued regions that arise when mapping from redshift to real-space distance. While the BORG posterior samples have a resolution of 2.6 h^{-1} Mpc in the initial conditions, the `Manticore-Local` simulations are run at a higher resolution with small-scale modes added below the inference grid scale, over which we then marginalize, effectively accounting for the likely range of smaller-scale density field realizations. The `Manticore-Local` present-day fields extrapolate below the initial condition grid scale to accurately pinpoint positions of galaxy clusters (see fig. 9 of S. McAlpine et al. 2025) Optionally, we also allow σ_v to vary with the local density, following equation (B4), to capture any residual small-scale velocity dispersion associated with galaxy clusters. However, this makes little difference and our fiducial model is that of constant σ_v. As in C15, we evaluate the LOS density and radial velocity, introduce an additional $\boldsymbol{V}_{\rm ext}$ velocity vector, and model all residual velocities with an uncorrelated, constant σ_v term. It is this final model that provides our headline posterior for H_0.

3.3 Selection function modelling

3.3.1 General approach

An important systematic uncertainty is the manner in which the sample was selected, in particular the impact of ‘unobserved’ data (see e.g. A. Gelman et al. 2004; B. C. Kelly 2007; B. C. Kelly, X. Fan & M. Vestergaard 2008; C. Messenger & J. Veitch 2013). Here, we begin more generally with the problem of inference from a flux- or redshift-limited survey, largely following B. C. Kelly et al. (2008). To make things more concrete, this formalism will be used notably to describe the population of Cepheid hosts. It is nonetheless fairly general and we will keep it that way till Section 3.3.2.

Let $\boldsymbol{d}_i$ denote some observed data vector for the ith source, with source parameters $\boldsymbol{\theta}_i$ drawn from a population characterized by $\boldsymbol{\Lambda}$ (e.g. the luminosity function or the standardized SN absolute magnitude). To simplify, we assume that there is no angular selection, except maybe a simple masking operation on the sky. The vector of all angular positions $\hat{u}_i$ is defined as $\hat{\boldsymbol{u}}$. The prior on the angular positioning is thus

$$\pi(\hat{u}) = \frac{M(\hat{u})}{\int M(\hat{u}')\,\mathrm{d}\hat{u}'}, \tag{13}$$

where M is mask function being one if the region is observable and zero otherwise. We define S_i as the detection indicator, where $S_i = 1$ if the ith source is detected and $S_i = 0$ otherwise. If those observations are independent, apart through the population parameter, the likelihood of all data $\boldsymbol{d}$ is a product over the likelihoods of the n detected and $N - n$ undetected objects multiplied by their detection indicators. We must multiply by a binomial coefficient $C_n^N \equiv N!/(n!(N-n)!)$ to account for the number of ways of selecting $N - n$ undetected objects out of N:

$$\mathcal{L}(\boldsymbol{d} \mid \hat{\boldsymbol{u}},\, \boldsymbol{\Lambda},\, N) = C_n^N \prod_{i\in\mathcal{A}_{\rm obs}} p(S_i = 1 \mid \boldsymbol{d}_i,\, \hat{u}_i)\mathcal{L}(\boldsymbol{d}_i \mid \boldsymbol{\Lambda}) \times \prod_{j\in\mathcal{A}_{\rm mis}} p(S_j = 0 \mid \boldsymbol{d}_j,\, \hat{u}_j)\mathcal{L}(\boldsymbol{d}_j \mid \boldsymbol{\Lambda}), \tag{14}$$

where $\mathcal{A}_{\rm obs}$ and $\mathcal{A}_{\rm mis}$ are the set of observed and unobserved sources respectively. N, the total number of galaxies that would have been detected were there no selection, acts as a nuisance parameter in this likelihood. We may multiply by the prior of angular positioning to get the joint likelihood of $\boldsymbol{d}$ and $\hat{\boldsymbol{u}}$:

$$\begin{aligned}&\mathcal{L}(\boldsymbol{d},\, \hat{\boldsymbol{u}} \mid \boldsymbol{\Lambda},\, N) = \\ &C_n^N \prod_{i\in\mathcal{A}_{\rm obs}} p(S_i = 1 \mid \boldsymbol{d}_i,\, \hat{u}_i)\mathcal{L}(\boldsymbol{d}_i \mid \boldsymbol{\Lambda},\, \hat{u}_i)\pi(\hat{u}_i) \times \\ &\prod_{j\in\mathcal{A}_{\rm mis}} p(S_j = 0 \mid \boldsymbol{d}_j,\, \hat{u}_j)\mathcal{L}(\boldsymbol{d}_j \mid \boldsymbol{\Lambda},\, \hat{u}_j)\pi(\hat{u}_j),\end{aligned} \tag{15}$$

We can now integrate equation (14) over the unobserved data, including their unknown angular positioning, to produce the observed data likelihood:

$$\begin{aligned}&\mathcal{L}(\boldsymbol{d}_{\rm obs},\, \hat{\boldsymbol{u}}_{\rm obs} \mid \boldsymbol{\Lambda},\, N) = \\ &C_n^N \prod_{i\in\mathcal{A}_{\rm obs}} p(S_i = 1 \mid \boldsymbol{d}_i)\mathcal{L}(\boldsymbol{d}_i \mid \boldsymbol{\Lambda},\, \hat{u}_i)\pi(\hat{u}_i) \times \\ &\prod_{j\in\mathcal{A}_{\rm mis}} \iint \mathrm{d}\hat{u}_j \mathrm{d}\boldsymbol{d}_j\; p(S_j = 0 \mid \boldsymbol{d}_j,\, \hat{u}_j)\mathcal{L}(\boldsymbol{d}_j \mid \boldsymbol{\Lambda},\, \hat{u}_j)\pi(\hat{u}_j) \\ &\propto C_n^N [p(S = 0 \mid \boldsymbol{\Lambda})]^{N-n} \prod_{i\in\mathcal{A}_{\rm obs}} \mathcal{L}(\boldsymbol{d}_i \mid \boldsymbol{\Lambda},\, \hat{u}_i),\end{aligned} \tag{16}$$

where we have used the fact that

$$\begin{aligned}&p(S = 0 \mid \boldsymbol{\Lambda}) = \\ &\iint \mathrm{d}\hat{u}\mathrm{d}\boldsymbol{d}\; p(S = 0 \mid d,\, \hat{u})\mathcal{L}(\boldsymbol{d} \mid \boldsymbol{\Lambda},\, \hat{u})\pi(\hat{u}),\end{aligned} \tag{17}$$

because all the j indices become irrelevant thanks to double integrations. The joint posterior of $\boldsymbol{\Lambda}$ and N is

$$\begin{aligned}\mathcal{P}(\boldsymbol{\Lambda},\, N \mid \boldsymbol{d}_{\rm obs}) &\propto \pi(\boldsymbol{\Lambda})\pi(N)C_n^N [p(S = 0 \mid \boldsymbol{\Lambda})]^{N-n} \\ &\times \prod_{i\in\mathcal{A}_{\rm obs}} \mathcal{L}(\boldsymbol{d}_i \mid \boldsymbol{\Lambda},\, \hat{u}_i),\end{aligned} \tag{18}$$

where $\pi(\boldsymbol{\Lambda})$ and $\pi(N)$ are the priors on $\boldsymbol{\Lambda}$ and N, respectively. At this point it is useful to marginalize over N to obtain the marginal posterior $\mathcal{P}(\boldsymbol{\Lambda} \mid \boldsymbol{d}_{\rm obs})$. For scale-invariance and computational simplicity, we choose a log-uniform prior on N, $\pi(N) \propto 1/N$. We then obtain the marginal by summing over all possible values of N:

$$\begin{aligned}&\mathcal{P}(\boldsymbol{\Lambda} \mid \boldsymbol{d}_{\rm obs}) \propto \\ &\pi(\boldsymbol{\Lambda}) \left[\prod_{i\in\mathcal{A}_{\rm obs}} \mathcal{L}(d_i \mid \boldsymbol{\Lambda})\right] \sum_{N=n}^{\infty} \frac{C_n^N [p(S = 0 \mid \boldsymbol{\Lambda})]^{N-n}}{N} \\ &\propto \pi(\boldsymbol{\Lambda})\, [p(S = 1 \mid \boldsymbol{\Lambda})]^{-n} \left[\prod_{i\in\mathcal{A}_{\rm obs}} \mathcal{L}(\boldsymbol{d}_i \mid \boldsymbol{\Lambda})\right] \\ &\times \sum_{N=n}^{\infty} C_{n-1}^{N-1}\, [p(S = 0 \mid \boldsymbol{\Lambda})]^{N-n}\, [p(S = 1 \mid \boldsymbol{\Lambda})]^{n}\,,\end{aligned} \tag{19}$$

where the second proportionality follows from multiplying and dividing by $p(S = 1 \mid \boldsymbol{\Lambda})^n$ and utilizing $C_n^N = C_{n-1}^{N-1}(N/n)$. The probability $p(S = 1 \mid \boldsymbol{\Lambda})$ follows from its definition: $p(S = 1 \mid \boldsymbol{\Lambda}) = 1 - p(S = 0 \mid \boldsymbol{\Lambda})$. We derive an expression herein below. The sum in the second proportionality is exactly the expression for the negative binomial distribution as a function of N, which must equal unity by conservation of probability when all possible values are summed over. This produces the final result

$$\mathcal{P}(\boldsymbol{\Lambda} \mid \boldsymbol{d}_{\rm obs}) \propto \pi(\boldsymbol{\Lambda})[p(S = 1 \mid \boldsymbol{\Lambda})]^{-n} \prod_{i\in\mathcal{A}_{\rm obs}} \mathcal{L}(\boldsymbol{d}_i \mid \boldsymbol{\Lambda},\, \hat{u}_i). \tag{20}$$

We note that the second term on the right-hand side must be constant for all data elements. If the posterior on N is desired, it can be determined from the other marginal

$$p(N \mid n,\ \mathbf{\Lambda}) = C_{n-1}^{N-1}[p(S = 1 \mid \mathbf{\Lambda})]^{n}[p(S = 0 \mid \mathbf{\Lambda})]^{N-n}, \tag{21}$$

which can be further marginalized over $\mathbf{\Lambda}$ as

$$p(N \mid n) = \int \mathrm{d}\mathbf{\Lambda}\ p(N \mid n,\ \mathbf{\Lambda})\pi(\mathbf{\Lambda}), \tag{22}$$

or equivalently estimated directly from the posterior samples of $\mathbf{\Lambda}$.

We use the marginalized posterior of equation (20), where the selection is entirely encapsulated through the $p(S = 1 \mid \mathbf{\Lambda})$ term, which represents the fraction of detected samples from the total population. It is calculated as

$$\begin{aligned} p(S = 1 \mid \mathbf{\Lambda}) = \iint & \mathrm{d}\hat{u}\ \mathrm{d}\boldsymbol{d}_{\mathrm{pred}}\ \mathrm{d}\boldsymbol{\theta}\ p(S = 1 \mid \boldsymbol{d}_{\mathrm{pred}},\ \hat{u}) \\ & \times \mathcal{L}(\boldsymbol{d}_{\mathrm{pred}} \mid \boldsymbol{\theta},\ \mathbf{\Lambda})\pi(\boldsymbol{\theta} \mid \mathbf{\Lambda},\ \hat{u})\pi(\hat{u}), \end{aligned} \tag{23}$$

where $\boldsymbol{d}_{\mathrm{pred}}$ denotes the predicted vector data (e.g. the redshift or apparent magnitude) and θ the unknown parameters of the source (e.g. distance). The first term in the integrand is the detection indicator, expressed as a function of $\boldsymbol{d}_{\mathrm{pred}}$ and the angular position $\hat{u}$. The second term is the likelihood of the $\boldsymbol{d}_{\mathrm{pred}}$, the third term is the prior on the source parameters, and the fourth term the prior on the angular position of sources in the sky. We note that this probability is now independent of sources through the double integration over angle and distance.

For a catalogue with a hard detection limit d_{lim} (e.g. a flux or redshift limit), the detection indicator is given by

$$p(S = 1 \mid d_{\mathrm{pred}},\ \mathbf{\Lambda}) = \begin{cases} 1 & \text{if} \quad d_{\mathrm{pred}} < d_{\mathrm{lim}}, \\ 0 & \text{if} \quad d_{\mathrm{pred}} \geq d_{\mathrm{lim}}, \end{cases} \tag{24}$$

Furthermore, if we have a Gaussian likelihood for the data with uncertainty σ_d, and no angular selection, then the probability of detection in equation (23) simplifies to

$$p(S = 1 \mid \mathbf{\Lambda}) = \iint \mathrm{d}\hat{u}\mathrm{d}\boldsymbol{\theta}\ \Phi\left(\frac{d_{\mathrm{lim}} - d_{\mathrm{pred}}}{\sigma_d}\right)\pi(\boldsymbol{\theta} \mid \mathbf{\Lambda},\ \hat{u})\pi(\hat{u}), \tag{25}$$

where d_{pred} is a function of $\boldsymbol{\theta}$ and $\mathbf{\Lambda}$, and $\Phi(x)$ is the cumulative distribution function (CDF) of the standard normal distribution, defined as

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}\mathrm{d}t. \tag{26}$$

Moreover, if in equation (25) the predicted data or the prior on $\boldsymbol{\theta}$ were independent of $\hat{u}$, the integral over $\hat{u}$ would reduce to unity. However, in practice we model a smooth probability of selection as

$$p(S = 1 \mid d_{\mathrm{obs}},\ \mathbf{\Lambda}) = \Phi\left(\frac{d_{\mathrm{lim}} - d_{\mathrm{obs}}}{\sigma_{\mathrm{sel}}}\right), \tag{27}$$

where d_{lim} is the truncation point and σ_{sel} sets the smoothness of the transition. For $d_{\mathrm{pred}} \ll d_{\mathrm{cut}}$ the probability tends to unity, while for $d_{\mathrm{pred}} \gg d_{\mathrm{lim}}$ it falls to zero. If $\sigma_{\mathrm{sel}} \to 0$, then this would reduce to equation (24). Within equation (23), the integral of a Gaussian CDF with a Gaussian density (assuming a Gaussian likelihood) evaluates to another CDF:

$$\begin{aligned} \int \mathrm{d}d_{\mathrm{obs}}\ \Phi\left(\frac{d_{\mathrm{lim}} - d_{\mathrm{obs}}}{\sigma_{\mathrm{sel}}}\right)\mathcal{N}(d_{\mathrm{obs}};\ d_{\mathrm{pred}},\ \sigma_d) = \\ = \Phi\left(\frac{d_{\mathrm{lim}} - d_{\mathrm{pred}}}{\sqrt{\sigma_{\mathrm{sel}}^2 + \sigma_d^2}}\right). \end{aligned} \tag{28}$$

As expected, if σ_{sel} is much smaller than σ_d, then the smooth selection approaches the expression in equation (25). Below, we consider the regime where they are comparable.

We now calculate the selection probability separately for SN magnitude-selected and host-galaxy redshift-selected samples. We verify on mock data that the inference of H_0 is unbiased when modelling selection in this way, as detailed in Appendix A.

3.3.2 *Supernova magnitude selection*

For SN apparent magnitude selection, we have that

$$p(S = 1 \mid M_B) = \iint \mathrm{d}\hat{u}\mathrm{d}r\ \Phi\left(\frac{m_{\mathrm{SN}}^{\mathrm{lim}} - m_{\mathrm{SN}}^{\mathrm{pred}}}{\sqrt{\tilde{\sigma}_{\mathrm{SN}}^2 + \sigma_{\mathrm{SN}}^2}}\right)\pi(r \mid \hat{u})\pi(\hat{u}), \tag{29}$$

where $m_{\mathrm{SN}}^{\mathrm{lim}}$ is the SN apparent magnitude threshold,

$$m_{\mathrm{SN}}^{\mathrm{pred}} = \mu(r) + M_B \tag{30}$$

is the predicted apparent magnitude of an SN at distance r with absolute magnitude M_B, $\tilde{\sigma}_{\mathrm{SN}}$ is the selection smoothness term, and σ_{SN} is the uncertainty in SN apparent magnitude. We set $\tilde{\sigma}_{\mathrm{SN}}$ to be 0.15 mag and σ_{SN} to the average square root of the diagonal of the SN covariance (approximately 0.13 mag), though the inference is not very sensitive to either choice.

Without considering the inhomogeneous Malmquist bias, the prior on distance is given by $\pi(r) \propto r^2$ and is independent of $\hat{u}$. At small physical distances, the distance modulus is related to distance (expressed in Mpc) as $\mu \approx 5\log(r/1\ \mathrm{Mpc}) + 25$, and the detection probability simplifies to

$$p(S = 1 \mid M_B) \propto 10^{-3M_B/5}. \tag{31}$$

This result is independent of σ_{SN}. In the expression above, we wrote the dependence as $10^{-3M_B/5}$. More precisely, the relevant factor is $10^{3(M_B - m_{\mathrm{SN}}^{\mathrm{lim}})/5}$, which depends on the difference between M_B and the limiting magnitude $m_{\mathrm{SN}}^{\mathrm{lim}}$. Since the posterior dependence on $m_{\mathrm{SN}}^{\mathrm{lim}}$ is separable, it can be marginalized out and absorbed into the overall proportionality constant together with σ_{SN}. We can also consider the case when the selection is a smooth function of the magnitude. Going back to equation (23), we have that

$$\begin{aligned} & p(S = 1 \mid M_B) = \\ & \int \mathrm{d}r\ \mathrm{d}m_{\mathrm{SN}}^{\mathrm{pred}}\ p(S = 1 \mid m_{\mathrm{SN}}^{\mathrm{pred}})\mathcal{L}(m_{\mathrm{SN}}^{\mathrm{pred}} \mid M_B,\ r)\pi(r). \end{aligned} \tag{32}$$

Through a similar change of variables, we can pull out the dependence on M_B outside of the integral, so that $p(S = 1 \mid M_B) \propto 10^{-3M_B/5}$ and the shape of the selection function only changes the normalization constant such that the posterior is independent of it.

However, because the inhomogeneous Malmquist bias makes the distance prior dependent on $\hat{u}$, the integral over $\hat{u}$ no longer vanishes and must be evaluated explicitly. We model the galaxy bias using the LOS density from the reconstruction and marginalize by drawing 1000 random LOS, computing the selection probability along each, and averaging the results. We verify that this number of random LOS is sufficient to ensure independence of the results on the sampling.

When accounting for the selection in SN apparent magnitude, the inference is no longer independent of the SNe and depends explicitly on M_B. The intrinsically brighter the SNe, the more of them pass the selection threshold and are observed at greater distances. This behaviour holds regardless of SNe being observed at much greater distances than those considered here; only the host galaxies included in the sample of 35 matter. Moreover, their absolute magnitudes must be modelled self-consistently with the inferred distances to the Cepheid host galaxies and with the SN apparent magnitudes used in the selection. We treat M_B as a free parameter and sample it explicitly, and therefore we also forward-model the apparent magnitudes of SNe in Cepheid host galaxies with the likelihood term to constrain it,

$$\mathcal{L}\left(\boldsymbol{m}_{\mathrm{SN}} \mid \boldsymbol{m}_{\mathrm{SN}}^{\mathrm{pred}}\right) = \mathcal{N}\left(\boldsymbol{m}_{\mathrm{SN}}; \boldsymbol{m}_{\mathrm{SN}}^{\mathrm{pred}}, \boldsymbol{\Sigma}_{\mathrm{SN}}\right), \tag{33}$$

where $\boldsymbol{m}_{\mathrm{SN}}$ are the bias-corrected SN apparent magnitudes from the SH0ES sample, $\boldsymbol{m}_{\mathrm{SN}}^{\mathrm{pred}}$ are the predicted SN apparent magnitudes, and $\boldsymbol{\Sigma}_{\mathrm{SN}}$ is the covariance matrix of the SN apparent magnitudes. Some of the 35 Cepheid host galaxies contain multiple SNe; however, in such cases we select only the brightest SN per host. In deriving equation (29), we assumed the SN observations are independent, which is not valid due to correlations introduced by the standardization procedure (i.e. non-zero off-diagonal elements in the covariance matrix). In contrast, the corresponding likelihood in equation (33) uses the covariance matrix. We verify that neither the downsampling of SNe to one per galaxy nor setting all the off-diagonal elements of the SN magnitude covariance matrix to zero affects the H_0 posterior distribution.

3.3.3 *Redshift-limited selection*

For a redshift-limited survey,

$$p(S = 1 \mid H_0) = \iint \mathrm{d}\hat{\boldsymbol{u}}\mathrm{d}r\, \Phi\left(\frac{cz^{\mathrm{lim}} - cz^{\mathrm{pred}}}{\sqrt{\tilde{\sigma}_v^2 + \sigma_v^2}}\right) \pi(r \mid \hat{\boldsymbol{u}}) \pi(\hat{\boldsymbol{u}}), \tag{34}$$

where z^{pred} is the predicted redshift of a galaxy at distance r and sky direction $\hat{\boldsymbol{u}}$, and $\tilde{\sigma}_v$ is the redshift truncation width, assumed to be 300 km s^{-1}, though the inference is not sensitive to this choice. For illustration, if peculiar velocities and inhomogeneous Malmquist bias are neglected (so that the integrand is independent of $\hat{\boldsymbol{u}}$) and σ_v is fixed, then at low redshift, where $cz_{\mathrm{CMB}} \approx H_0 r$, the detection probability for a redshift-selected survey simplifies to

$$p(S = 1 \mid H_0) \propto H_0^{-3}, \tag{35}$$

per host (and these probabilities multiply, yielding an effective H_0^{+105} dependence, which can shift the peak significantly), with the redshift cut-off absorbed into the proportionality constant. In practice, as before, the inhomogeneous Malmquist bias and peculiar velocities introduce a dependence on $\hat{\boldsymbol{u}}$, requiring the selection probability to be marginalized over all plausible LOSs. We follow the same procedure as previously, but now also use the LOS peculiar velocities from the reconstruction.

As with the SN magnitude selection, the detection probability in equation (34) assumes mutually independent source redshifts, corresponding to a diagonal covariance matrix with entries of σ_v^2. This assumption holds for `Manticore-Local` where we take each realization (without covariance) independently and combine them self-consistently, and also for C15 which provides a single realization. We assume any residual correlations to be negligible. However, it fails when adopting the Λ-CDM peculiar velocity covariance matrix. As demonstrated in Fig. 3, the covariance matrix for the 35 host galaxies exhibits significant off-diagonal correlations.

To 'correct' for the loss of information due to correlated samples, we estimate the effective number of samples: given a generic covariance matrix $\boldsymbol{C}$ of dimension $N \times N$ we can calculate the effective rank as the exponential of the Shannon entropy of its normalized eigenvalues,

$$N_{\mathrm{eff}}(\boldsymbol{C}) = \exp\left[-\sum_{i=1}^{N} p_i \log p_i\right], \tag{36}$$

where $p_i = \lambda_i / \sum_{j=1}^{N} \lambda_j$, and λ_i are the eigenvalues of $\boldsymbol{C}$ (O. Roy & M. Vetterli 2007). $N_{\mathrm{eff}}(\boldsymbol{C})$ provides a continuous measure of the matrix's effective dimensionality, with $N_{\mathrm{eff}}(\boldsymbol{C}) = N$ for a diagonal covariance matrix. This produces $N_{\mathrm{eff}} = 21.0$ for the Λ-CDM peculiar velocity covariance matrix for the 35 host galaxies. Accordingly, we downweight the contribution of equation (34) in the model likelihood by a factor of $N_{\mathrm{eff}}/N \approx 0.57$ for that peculiar velocity model. However, we stress that this correction is only approximate. Moreover, no such correction is required for either the SN or Cepheid magnitude selection; the resulting constraints on H_0 are unchanged when their respective covariances are diagonalized in the likelihood.

3.3.4 *Joint supernova magnitude- and redshift-limited selection*

Although not a fiducial model, we also consider the case of selection on *both* SN host magnitude and galaxy redshift. In this case, the selection probability is given by

$$\begin{aligned} &p(S = 1 \mid H_0, M_B) \\ &= \iint \mathrm{d}\hat{\boldsymbol{u}}\mathrm{d}r\, \Phi\left(\frac{m_{\mathrm{SN}}^{\mathrm{lim}} - m_{\mathrm{SN}}^{\mathrm{pred}}}{\sqrt{\tilde{\sigma}_{\mathrm{SN}}^2 + \sigma_{\mathrm{SN}}^2}}\right) \Phi\left(\frac{cz^{\mathrm{lim}} - cz^{\mathrm{pred}}}{\sqrt{\tilde{\sigma}_v^2 + \sigma_v^2}}\right) \\ &\times \pi(r \mid \hat{\boldsymbol{u}}) \pi(\hat{\boldsymbol{u}}), \end{aligned} \tag{37}$$

which is computed analogously to the previous cases, except that we have a product of the CDFs. For this selection, when modelling the peculiar velocities with the Λ-CDM covariance (but crucially not with either `Manticore-Local` or C15), we use the same correction for the number of effective samples as in Section 3.3.3.

4 RESULTS

As a preliminary to our H_0 inference, in Section 4.1 we infer the host galaxy distances without using their redshifts. This enables us to quantify the effect of priors and selection modelling on the inferred distances. Then, in Section 4.2 we add the redshift information to constrain H_0 with the full model (i.e. including the part in the dashed box of Fig. 2). In Table D1, we provide the inferred Cepheid host distances and the peculiar velocities from both C15 and `Manticore-Local`.

4.1 Host distances

We now consider only the Cepheid host distances, calibrated using the CPLR and two geometric anchors (LMC and NGC 4258). We examine four modelling scenarios:

(i) a uniform prior in distance modulus without selection modelling (a baseline reference);
(ii) a uniform prior in volume without selection modelling;
(iii) a uniform prior in volume with modelling of SN magnitude selection;
(iv) a uniform prior in volume with modelling of Cepheid magnitude selection.

The uniform-in-volume prior, reflecting the fact that the Universe is 3D, yields unbiased distance-ladder inference. In contrast, the uniform-in-distance-modulus has no physical motivation and is implicitly adopted in K22 and the SH0ES analysis (A. G. Riess et al. 2022a) where maximum-likelihood parameters are estimated by minimizing a joint χ^2 statistic across the distance ladder including

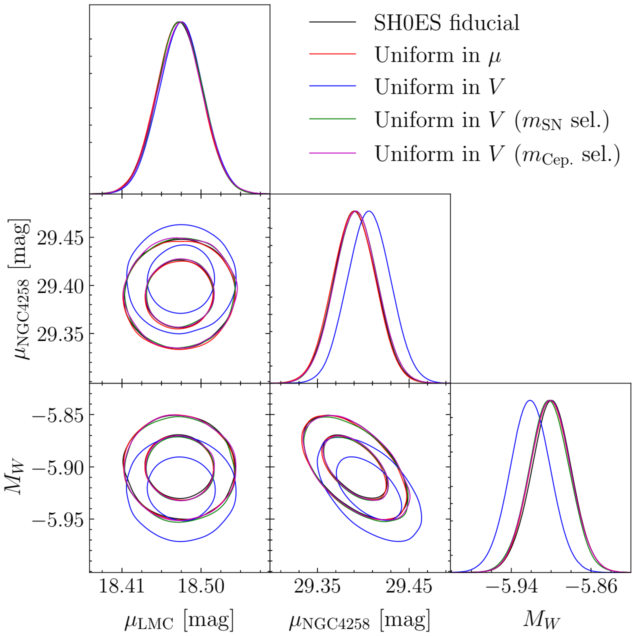

Figure 4. Corner plot of the inferred distance moduli to the LMC, NGC 4258, and the CPLR zero-point calibration, from analyses that do not incorporate redshift information. We compare four scenarios: a uniform-in-volume prior on host galaxy distances without selection modelling (blue), a uniform-in-distance-modulus prior without selection modelling (red), a uniform-in-volume prior with SN magnitude selection modelled (green), and a uniform-in-volume prior with Cepheid magnitude selection (violet). The uniform-in-distance-modulus prior agrees with the SH0ES analysis (A. G. Riess et al. 2022a), which implicitly assumes this prior (black). The uniform-in-volume posteriors with either SN or Cepheid magnitude selection are in excellent agreement with the SH0ES calibration. Contours denote 1σ and 2σ confidence intervals.

the host galaxy distance moduli as free parameters. The uniform-in-volume prior follows $p(r_i) \propto r_i^2$, while the uniform-in-distance-modulus prior corresponds to $p(r_i) \propto 1/r_i$. Since the uniform-in-distance-modulus prior biases the inferred distances low relative to a uniform-in-volume prior, for fixed redshifts it biases H_0 high. This is discussed in more detail and generality in H. Desmond et al. (2025). Fig. 4 shows the inferred distance moduli to the LMC and NGC 4258 along with the Cepheid zero-point M_W. We compare our results to the SH0ES analysis (A. G. Riess et al. 2022a). The no-selection, uniform-in-distance-modulus prior (red contours) produces results in excellent agreement with the SH0ES analysis, as expected. In contrast, the no-selection, uniform-in-volume prior (blue contours) yields larger inferred distances and a brighter Cepheid absolute magnitude zero-point. However, interestingly, incorporating either the SN or Cepheid selection modelling with a uniform-in-volume prior shifts the inferred distances to the LMC and NGC 4258 and Cepheid zero-point back to values identical with those obtained under a uniform-in-distance-modulus prior without selection modelling (see Fig. 4).

We verify that in any case, the inferred Cepheid-zero-point M_W is not in any tension with the MW calibration, which was used as a constraint in the model. Similarly, the inferred distance to the LMC and NGC 4258 remain consistent with the imposed external calibration. Moreover, Fig. 4 also shows that, in the absence of selection modelling, the choice of prior shifts the distance to NGC4258 but leaves the distance to the LMC unchanged. This suggests that the MW Cepheid zero-point calibration, with precise parallax distances, is unaffected by the prior choice.

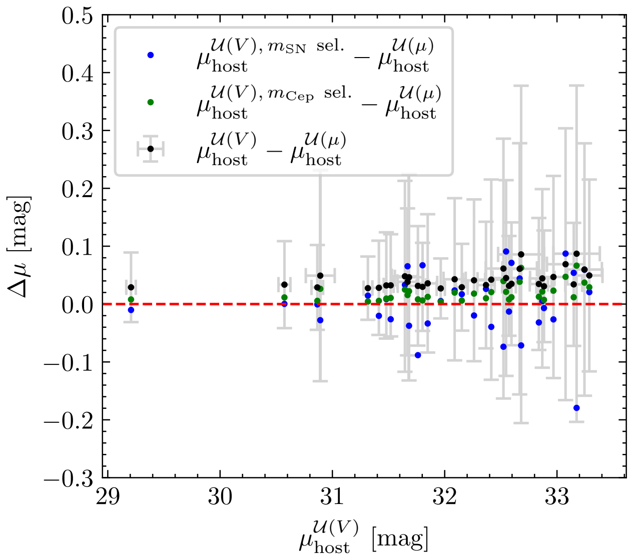

Figure 5. Comparison of distances to the 35 Cepheid host galaxies between a Cepheid-only distance inference using a uniform-in-volume prior without any selection (black), or with SN magnitude selection (blue), or with Cepheid magnitude selection (green) relative to a uniform-in-distance-modulus prior without selection modelling. The first yields an average distance shift of 0.045 mag (which corresponds to H_0 being biased high by $1.79^{+1.00}_{-0.42}$ per cent), the second yields no systematic offset, whereas the third yields an average distance shift of 0.0218 mag. The error bars represent 1σ uncertainties.

In Fig. 5, we extend the comparison of inferred distances to all 35 Cepheid host galaxies. In the absence of selection modelling, we find that the uniform-in-volume prior yields, on average, a distance modulus larger by 0.045 ± 0.15 mag compared to the uniform-in-distance-modulus prior. Thus, using a uniform-in-distance-modulus prior instead of uniform-in-volume potentially biases the inferred value of H_0 high by $1.79^{+1.00}_{-0.42}$ per cent. On the other hand, a uniform-in-volume prior with SN selection modelled yields distance moduli that are not systematically offset from those obtained with a uniform-in-distance-modulus prior without selection modelling; the mean difference is -0.0049 ± 0.13 dex. Finally, the uniform-in-volume prior with Cepheid magnitude selection yields distances that are systematically offset from those inferred with a uniform-in-distance-modulus prior by 0.022 ± 0.15 dex, approximately half the shift observed when selection effects are neglected. From this point onward, we adopt the physically motivated uniform-in-volume prior. We note that in Section 4.2, when inferring H_0 using either `Manticore-Local` or C15, we apply this prior to host galaxy distances and account for inhomogeneous Malmquist bias due to source density fluctuations induced by underlying density fluctuations.

4.2 *Hubble* constant inference

To set the stage for our inference of H_0, and visualize the data, we present in Fig. 6 the *Hubble* diagram showing the relation between the observed redshifts of the 35 host galaxies converted to the CMB frame and their inferred distance moduli under the uniform-in-volume prior and the SN magnitude selection modelling from the previous subsection. The red and blue lines indicate the

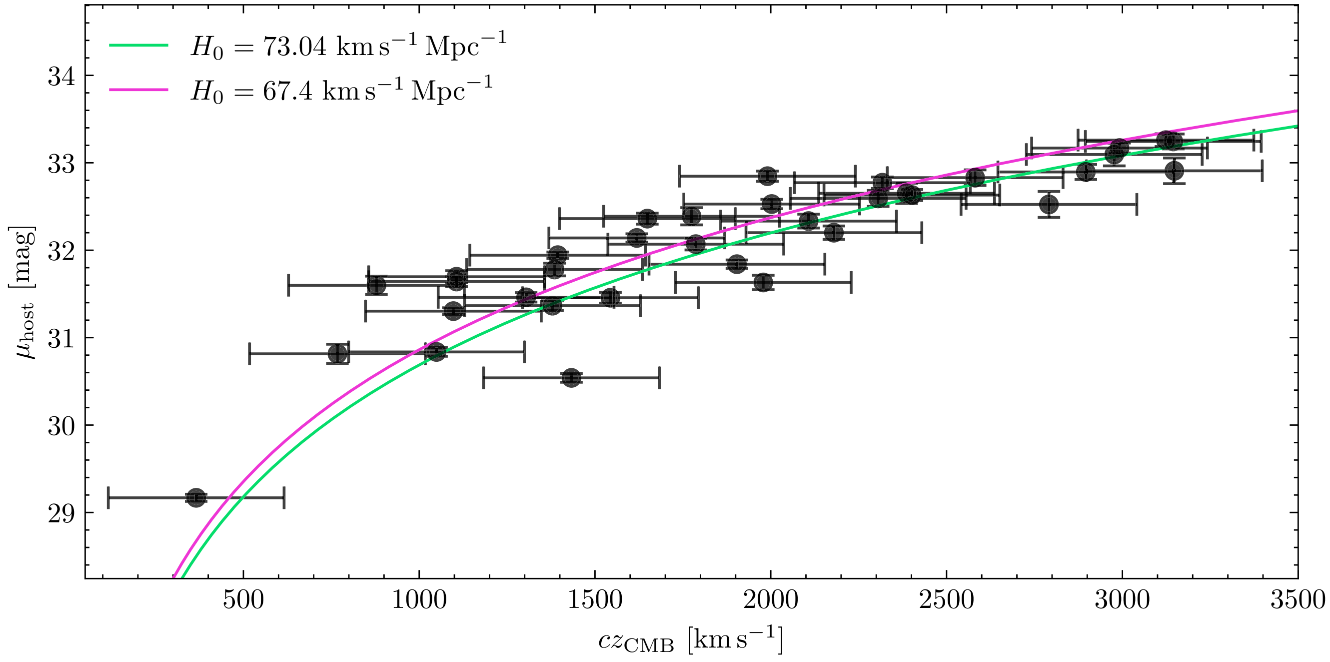

Figure 6. The *Hubble* diagram for the 35 Cepheid host galaxies, showing the relation between CMB-frame redshifts $z_{\rm CMB}$ and the inferred distance moduli from our analysis assuming a uniform-in-volume prior and a selection in SN apparent magnitude. The red curve indicates the predicted distance modulus–redshift relation for the SH0ES best-fitting value of $H_0 = 73.0\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (A. G. Riess et al. 2022a), while the blue curve shows the relation for the *Planck* value $H_0 = 67.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (Planck Collaboration VI 2020). For plotting purposes, we assume a redshift uncertainty corresponding to a velocity dispersion of $250\,\mathrm{km\,s^{-1}}$ though the exact value would depend on peculiar velocity modelling. The error bars represent 1σ uncertainties.

predictions for the best-fitting SH0ES and *Planck* values of H_0, respectively.

We now use the full forward model (see Fig. 2) to infer H_0 as well as the distances by folding in the Cepheid host galaxy redshift information. In Table 3, we tabulate the inferred values of H_0 for various peculiar velocity and selection function models considered in Section 3, all assuming a uniform-in-volume prior on the host galaxy distances, Fig. 7 compares the inferred H_0 for the various selection models while modelling peculiar velocities with `Manticore-Local`. We present a stacked H_0 comparison in Fig. 8.

4.2.1 No selection

As the baseline case we completely omit selection modelling and assume a uniform-in-volume distance prior, although we note that this cannot be a realistic model as the sample is confined within approximately 40 Mpc. Assuming the underlying population of sources is indeed distributed uniformly in volume, this implies selection based on a distance-dependent observable which should be either redshift or apparent magnitude. We find $H_0 = 68.9 \pm 1.6\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ using the C15 reconstruction and $H_0 = 71.1 \pm 1.2\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ with `Manticore-Local`. The values are lower than with any selection modelling, as described below.

4.2.2 Supernova magnitude selection

We now consider in more detail the case where the sample is assumed to be selected by SN apparent magnitude solely (described in Section 3.3.2). As outlined in Section 3.2, our simplest models assume uncorrelated peculiar velocities, either setting them to zero on average, or modelling them with a constant velocity vector $\boldsymbol{V}_{\rm ext}$ that is sampled jointly. The redshift scatter is described by a free parameter σ_v that is also sampled. In these cases, we find $H_0 = 68.8 \pm 2.0$ and $69.3 \pm 1.7\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, respectively. However, these results should be interpreted with caution, as they rely on highly simplified treatments of peculiar velocities. An alternative and more conservative approach adopts the Λ−CDM peculiar velocity covariance matrix (described in Appendix C). This effectively marginalizes over all plausible peculiar velocity realizations consistent with the Λ-CDM power spectrum. It yields $H_0 = 70.1 \pm 3.0$, or $70.1 \pm 3.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ when including a global scaling factor A for the covariance matrix, with it constrained to 1.2 ± 0.3.

We now account for peculiar velocities using local Universe reconstructions. First, we use the C15 reconstruction, with $\boldsymbol{V}_{\rm ext}$ to capture large-scale flows sourced by structure outside the reconstructed volume. Smaller scale motions are treated as uncorrelated and modelled with a velocity dispersion σ_v. As shown in Fig. 7, this model yields $H_0 = 72.1 \pm 1.6\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. When using C15, we find $\sigma_v = 215 \pm 31\,\mathrm{km\,s^{-1}}$, $|\boldsymbol{V}_{\rm ext}| = 199 \pm 82\,\mathrm{km\,s^{-1}}$, and $\beta = 0.42 \pm 0.02$ (sampled with a Gaussian prior with mean 0.43 and standard deviation 0.02, following C15). To account for SN magnitude selection, we infer and marginalize over the standardized SN absolute magnitude M_B. We verify that, when no selection is modelled and only the SN magnitude likelihood of equation (33) is included, we obtain $H_0 = 69.0 \pm 1.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, in good agreement with the no-selection inference without the SNe, which yields $H_0 = 68.9 \pm 1.6\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. Thus the modification to H_0 is driven by the selection modelling, not by the additional SN likelihood.

Second, we use the `Manticore-Local` reconstruction, finding $H_0 = 71.7 \pm 1.3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. In this setup, the velocity field scaling parameter β is fixed to unity (since `Manticore-Local` is

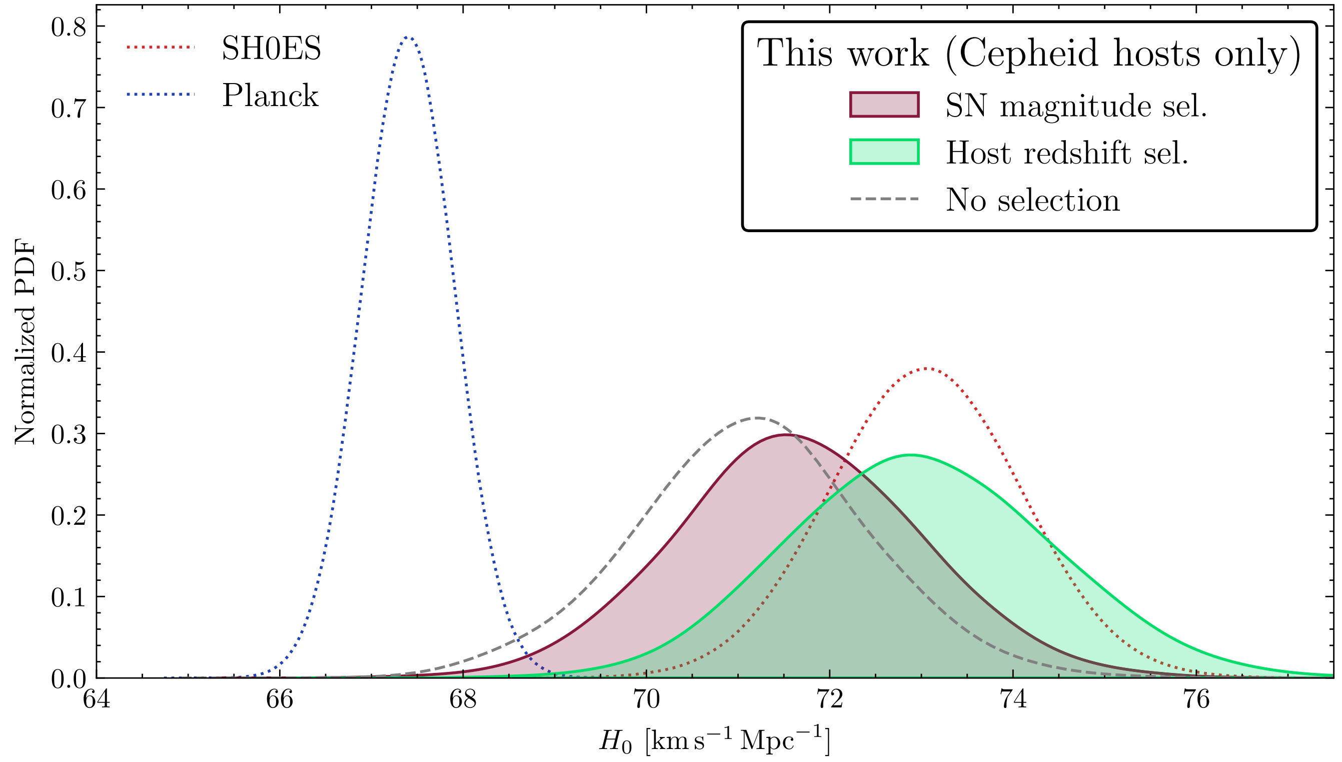

Figure 7. Posterior distributions of H_0 for the two main selection scenarios – SN magnitude selection (purple) and host galaxy redshift selection (green) – when accounting for peculiar velocities using `Manticore-Local` (S. McAlpine et al. 2025), compared to *Planck* ($67.4 \pm 0.5\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$; Planck Collaboration VI 2020) and SH0ES ($73.0 \pm 1.0\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$; A. G. Riess et al. 2022a). For completeness, we also show the H_0 posterior without modelling any selection effects, which leads to a biased inference. These results agree well with SH0ES, though they peak slightly lower except under redshift selection. Model variations are listed in Table 3 and Fig. 8 compares our results using both `Manticore-Local` and J. Carrick et al. (2015) with those from the literature.

Table 3. Inferred values of the *Hubble* constant H_0 in units of $\mathrm{km\,s^{-1}\,Mpc^{-1}}$ for different combinations of peculiar velocity models (rows) and selection models (columns). All results are based on the forward modelling approach described in Section 3.1. See Section 4.2 for details.

Peculiar velocity model	None	Selection model SN mag.	Redshift
No pec. vel., σ_v	67.3 ± 1.9	68.8 ± 2.0	73.9 ± 2.7
Constant inferred flow $V_{\rm ext}$, σ_v	67.9 ± 1.7	69.3 ± 1.7	73.2 ± 2.3
Λ-CDM covariance, σ_v^2	68.3 ± 3.1	70.1 ± 3.0	76.1 ± 3.1
Λ-CDM covariance, scaling A, σ_v^2	68.4 ± 3.3	70.1 ± 3.4	78.5 ± 4.0
J. Carrick et al. (2015), $V_{\rm ext}$, σ_v	68.9 ± 1.6	72.1 ± 1.6	74.4 ± 1.8
`Manticore-Local`, $V_{\rm ext}$, σ_v	71.1 ± 1.2	**71.7 ± 1.3**	73.0 ± 1.4
`Manticore-Local`, $V_{\rm ext}$, $\sigma_v(\delta)$	71.1 ± 1.2	71.6 ± 1.3	72.9 ± 1.4
`Manticore-Local`, $V_{\rm ext}$, σ_v, β	70.9 ± 1.2	71.5 ± 1.3	72.7 ± 1.4
External comparison			
CMB (Planck Collaboration VI 2020)	67.4 ± 0.5		
SH0ES (A. G. Riess et al. 2022a)	73.0 ± 1.0		
SH0ES (with the SMC, L. Breuval et al. 2024)	73.2 ± 0.9		
SH0ES Cepheid-only (W. D. Kenworthy et al. 2022)	72.9 ± 2.3		

Notes. We highlight in bold peculiar velocity modelling with `Manticore-Local` (S. McAlpine et al. 2025), which we consider our fiducial model. The lower block contains literature constraints for external comparison. All uncertainties are 1σ statistical except for that of W. D. Kenworthy et al., which combines statistical and systematic uncertainty.

a dark matter resimulation assuming a fixed cosmology) and we find that $\sigma_v = 145 \pm 30\,\mathrm{km\,s^{-1}}$ and $|V_{\rm ext}| = 223 \pm 64\,\mathrm{km\,s^{-1}}$. In Fig. 9, we show the posterior distribution of all model parameters besides the 35 host galaxy distances. In particular, the posterior on M_B agrees well with the reported value of $M_B = -19.253$ in fig. 14 of A. G. Riess et al. (2022a). Letting σ_v follow the density-dependent form of equation (B4) yields $H_0 = 71.6 \pm 1.3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, almost unchanged from the fiducial value. We also consider the case where β is allowed to vary, imposing a Gaussian prior with mean 1 and standard deviation 0.5, loosely corresponding to varying the

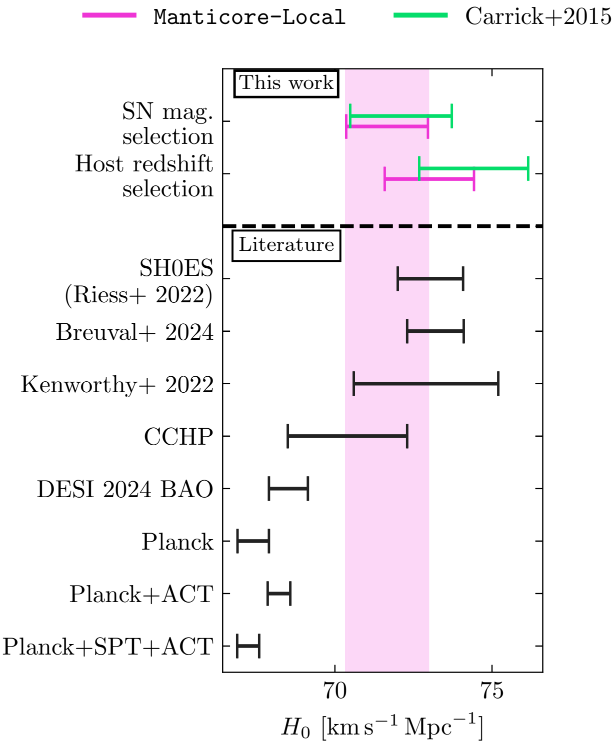

Figure 8. Stacked representation of the inferred H_0 using two local Universe reconstructions to model galaxy bias and account for peculiar velocities: `Manticore-Local` (S. McAlpine et al. 2025) and J. Carrick et al. (2015). Results are shown for the fiducial selection scenario: SN magnitude selection, but for illustration we also show the inferred *Hubble* parameter if selection were made on host redshift, and compared to literature values from Planck Collaboration VI (2020); W. D. Kenworthy et al. (2022); A. G. Riess et al. (2022a); L. Breuval et al. (2024); A. G. Adame et al. (2025); E. Camphuis et al. (2025); W. L. Freedman et al. (2025); and T. Louis et al. (2025), respectively. The pink band carries our fiducial result down the plot and the bands represent 1σ uncertainty.

assumed value of σ_8. This produces $H_0 = 71.5 \pm 1.3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ with $\beta = 0.81 \pm 0.17$, which is in good agreement with the $\beta = 1$ inference and β itself is consistent with unity at $\sim 1\sigma$. Using `Manticore-Local` yields the smallest uncertainty on H_0 due to its most precise (and accurate; S. McAlpine et al. 2025; R. Stiskalek et al. 2025) determination of the peculiar velocity field. The smaller inferred σ_v than C15 also indicates that it more successfully captures local flows.

4.2.3 *Host galaxy redshift selection*

We now consider the case in which the sample of host galaxies is assumed to be selected solely based on observed redshift, as described in Section 3.3.3. Our two simplest models, both using a diagonal covariance matrix with a free parameter σ_v, assume either zero-mean peculiar velocities or model them with a constant velocity vector $\boldsymbol{V}_{\rm ext}$. These yield $H_0 = 73.9 \pm 2.7$ and $73.2 \pm 2.3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, respectively. In contrast, assuming zero-mean peculiar velocities but including the Λ-CDM peculiar velocity covariance matrix, we find $H_0 = 76.1 \pm 3.1$, or $78.5 \pm 4.0\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ when allowing for a global scaling factor A, which is found to be 1.4 ± 0.3. However, these results require an approximate correction to the selection function to account for the effective number of independent samples (see Section 3.3.3), so the results obtained using the Λ-CDM covariance should be interpreted with caution. No such correction is necessary when peculiar velocities are explicitly modelled and the covariance is diagonal.

Then, we again consider our two local Universe reconstruction models. Using the C15 reconstruction, we find $H_0 = 74.4 \pm 1.8\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, with $\beta = 0.42 \pm 0.02$, which is in good agreement with SH0ES. In contrast, using the `Manticore-Local` reconstruction yields $H_0 = 73.0 \pm 1.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. Fig. 9 presents the posterior distribution assuming `Manticore-Local` and redshift-limited selection, showing that the inferred $\boldsymbol{V}_{\rm ext}$ and σ_v are consistent with the SN selection case. Applying the density-dependent σ_v model of equation (B4) gives $H_0 = 72.9 \pm 1.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. If instead of fixing β to unity for `Manticore-Local` we sample it with a Gaussian prior of mean of unity and standard deviation 0.5, we obtain consistent values of $H_0 = 72.7 \pm 1.4\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ and $\beta = 0.87 \pm 0.18$.

4.2.4 *Joint magnitude and redshift selection*

For completeness, we briefly consider joint selection scenarios. If the entire host sample were selected on *both* SN apparent magnitude *and* observed redshift (see Section 3.3.4), then equation (29) is modified so that the integrand includes the selection in both the magnitude and redshift limits (equation 37). Such a joint selection yields H_0 values intermediate between those from SN magnitude-only and host redshift-only selections. Using C15 gives $H_0 = 73.3 \pm 1.7\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ with $\beta = 0.42 \pm 0.02$, while `Manticore-Local` yields $H_0 = 72.3 \pm 1.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. Adopting the density-dependent σ_v of equation (B4) yields $H_0 = 72.1 \pm 1.3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. Finally, allowing β to vary for `Manticore-Local` with a Gaussian prior of $\mathcal{N} = (1, 0.5)$, we find $H_0 = 72.0 \pm 1.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and $\beta = 0.86 \pm 0.18$ indicating no significant effect on H_0 from fixing β to unity, i.e. the assumed cosmology of `Manticore-Local`.

Alternatively, the sample could be interpreted as a concatenation of two subsamples: one selected purely on SN magnitude and the other purely on host redshift. Fig. 10 shows the mean inferred H_0 as a function of the number of galaxies selected by SN magnitude, varying this number from zero to 35. The left bound corresponds to purely redshift-based selection and the right bound to purely SN-magnitude-based selection. We find an approximately linear scaling between these two limiting cases, with an average H_0 uncertainty very close to $\sim 1.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ in all cases.

5 DISCUSSION

In this section, we discuss the impact of selection function modelling on the inferred H_0 (Section 5.1) and then compare our results with those of K22 (Section 5.2).

5.1 Importance of selection function modelling

Typically (or at least ideally), the selection strategy for a survey is well-defined and imposed, for example, on apparent magnitude (i.e. a flux limit), reflecting the detector's sensitivity. However, as the sample of Cepheid host galaxies has been assembled over many years as follow-up to nearby SN hosts from various proposals, it is unclear which selection is most applicable (K22). None the less, given that all galaxies in the sample are relatively nearby, some form of selection must be present. Therefore, we consider two main selection scenarios: an SN apparent magnitude limit and a host galaxy

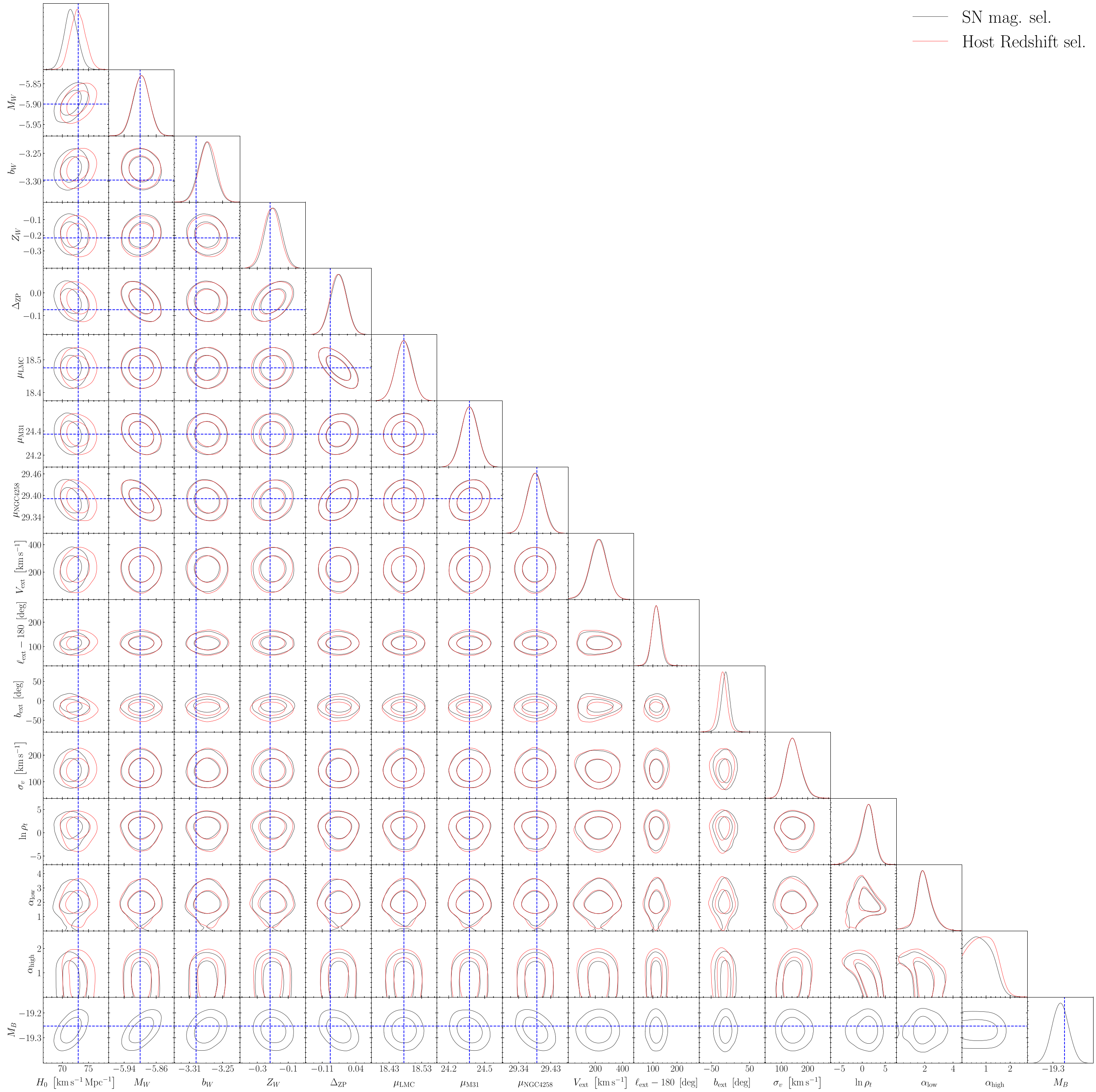

Figure 9. Constraints on all model parameters, except the 35 Cepheid host distances, for our fiducial peculiar velocity model based on the `Manticore-Local` reconstruction (S. McAlpine et al. 2025), assuming either SN magnitude (black) or host redshift selection (red). The blue lines show the best-fitting parameter values from the SH0ES analysis (A. G. Riess et al. 2022a). The contours represent 1σ and 2σ confidence intervals.

redshift limit. K22 considered these same two cases, arguing for an SN-based selection on the grounds that Cepheid hosts are typically targeted for follow-up after an SN is observed, with a preference for brighter (nearer) SNe within which (once selected) Cepheids are always observed.

Given the distributions of SN magnitudes and host redshifts shown in Fig. 1, we consider selection based solely on SN magnitude to be the more plausible scenario. Compared to the PANTHEON + sample, such a selection can reproduce the Cepheid host redshift distribution, whereas selection on redshift alone fails to reproduce the SN magnitude distribution. We briefly explore mixed selection scenarios in Section 3.3.4, finding that they simply interpolate between the two cases. The effect of Cepheid selection may however also need to be considered. This would require modelling not only the Cepheid population (e.g. period and metallicity distributions) but also Cepheids' spatial distribution within host galaxies, which affects detectability, and propagating the resulting non-trivial covariance matrix. We leave this for future work.

If the sample was selected based the SN magnitude, neglecting this selection would bias H_0 low, as the selection preferentially retains

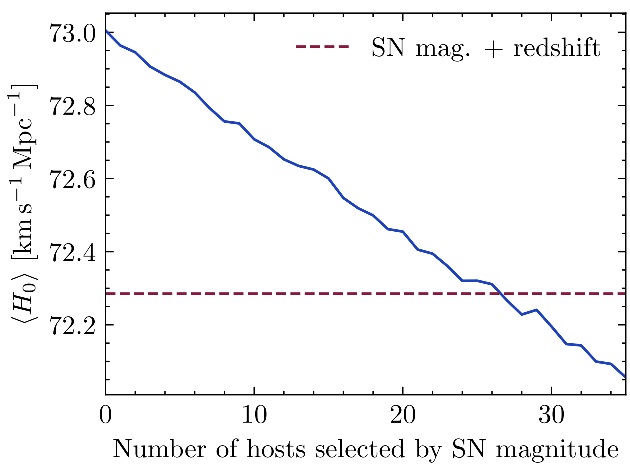

Figure 10. Mean inferred H_0 using `Manticore-Local` as a function of the number of host galaxies (out of 35) assumed to be selected by SN magnitude, with the remainder selected by host redshift. The left bound corresponds to purely redshift-based selection, whilst the right bound corresponds to purely SN magnitude-based selection. The horizontal dashed line shows the result of assuming the entire sample is selected on both SN magnitude and redshift.

nearby sources. If the forward model does not account for this, it may place sources at greater distances which, at fixed redshift, leads to a lower and biased estimate of H_0. Because the SN and Cepheid apparent magnitudes (which are directly used within the forward model) are correlated via the host galaxy distance, selection on SNe magnitude alters the distribution of 'true' source distances retained in the sample, thus necessitating explicit modelling as described in Section 3.3.2 that (weakly) couples the inference to the SNe. Consequently, the SN magnitude selection analysis is no longer independent of SNe; however, we verify that including SN apparent magnitudes in the forward model alone (without modelling their selection) has no effect on the inferred H_0. On the other hand, if selection is assumed to be imposed on either the Cepheid apparent magnitude or host galaxy redshift, then there is no need to account for SN apparent magnitudes, and the analysis remains independent of SN data.

5.2 Comparison to Kenworthy et al. (2022)

While using the exact same 35 host galaxies, our approach has three principal differences to that of K22:

(i) We forward model the Cepheid apparent magnitudes, rather than adopt *a-priori* calibrated distance estimates.

(ii) We employ a principled treatment of selection effects, avoiding arbitrary modifications to the prior on object distances (unless considering an effective selection model).

(iii) We use a state-of-the-art local Universe reconstruction `Manticore-Local`, rather than relying solely on linear-theory reconstructions. This marginalizes over the full posterior rather than assuming a Gaussian covariance.

We now compare the methods in detail, showing that neither the redshift- nor distance-limited selection of K22 is well justified. A distance-based selection cannot be considered in the first place, since selection must act on the level of an observable quantity, which distance is not.

5.2.1 *Redshift selection*

For redshift selection, K22 model the case in which only the host galaxy redshift was effectively used in the selection procedure. It appears that they mean this to imply a prior on the comoving distance $\boldsymbol{\chi}$ of the Cepheid hosts that models homogeneous and inhomogeneous Malmquist bias in the usual way (their equation 15), without any further selection modifications. However, they then write 'Tests on simulated data found that the use of this prior tended to bias the recovered distances and H_0 high', and proceed to modify the prior to:

$$\pi(\boldsymbol{\chi}) = \max(\boldsymbol{\chi})^{\alpha} \prod_i \frac{\chi_i^2}{\int_0^{\max(\chi)} \chi^2 \mathrm{d}\chi}, \tag{38}$$

where $\boldsymbol{\chi}$ is expressed in h^{-1} Mpc and α is set to 60. This entails that:

(i) the individual χ_i^2 terms are normalized at each step to the current maximum inferred host distance, effectively truncating the χ^2 prior at $\max(\boldsymbol{\chi})$;

(ii) the prior is not a valid probability density function, as the $\max(\boldsymbol{\chi})^{\alpha}$ factor breaks the normalization. Moreover, it is not merely an improper prior on distance but it explicitly depends on h, which is a separate (and the crucial) parameter.

To quantify the bias on H_0 that this produces, we consider the relation between priors on χ_i and r_i, where r_i is physical distance in Mpc such that $\chi_i = r_i h$. If $\pi(\chi_i) = 3\chi_i^2/\chi_{\max}^3$, which corresponds to a prior on χ_i that goes as χ_i^2 but truncated at $\chi_{\max}$, the corresponding prior on r_i is

$$\pi(r_i) = \pi(\chi_i) \left| \frac{\mathrm{d}\chi}{\mathrm{d}r} \right|_{r_i} = \frac{3r_i^2}{r_{\max}^3}, \tag{39}$$

where $r_{\max} = \chi_{\max}/h$. Therefore, provided that this term is normalized as it is in K22's equation (17), it does not introduce spurious factors of h. However, the factor of $\max(\boldsymbol{\chi})^{\alpha}$ introduces a problem by modifying the 'prior' on individual distances to

$$\pi(\chi_i) = \chi_{\max}^{\alpha/n} \frac{3\chi_i^2}{\chi_{\max}^3}, \tag{40}$$

where n is the number of Cepheid host galaxies. This distribution is no longer normalized and alters the prior on r in a strange way. Since $\chi_{\max} = r_{\max}h$, this introduces a spurious factor of $(r_{\max}h)^{\alpha/n}$ per host galaxy. Thus, the $\max(\boldsymbol{\chi})^{\alpha}$ term implicitly imposes a prior on H_0 proportional to H_0^{α}, which K22 set to H_0^{60}. The simulations from which they 'derive' this prior include a selection in the observed redshift of $z < 0.011$. We have showed that without modelling peculiar velocities or the inhomogeneous Malmquist bias, the correct method for accounting for a redshift selection introduces a factor of H_0^3 per host galaxy, as discussed in Section 3.3.3. Thus, for a uniform-in-volume prior and a redshift-selected sample of 35 host galaxies, the correct treatment of the selection introduces a factor of H_0^{105}, rather than H_0^{60}.

Adopting C15 as the peculiar velocity model, K22 found $H_0 = 76.4^{+2.6}_{-2.4}$ or $74.1^{+2.4}_{-2.1}$ km s^{-1} Mpc^{-1}, depending on whether the redshift scatter σ_v is fixed or inferred. In contrast, with the same field and inferring σ_v , we find $H_0 = 74.4 \pm 1.8$ km s^{-1} Mpc^{-1}. Although in principle our selection treatment effectively introduces a factor of H_0^{105} rather than H_0^{60}, direct comparison is complicated by our explicit treatment of both peculiar velocities and inhomogeneous Malmquist. While K22 introduced an ad hoc prior modification in an attempt to to minimize bias in redshift-limited mocks (which is ultimately incorrect), we present a principled method for modelling redshift selection, as well as properly accounting for the effects of peculiar velocities and inhomogeneous Malmquist bias on it. We demonstrate that our method is unbiased in Appendix A.

5.2.2 Distance selection

The second case considered by K22 is a distance-limited selection, which is presumably meant to model a selection on the observable SN magnitudes. In this model, they set the prior on the host galaxy distances to

$$\pi(\chi_i | d_T) = \frac{f(\chi_i, d_T)}{\int_0^{80} f(x, d_T) \mathrm{d}x}, \tag{41}$$

where d_T is the distance cutoff in h^{-1} Mpc, the dummy variable x is also in h^{-1} Mpc and

$$f(\chi, d_T) = \chi^2 \, \Phi\left(\frac{\mu(\chi) - \mu(d_T)}{0.15}\right). \tag{42}$$

(The inhomogeneous Malmquist term is also included but we omit it for brevity here.) Φ is the CDF of a standard Gaussian distribution defined in equation (26). However, since the CDF has the limits

$$\Phi(x) \to 0 \quad \text{as } x \to -\infty; \quad \Phi(x) \to 1 \quad \text{as } x \to +\infty, \tag{43}$$

this implies that χ values *below* d_T are cut-off. We assume this is a typo and instead interpret Φ as the survival function, defined as

$$\mathrm{SF}(x) = 1 - \Phi(x), \tag{44}$$

with which we can reproduce the reported results. Using the CDF would force d_T to be less than the distance to the nearest host, since the model cannot assign host galaxies to distances where the prior is forced to zero by the CDF.

To verify this, we implement the model of K22 and their distance selection (described in Section 5.2). Unlike our analysis, this involves predicting some *a priori*-calibrated distance estimates, rather than deriving them simultaneously with the other model parameters by forward-modelling the Cepheid magnitudes. For this, we use the Cepheid-derived distances from our redshift-independent analysis in Section 4.1 with a uniform-in-distance-modulus prior, matching the data and assumptions they used. When working in h^{-1} Mpc as they do and marginalizing over peculiar velocities with the Λ-CDM covariance matrix, we find $H_0 = 71.6 \pm 3.2 \, \mathrm{km\,s^{-1}\,Mpc^{-1}}$, consistent with their reported value of $71.6 \pm 4.6 \, \mathrm{km\,s^{-1}\,Mpc^{-1}}$ as shown in the second row of their table 2. (Their uncertainty is larger because they overestimate the Λ -CDM peculiar velocity covariance; see the last paragraph of Appendix C.) Conversely, if we use the model of K22 in terms of predicting 'observed' distances derived from a prior analysis, but work consistently in Mpc without applying any 'selection' modification to the prior, we again find a lower value of $H_0 = 68.4 \pm 3.0 \, \mathrm{km\,s^{-1}\,Mpc^{-1}}$, consistent with our no-selection results.

However, the approach of K22 does not provide a principled method for modelling selection in apparent magnitude. In Section 3.3.2, we derive the correct treatment: in a Bayesian hierarchical framework, modifying the prior of latent variables is not justified; instead, the appropriate approach is to apply a correction for the unobserved sources, which is independent of the properties of the detected sources.

Although K22 do not consider the case of no selection, it is interesting to see how it would play out when measuring distances in h^{-1} Mpc units as they do. This would correspond to $\pi(\chi_i) \propto \chi_i^2$, which, if treated as an improper prior without normalization (as would appear from their equation 15), would produce a factor of h^3 per galaxy when mapping it to the prior on r_i: a factor of h^2 derives from χ^2, and h from the Jacobian of the transformation from χ to r. This yields a prior scaling as H_0^{105}, which, as we have shown, is effectively the result of redshift selection. We therefore see that if one works in redshift as the independent variable one does not need to explicitly model redshift selection, analogously to how one does not need to explicitly model selection effects for a volume-limited sample when working in physical distances.

While most of the issues we find with K22 are specific to that analysis, we note that the issue of distance priors are pervasive to distance-ladder studies. In general, adopting a uniform-in-distance-modulus or uniform-in-distance prior (r^0 or r^{-1}) puts galaxies at smaller distances relative to the physically motivated uniform-in-volume prior (r^2), biasing H_0 high. This may be contributing to a community-wide overestimation of H_0 (H. Desmond et al. 2025), although it is complicated by selection effects.

5.2.3 Modelling the peculiar velocity covariance

We effectively considered three discrete velocity models: `Manticore-Local`, the C15 field (both with additional variance), and the Λ-CDM covariance without a mean field. K22 add an additional covariance to C15 describing contributions to the velocity field at small enough scales not captured by C15 yet large enough to couple galaxies, i.e. not simply a diagonal σ_v contribution. They find that the inclusion of this covariance term yields a relatively large shift in H_0, pushing it higher. To derive this additional covariance, K22 use the reconstruction of R. Lilow & A. Nusser (2021), which is a linear reconstruction based on the 2MASS Redshift Survey (2MRS) galaxy density field (J. P. Huchra et al. 2012) within a $200 \, h^{-1}$ Mpc radius. We do not adopt this approach: a covariance of unmodelled peculiar velocities cannot be reliably constructed from two discrete fields. However, it can be estimated from `Manticore-Local`, which provides self-consistent posterior samples of the large-scale structure, and accounted for directly by marginalizing over field realizations, as we do. Moreover, this approach ensures a self-consistent treatment of the peculiar velocity correlations between host galaxies.

We also investigate directly whether the covariance between galaxies on scales smaller than those resolved by C15 is significant: we consider the contribution to the Λ-CDM peculiar velocity covariance matrix from scales not captured by C15. The effective smoothing applied to the galaxy density field by C15 is $4 \, h^{-1}$ Mpc, which corresponds to a wavenumber $k \approx 1.6 \, h \, \mathrm{Mpc}^{-1}$. (K22 apply an additional Gaussian smoothing of $3 \, h^{-1}$ Mpc.) To be conservative we adopt a cut-off of $k > 0.5 \, h \, \mathrm{Mpc}^{-1}$ (corresponding to a physical scale of $12.6 \, h^{-1}$ Mpc) and set the power spectrum to zero below this limit when computing the Λ-CDM peculiar velocity covariance matrix. Fig. 11 shows the resulting correlation matrix for all 35 Cepheid host galaxies. Compared to Fig. 3 (computed using the full power spectrum), the correlations are significantly reduced, since it is the large-scale flows that are primarily responsible for coupling the velocities. Of the 595 possible pairs of galaxies, only 7 have an absolute expected Λ-CDM peculiar velocity correlation exceeding 0.1, and only 57 exceed 0.02. In comparison, for the full covariance matrix these numbers were 350 and 534, respectively.

To quantify this effect, we infer H_0 using the same model as in Section 4.2, but compare two cases (each assuming SN magnitude selection). In the fiducial analysis, when peculiar velocities are accounted for using the C15 reconstruction we assume the covariance matrix to be diagonal with variance σ_v^2. In the second

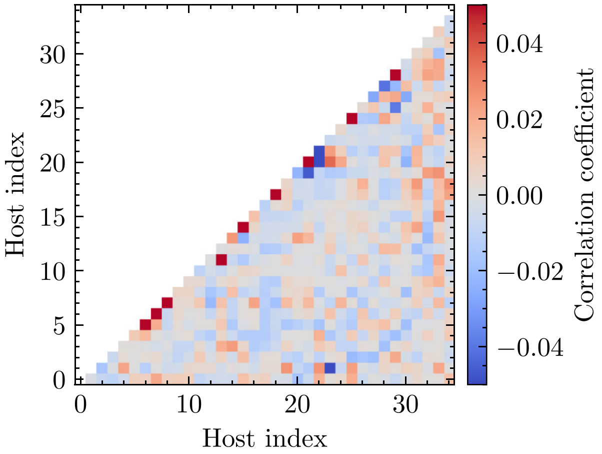

Figure 11. Correlation matrix of the Λ -CDM peculiar velocity covariance computed with a non-linear power spectrum including only modes with $k > 0.5\,h\,\mathrm{Mpc}^{-1}$, conservatively modelling scales not reconstructed by J. Carrick et al. (2015). The matrix is shown for the 35 Cepheid host galaxies. Excluding large-scale contributions significantly reduces the correlations between hosts compared to Fig. 3, where all scales were included. The resulting correlation coefficients are negligible, indicating that scales below the reconstruction limit of J. Carrick et al. (2015) can be well approximated by a diagonal covariance matrix as in our fiducial model.

case, we additionally include the Λ-CDM covariance matrix, which accounts for contributions from all scales with $k > 0.5\,h\,\mathrm{Mpc}^{-1}$. This fiducial analysis yields $H_0 = 72.1 \pm 1.6$, while the latter yields $72.1 \pm 1.6\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. This demonstrates that the off-diagonal contributions are negligible for scales not captured by C15, justifying our fiducial model where only a diagonal contribution (σ_v) is added to that. A similar conclusion holds for the other selection models. This disagrees with the finding of K22, which we attribute to them modelling the peculiar velocity covariance from the two distinct fields of C15 and R. Lilow & A. Nusser (2021).

Moreover, we use `Manticore-Local` which has a spatial resolution of $0.7\,h^{-1}\,\mathrm{Mpc}$ – hence self-consistently modelling scales smaller than those resolved by C15 – and provides an ensemble of plausible realizations of the local peculiar velocity field describing the reduced cosmic variance given the BORG constraints. However, in all cases we find that `Manticore-Local` and C15 agree to about 1σ. In S. McAlpine et al. (2025) and R. Stiskalek et al. (2025), we benchmarked `Manticore-Local`, C15 and other reconstructions against direct distance data, finding `Manticore-Local` to be the most accurate model of the local Universe to date, with the C15 reconstruction also performing well. Compared to C15, the primary advantages of `Manticore-Local` are its more accurate gravity and galaxy bias models which allow the modelling of the density and velocity field at a higher resolution by going beyond linear theory, and its independent posterior samples of the large-scale structure, over which we marginalize.

The higher fidelity of the `Manticore-Local` reconstruction is particularly important for host galaxies located within or near galaxy clusters. In R. Stiskalek et al. (2025), we showed that BORG-based models provide significantly improved modelling of galaxies in overdense regions compared to linear modelling. This is particularly relevant here because at such low redshifts, peculiar velocities are the dominant source of uncertainty, and the local dynamics is dominated by infall to nearby clusters, most notably Virgo, Fornax, and Leo. Indeed at least 7 of the 35 host galaxies lie within the infall region of these clusters (this number is approximate as we do not carefully test for cluster membership). In contrast, C15 provides only a single mean field estimate.

5.3 Ramifications of the results and future work

We have shown that a Cepheid-only distance ladder yields H_0 values consistent with SH0ES (A. G. Riess et al. 2022a), though using the state-of-the-art `Manticore-Local` reconstruction to model galaxy bias and account for peculiar velocities gives somewhat lower values unless the sample is assumed to be redshift-selected. We adopt `Manticore-Local` as our fiducial peculiar velocity model and compare the inferred H_0 with results obtained using the C15 reconstruction. We find good agreement between the two, though `Manticore-Local` yields slightly lower values of H_0 (see Fig. 8 or Table 3 for a summary of all H_0 variations). We picked the `Manticore-Local` reconstruction because of its validated peculiar velocity field and improved performance relative to other available reconstructions (S. McAlpine et al. 2025; R. Stiskalek et al. 2025). Additionally, `Manticore-Local` accurately predicts both the masses and positions of all nearby clusters across all several validation metrics (see S. McAlpine et al. 2025). We do not include any SNe in our analysis (except for modelling an SN selection), instead focusing on a detailed investigation of the two lower rungs of the distance ladder.

Using `Manticore-Local` to account for peculiar velocities (and their covariance), the uncertainty on the inferred H_0 with only 35 Cepheid host galaxies ranges between 1.2 to $1.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, which is less than two per cent. Relative to the C15 model, `Manticore-Local` yields a 20 per cent reduction in the uncertainty on H_0. The relatively small volume probed by the second-rung objects underscores the need for increasingly faithful representations of the local Universe such as those enabled by the BORG programme (J. Jasche & B. D. Wandelt 2013; J. Jasche, F. Leclercq & B. D. Wandelt 2015; G. Lavaux & J. Jasche 2016; F. Leclercq et al. 2017; J. Jasche & G. Lavaux 2019; G. Lavaux, J. Jasche & F. Leclercq 2019; N. Porqueres et al. 2019; S. Stopyra et al. 2024; S. McAlpine et al. 2025) to suppress both the statistical and systematic uncertainty due to peculiar velocities. However, assuming these (and other) systematic uncertainties are already under control, further crucial evidence concerning the *Hubble* tension could be obtained by incorporating more second-rung data.

Even restricting to Cepheids, the host galaxies we used represent only a small subset of available Cepheid measurements. We have used the SH0ES sample because of its high-quality (e.g. A. G. Riess et al. 2022b,2023,2024; A. Bhardwaj et al. 2023; L. Breuval et al. 2023, 2024) and self-consistent (J. A. Nájera & H. Desmond 2025) measurements, including well-characterized covariances. However, in principle more Cepheid hosts could be incorporated. Beyond that, several alternative second-rung indicators exist, including the Tip of the Red Giant Branch (TRGB; W. L. Freedman et al. 2019; S. Li et al. 2024), Type II SNe (C. Vogl et al. 2025), J-region asymptotic giant branch (JAGB; W. L. Freedman & B. F. Madore 2020), Mira variables (C. D. Huang et al. 2024), surface brightness fluctuations (SBF; M. Cantiello & J. P. Blakeslee 2023; J. B. Jensen et al. 2025) and potentially other less explored methods (J. A. Nájera & H. Desmond 2025). J. P. Blakeslee et al. (2021) used an SBF calibration anchored to both Cepheid and TRGB distances to infer $H_0 = 73.3 \pm 0.7\,(\mathrm{stat}) \pm 2.4\,(\mathrm{sys})\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. This measurement has recently been updated with *HST* calibration by J. B. Jensen et al. (2025), who found $H_0 = 73.8 \pm 0.7\,(\mathrm{stat}) \pm 2.3\,(\mathrm{sys})\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. The dominant systematic error arises primarily from the distance calibration, suggesting the value of a principled joint calibration approach similar to the one we undertake here. A future analysis should revisit this kind of approach to model jointly the Cepheid, TRGB, and SBF (and perhaps other) distances using the framework outlined here.

It is likely that with additional distance data and ever-improving peculiar velocity modelling (such as with `Manticore-Local`), second-rung inferences could become even more precise, with the key added benefit of being independent of SNe and associated potential systematics. This is even before considering future data: G. Anand et al. (2025) argue that only 24 new Cepheid hosts observed by the proposed Habitable Worlds Observatory would be sufficient to reach a one per cent determination of H_0 using the SH0ES error model. As of now, with `Manticore-Local` we achieve a determination of H_0 to better than two per cent (up to the caveat of uncertain selection).

A two-rung ladder is particular useful because of its complete elimination of SNe, beyond their use to select nearby hosts. Indeed there have been several studies suggesting possible miscalibration of SNe (e.g. R. Wojtak & J. Hjorth 2022; G. Efstathiou 2025; T. J. Hoyt et al. 2025; A. Seifert et al. 2025). A recent exploration by R. Wojtak & J. Hjorth (2025) modelled the SH0ES SN sample as comprising two populations with distinct intrinsic and extrinsic properties – primarily differences in mean absolute magnitude and extinction coefficients – finding a preference for lower H_0 values and reducing the *Hubble* tension by 30–50 per cent. Indeed, an alternative programme to SH0ES – the Chicago-Carnegie Hubble Program (W. L. Freedman et al. 2019) – infers $H_0 = 69.8 \pm 0.8$ (stat) $\pm$ 1.7 (sys) $\mathrm{km\,s^{-1}\,Mpc^{-1}}$ using the TRG. This was recently refined using *James Webb Space Telescope* data of TRGB, Cepheids and the JAGB, finding very similar results (W. L. Freedman et al. 2025).

6 CONCLUSION

The persistence of the *Hubble* tension demands that we multiply cross-check the cosmic distance ladder to vet the local H_0 measurement fully. To this end, we have considered dropping the SNe from the SH0ES pipeline to constrain H_0 from the Cepheid distances (with geometric anchors) and host galaxy redshifts alone. To minimize the risk of introducing new systematics, we keep our data as close as possible to the SH0ES pipeline (A. G. Riess et al. 2022a). This follows in part the analysis of K22, who find $H_0 = 72.9^{+2.4}_{-2.2}\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, where the uncertainty includes both statistics and systematics. We find several issues with this analysis, mostly to do with selection modelling. We introduce a principled framework for modelling selection effects and also employ state-of-the-art peculiar velocity modelling with `Manticore-Local` (S. McAlpine et al. 2025), including marginalization over plausible realizations of the local Universe. We summarize the inferred H_0 in Fig. 8. Our main findings are as follows:

(i) Our fiducial model, which accounts for peculiar velocities using `Manticore-Local`, yields $71.7 \pm 1.3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ when the Cepheid host sample is assumed to be selected on the basis of SN apparent magnitudes, which we consider the more plausible scenario. If however the sample were selected on observed redshift, we would infer $H_0 = 73.0 \pm 1.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. These two cases bracket the range of plausible H_0 values, with mixed selection scenarios lying in between. Our headline result is in 3.3σ tension with combined CMB data from the Atacama Cosmology Telescope (ACT), South Pole Telescope (SPT), and *Planck* (E. Camphuis et al. 2025).

(ii) While S. McAlpine et al. (2025) and R. Stiskalek et al. (2025) have demonstrated that `Manticore-Local` provides the most accurate velocity field among all currently in the literature, we explore others to investigate possible systematic uncertainties. Using the C15 reconstruction shifts the inferred H_0 marginally higher, as well as increasing the uncertainty on H_0 by twenty per cent (producing $H_0 = 72.1 \pm 1.6$ and $74.4 \pm 1.8\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ for SN magnitude and redshift selection, respectively). Any H_0 inference with second-rung objects only is limited by the peculiar velocity modelling, highlighting the need for a state-of-the-art reconstruction like `Manticore-Local`.

(iii) We identify several problems in the K22 analysis of the same data: (1) they adopt an unphysical prior on the galaxy distances used as input to their analysis (r^{-1} rather than the uniform-in-volume r^2; see also H. Desmond et al. 2025), (2) they model selection effects as acting on the prior of the latent true distances, rather than modelling the unobserved sources, and (3) by representing distances in h^{-1} Mpc rather than Mpc units (i.e. actually redshifts), they introduce spurious high powers of H_0 in the prior.

We conclude that the SH0ES Cepheid sample alone yields a *Hubble* constant consistent with the full three-rung SH0ES analysis and with only slightly larger uncertainty, though with a $\sim 1\sigma$ lower central value in the fiducial magnitude-selection scenario. This conclusion relies on careful modelling of both the selection function and peculiar velocities using the state-of-the-art `Manticore-Local` model. The ongoing challenge of understanding and resolving the *Hubble* tension motivates further work on both the second rung of the distance ladder, independent of SNe, and on the modelling of SNe themselves (e.g. `BayeSN`; K. S. Mandel et al. 2022; M. Grayling et al. 2024).

More constructively, our analysis demonstrates that H_0 may be inferred with high precision (less than two per cent statistical uncertainty) from first-plus-second-rung data alone, extending only to $z \approx 0.01$. This suggests that bringing further such data to bear may readily afford a *Hubble* constant inference even more precise than SH0ES and crucially avoiding the third rung (e.g. SNe) altogether. We stress that while the second-rung is independent of SNe (except when modelling the host selection), it is susceptible to significant systematics arising from selection effects or peculiar velocity modelling if these are not properly accounted for. Thus, this endeavour requires principled methods for accounting for selection (as we have implemented here) as well as accurate density and velocity field models of the local Universe, which are growing in quality through Bayesian forward modelling of the galaxy number density field (BORG). Finally, we remark that the sensitivity to selection effects underscores the need for observing programmes to adopt well-defined selection procedures that render the required cosmological inference procedure unambiguous.

ACKNOWLEDGEMENTS

We thank Indranil Banik, Karim Benabed, Nicholas Choustikov, Julien Devriendt, George Efstathiou, Pedro Ferreira, Sebastian von Hausegger, Mike Hudson, D'Arcy Kenworthy, Florent Leclercq, Gabriele Montefalcone, José Antonio Nájera, John Peacock, Erik Peterson, Adam Riess, Subir Sarkar, Dan Scolnic, Ian Steer, Licia Verde, and Tariq Yasin for useful inputs and discussions. This work was done within the Aquila Consortium.[3]

RS acknowledges financial support from Science and Technology Facilities Council (STFC) grant no. ST/X508664/1, the Snell

---

[3] https://aquila-consortium.org

Exhibition of Balliol College, Oxford, and the CCA Pre-doctoral Program. HD is supported by a Royal Society University Research Fellowship (grant no. 211046). ET was supported by STFC through Imperial College Astrophysics Consolidated Grant ST/W000989/1. JJ, GL, and SM acknowledge support from the Simons Foundation through the Simons Collaboration on 'Learning the Universe'. JJ and SM acknowledge support by the research project grant 'Understanding the Dynamic Universe', funded by the Knut and Alice Wallenberg Foundation (Dnr KAW 2018.0067). Additionally, JJ acknowledges financial support from the Swedish Research Council (VR) through the project 'Deciphering the Dynamics of Cosmic Structure' (2020-05143) and GL acknowledges support from the CNRS IEA programme 'Manticore'.

The authors would like to acknowledge the use of the University of Oxford Advanced Research Computing (ARC) facility in carrying out this work.[4] In addition, this work has made use of the Infinity Cluster hosted by Institut d'Astrophysique de Paris, and was granted access to the HPC resources of TGCC (Très Grand Centre de Calcul), Irene-Joliot-Curie supercomputer, under the allocations A0170415682 and SS010415380.

DATA AVAILABILITY

The SH0ES and PANTHEON +data are available at https://github.com/PantheonPlusSH0ES/DataRelease. The `Manticore-Local` resimulations will be made publicly available at https://www.cosmictwin.org. The J. Carrick et al. (2015) velocity reconstruction is available at https://cosmicflows.iap.fr. The code and all other data will be made available on reasonable request to the authors. The code will also be released publicly on publication of the paper.

REFERENCES

Abazajian K. N. et al., 2009, ApJS, 182, 543
Abbott T. M. C. et al., 2022, Phys. Rev. D, 105, 023520
Adame A. G. et al., 2025, J. Cosmol. Astropart. Phys., 2025, 021
Aiola S. et al., 2020, J. Cosmol. Astropart. Phys., 2020, 047
Anand G., Durbin M., Beaton R., Jensen J., Riess A., 2025, preprint (arXiv:2507.02056)
Balkenhol L. et al., 2023, Phys. Rev. D, 108, 023510
Bartlett D. J., Kostić A., Desmond H., Jasche J., Lavaux G., 2022, Phys. Rev. D, 106, 103526
Bennett C. L. et al., 2013, ApJS, 208, 20
Bhardwaj A. et al., 2023, ApJ, 955, L13
Bingham E. et al., 2019, J. Mach. Learn. Res., 20, 1
Blakeslee J. P., Jensen J. B., Ma C.-P., Milne P. A., Greene J. E., 2021, ApJ, 911, 65
Boruah S. S., Hudson M. J., Lavaux G., 2020, MNRAS, 498, 2703
Boruah S. S., Hudson M. J., Lavaux G., 2021, MNRAS, 507, 2697
Boruah S. S., Geshnizjani G., Lavaux G., 2025, preprint (arXiv:2503.07974)
Boubel P., Colless M., Said K., Staveley-Smith L., 2024a, MNRAS, 531, 84
Boubel P., Colless M., Said K., Staveley-Smith L., 2024b, MNRAS, 533, 1550
Bouchet F. R., Colombi S., Hivon E., Juszkiewicz R., 1995, A&A, 296, 575
Bresolin F., Kudritzki R.-P., Urbaneja M. A., Gieren W., Ho I. T., Pietrzyński G., 2016, ApJ, 830, 64
Breuval L. et al., 2023, ApJ, 951, 118
Breuval L. et al., 2024, ApJ, 973, 30
Brout D. et al., 2022, ApJ, 938, 110
Burns C. R. et al., 2018, ApJ, 869, 56
Calabrese E. et al., 2025, J. Cosmol. Astropart. Phys., 2025, 063
Camphuis E. et al., 2025, preprint (arXiv:2506.20707)
Cantiello M., Blakeslee J. P., 2023, preprint (arXiv:2307.03116)
Carreres B. et al., 2025, A&A, 694, A8
Carrick J., Turnbull S. J., Lavaux G., Hudson M. J., 2015, MNRAS, 450, 317
Colombi S., Chodorowski M. J., Teyssier R., 2007, MNRAS, 375, 348
Davis T. M. et al., 2011, ApJ, 741, 67
de Jaeger T., Galbany L., Riess A. G., Stahl B. E., Shappee B. J., Filippenko A. V., Zheng W., 2022, MNRAS, 514, 4620
Desmond H., Stiskalek R., Najera J. A., Banik I., 2025, preprint (arXiv:2511.03394)
Dhawan S. et al., 2022, ApJ, 934, 185
Di Valentino E. et al., 2021, Class. Quantum Gravity, 38, 153001
Di Valentino E. et al., 2025, Phys. Dark Univ., 49, 101965
Efstathiou G., 2025, MNRAS, 538, 875
Efstathiou G., Rosenberg E., Poulin V., 2024, Phys. Rev. Lett., 132, 221002
Freedman W. L., Madore B. F., 2020, ApJ, 899, 67
Freedman W. L. et al., 2019, ApJ, 882, 34
Freedman W. L., Madore B. F., Hoyt T. J., Jang I. S., Lee A. J., Owens K. A., 2025, ApJ, 985, 203
Gall C., Izzo L., Wojtak R., Hjorth J., 2024, preprint (arXiv:2411.05642)
Gelman A., Carlin J. B., Stern H. S., Rubin D. B., 2004, Bayesian Data Analysis, 2nd edn. Chapman and Hall/CRC, Boca Raton
Grayling M., Thorp S., Mandel K. S., Dhawan S., Uzsoy A. S. M., Boyd B. M., Hayes E. E., Ward S. M., 2024, MNRAS, 531, 953
Guidorzi C. et al., 2017, ApJ, 851, L36
Hoffman M. D., Gelman A., 2011, preprint (arXiv:1111.4246)
Hoyt T. J., Jang I. S., Freedman W. L., Madore B. F., Owens K. A., Lee A. J., 2025, preprint (arXiv:2503.11769)
Huang C. D. et al., 2024, ApJ, 963, 83
Hubble E., 1929, Proc. Natl. Acad. Sci., 15, 168
Huchra J. P. et al., 2012, ApJS, 199, 26
Hui L., Greene P. B., 2006, Phys. Rev. D, 73, 123526
Huterer D., Shafer D. L., Schmidt F., 2015, J. Cosmol. Astropart. Phys., 2015, 033
Jasche J., Lavaux G., 2019, A&A, 625, A64
Jasche J., Wandelt B. D., 2013, MNRAS, 432, 894
Jasche J., Leclercq F., Wandelt B. D., 2015, J. Cosmol. Astropart. Phys., 2015, 036
Jensen J. B., Blakeslee J. P., Cantiello M., Cowles M., Anand G. S., Tully R. B., Kourkchi E., Raimondo G., 2025, ApJ, 987, 87
Jones D. H. et al., 2009, MNRAS, 399, 683
Kaiser N., 1988, MNRAS, 231, 149
Kelly B. C., 2007, ApJ, 665, 1489
Kelly B. C., Fan X., Vestergaard M., 2008, ApJ, 682, 874
Kenworthy W. D. et al., 2022, ApJ, 935, 83
Lavaux G., Hudson M. J., 2011, MNRAS, 416, 2840
Lavaux G., Jasche J., 2016, MNRAS, 455, 3169
Lavaux G., Jasche J., Leclercq F., 2019, preprint (arXiv:1909.06396)
Leavitt H. S., Pickering E. C., 1912, Harv. Coll. Obs. Circ., 173, 1
Leclercq F., Jasche J., Lavaux G., Wandelt B., Percival W., 2017, J. Cosmol. Astropart. Phys., 2017, 049
Lewis A., Challinor A., 2011, Astrophysics Source Code Library, record ascl:1102.026
Li S., Riess A. G., Busch M. P., Casertano S., Macri L. M., Yuan W., 2021, ApJ, 920, 84
Li S. et al., 2024, ApJ, 976, 177
Lilow R., Nusser A., 2021, MNRAS, 507, 1557
Louis T. et al., 2025, J. Cosmol. Astropart. Phys., 2025, 062
Mandel K. S., Thorp S., Narayan G., Friedman A. S., Avelino A., 2022, MNRAS, 510, 3939
McAlpine S., Jasche J., Ata M., Lavaux G., Stiskalek R., Frenk C. S., Jenkins A., 2025, MNRAS, 540, 716
Messenger C., Veitch J., 2013, New J. Phys., 15, 053027
Monaghan J. J., 1992, ARA&A, 30, 543
Nájera J. A., Desmond H., 2025, MNRAS, 541, 671
Pesce D. W. et al., 2020, ApJ, 891, L1
Peterson E. R. et al., 2022, ApJ, 938, 112
Pettini M., Pagel B. E. J., 2004, MNRAS, 348, L59

---

[4] https://doi.org/10.5281/zenodo.22558

Phan D., Pradhan N., Jankowiak M., 2019, preprint (arXiv:1912.11554)
Pietrzyński G. et al., 2019, Nature, 567, 200
Planck Collaboration VI, 2020, A&A, 641, A6
Planck Collaboration LI, 2017, A&A, 607, A95
Porqueres N., Kodi Ramanah D., Jasche J., Lavaux G., 2019, A&A, 624, A115
Reid M. J., Pesce D. W., Riess A. G., 2019, ApJ, 886, L27
Riess A. G. et al., 2018, ApJ, 861, 126
Riess A. G., Casertano S., Yuan W., Bowers J. B., Macri L., Zinn J. C., Scolnic D., 2021, ApJ, 908, L6
Riess A. G. et al., 2022a, ApJ, 934, L7
Riess A. G. et al., 2022b, ApJ, 938, 36
Riess A. G. et al., 2023, ApJ, 956, L18
Riess A. G. et al., 2024, ApJ, 962, L17
Roy O., Vetterli M., 2007, in 15th European Signal Processing Conference, p. 606
Said K., Colless M., Magoulas C., Lucey J. R., Hudson M. J., 2020, MNRAS, 497, 1275
Schaller M. et al., 2024, MNRAS, 530, 2378
Schombert J., McGaugh S., Lelli F., 2020, AJ, 160, 71
Schöneberg N., Abellán G. F., Sánchez A. P., Witte S. J., Poulin V., Lesgourgues J., 2022, Phys. Rep., 984, 1
Seifert A., Lane Z. G., Galoppo M., Ridden-Harper R., Wiltshire D. L., 2025, MNRAS, 537, L55
Skrutskie M. F. et al., 2006, AJ, 131, 1163
Stiskalek R., Desmond H., Devriendt J., Slyz A., Lavaux G., Hudson M. J., Bartlett D. J., Courtois H. M., 2025, MNRAS, 545, staf1960
Stopyra S., Peiris H. V., Pontzen A., Jasche J., Lavaux G., 2024, MNRAS, 527, 1244
Tristram M. et al., 2024, A&A, 682, A37
Tsaprazi E., Heavens A. F., 2025, MNRAS, 539, 1448
Tsaprazi E. et al., 2022, MNRAS, 510, 366
Vogl C. et al., 2025, A&A, 702, A41
Wang L., Steinhardt P. J., 1998, ApJ, 508, 483
Westover M., 2007, PhD thesis, Harvard University
Willick J. A., Batra P., 2001, ApJ, 548, 564
Wojtak R., Hjorth J., 2022, MNRAS, 515, 2790
Wojtak R., Hjorth J., 2024, MNRAS, 533, 2319
Wojtak R., Hjorth J., 2025, A&A, 702, A176
Yahil A., Strauss M. A., Davis M., Huchra J. P., 1991, ApJ, 372, 380

APPENDIX A: MOCK DATA BIAS TESTS

We assess the impact of selection on either SN apparent magnitude or host galaxy redshift using mock catalogues. We generate mock catalogues designed to match the properties of the observed data and demonstrate that the inferred value of H_0 remains unbiased.

The SN magnitude-selected sample is generated as follows. For simplicity, rather than modelling the full CPLR, we assume all Cepheid stars have identical absolute magnitudes. We draw source galaxy distances from a uniform-in-volume (r^2) distribution. For each trial, we compute the SN apparent magnitude, given an assumed $M_B = -19.25$, adding a scatter of $\sigma_{\rm SN} = 0.15$ dex. If the resulting apparent magnitude is below the selection threshold, we retain the sampled distance; otherwise, we repeat the process. For each accepted host galaxy, we generate five Cepheids (the precise number is unimportant), each at the same distance, with absolute magnitude $\tilde{M}_W = -18$ and a scatter in observed magnitude of $\sigma_{\rm Ceph} = 0.1$ dex. The observed redshift of the host galaxy is then computed, assuming a given $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and including a velocity dispersion of $\sigma_v = 250\,\mathrm{km\,s^{-1}}$ to mimic peculiar velocities. This procedure is repeated until a sample of 35 host galaxies is obtained. Finally, to mimic the inclusion of the LMC and NGC 4258 as calibrators, we add two nearby sources. For each, we assign a true distance modulus and sample an 'observed' distance modulus with a scatter of 0.25 dex, similar to the values reported in the literature. The host galaxy redshift-limited sample is generated analogously, except that SN apparent magnitudes are not modelled; instead, the rejection process is based on whether the host galaxy redshift falls below a specified threshold. When sources are drawn from an r^2 distribution, a maximum distance must be imposed for practical reasons. We adopt an upper limit of 150 Mpc, which ensures that all sources passing either the magnitude or redshift selection criteria lie well below this threshold (typically within 50 Mpc). We also test an alternative procedure in which we generate a fixed total number of galaxies and then retain those that pass the cut, rather than resampling until a fixed number pass the cut. In both cases our procedure is unbiased.

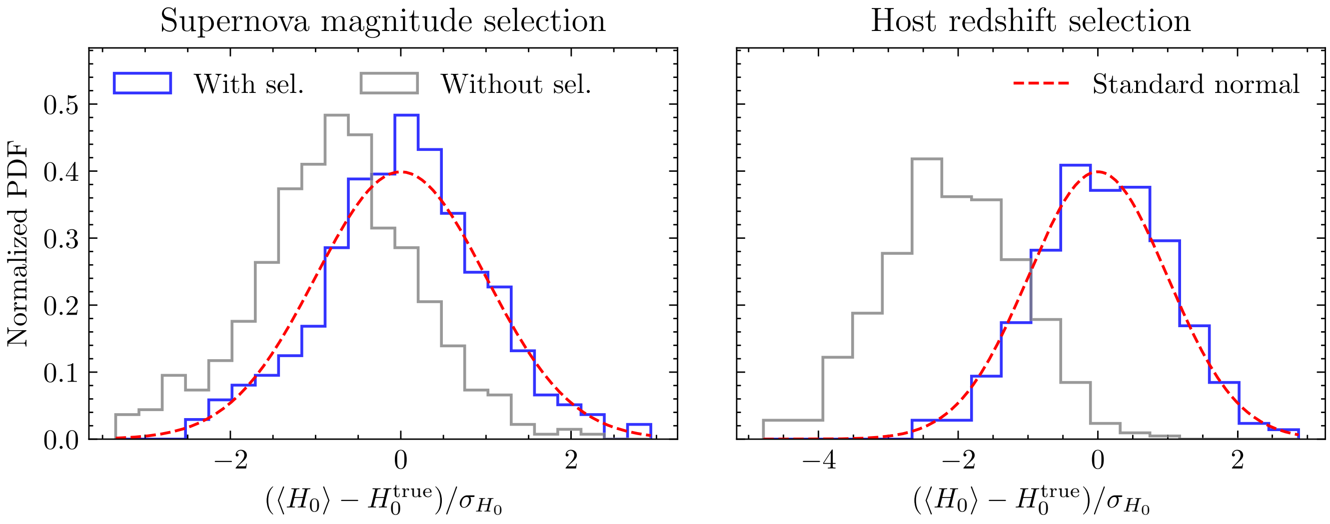

Figure A1. Bias test of H_0 for 'Cepheid'-only mock inferences, as described in Appendix A, under host selection by either SN magnitude or host redshift. When selection effects are not modelled (grey), the inferred H_0 is significantly biased low. By contrast, when selection is properly modelled (blue), the inference is unbiased in both cases, as indicated by excellent agreement with the standard normal distribution (red dashed).

In Fig. A1, we present bias tests in H_0 for the SN magnitude- and redshift-limited selection. In each case, we use identical injected parameter values but draw the mock catalogue using different random seeds. This procedure is repeated 500 times. The bias is defined as

$$\mathrm{Bias}(H_0) = \frac{\langle H_0 \rangle - H_0^{\mathrm{true}}}{\sigma_{H_0}}, \tag{A1}$$

where $\langle H_0 \rangle$ and σ_{H_0} denote the posterior mean and standard deviation of the inferred H_0, and H_0^{true} is the true (injected) value. If the inferred H_0 is unbiased (and its posterior is Gaussian), then the distribution of $\mathrm{Bias}(H_0)$ across the mock realizations should follow a standard normal distribution. Fig. A1 demonstrates that, when selection effects are properly modelled, the recovered H_0 is unbiased on average in all cases. In contrast, neglecting selection effects introduces a significant bias in the inferred H_0 for each scenario.

APPENDIX B: LOCAL UNIVERSE RECONSTRUCTIONS

As discussed in Section 3, some of our peculiar velocity models use the density and velocity field reconstruction of either `Manticore-Local` or C15. This enables us to both model the inhomogeneous Malmquist bias as a function of the local density, as well as the contribution of the LOS peculiar velocities to observed redshifts. Both reconstructions are based on the 2M++ galaxy catalogue (G. Lavaux & M. J. Hudson 2011), which contains 69160 galaxies and combines Two Micron All-Sky Survey photometry (2MASS; M. F. Skrutskie et al. 2006) with redshifts from 2MRS (J. P. Huchra et al. 2012), 6dF (D. H. Jones et al. 2009), and Sloan Digital Sky Survey (SDSS) DR7 (K. N. Abazajian et al. 2009). Apparent magnitudes are corrected for Galactic extinction, k-corrections, evolution, and surface brightness dimming. The 2M++ sample is magnitude-limited to $K < 11.5$ in the 2MRS region and $K < 12.5$ in the 6dF and SDSS regions.

The density and velocity field of either `Manticore-Local` or C15 is reconstructed in real space. Since the real-space distances to the host galaxies are not known a priori, we consider all plausible velocities along the LOS to each source. The predicted peculiar velocity of the ith host is given by

$$V_{\mathrm{pec},i}^{\mathrm{C15}} = (\beta \boldsymbol{V}(\boldsymbol{r}_i) + \boldsymbol{V}_{\mathrm{ext}}) \cdot \hat{\boldsymbol{r}}_i, \tag{B1}$$

where $\boldsymbol{V}(\boldsymbol{r}_i)$ is the velocity field evaluated at candidate galaxy position $\boldsymbol{r}_i$. β is set to unity for `Manticore-Local` and treated as a free parameter when using the reconstruction of C15 (where it is related to σ_8). The associated density field can be used to model galaxy bias $n_g(r)$, i.e. the so-called inhomogeneous Malmquist bias, which is the prior preference for galaxies to be located in regions of higher matter density. This leads to the following prior probability for the physical distance (in Mpc) to a given galaxy

$$\pi(r \mid \boldsymbol{\theta}) = \frac{r^2 n_g(r, \boldsymbol{\theta})}{\int_0^R r'^2 n_g(r', \boldsymbol{\theta}) \mathrm{d}r'}, \tag{B2}$$

where the integral is evaluated along the LOS to the galaxy. Without $n_g(r)$ the prior could not be normalized but this would not matter because the normalization would be independent of the model parameters and hence the prior could be left improper. However, because of the varying LOS densities per source, the normalization in modelling the galaxy bias is no longer source-independent. It may also depend on inferred parameters and, therefore, must be computed explicitly for each source. We discuss the galaxy bias choices for `Manticore-Local` and C15 below.

Ideally, the integral in equation (B2) should extend to $R \to \infty$, but the reconstructed density field is available only within approximately 200 Mpc. We therefore assume $n_g(r) = 1$ outside the reconstruction, so the integrand reduces smoothly to r^2. To avoid sharp discontinuities, we extrapolate the overdensity field $\delta(r)$ beyond the reconstructed volume with an exponential decay towards zero with a $5\,h^{-1}$ Mpc scale length. We have checked that the results are not sensitive to this choice. The normalization integral is then computed numerically up to R. Contributions beyond R are independent of the model parameters and can be neglected; the posterior is therefore independent of R provided it is > 200 Mpc. We note that using an improper prior proportional to χ^2, where χ denotes comoving distance in h^{-1} Mpc, leads to biased inference because the prior explicitly depends on h. This is effectively putting a prior on redshift, where it should be on physical distance.

B1 The `Manticore-Local` reconstruction

`Manticore` is the latest set of reconstructions using the BORG algorithm (J. Jasche & B. D. Wandelt 2013; J. Jasche et al. 2015; G. Lavaux & J. Jasche 2016; F. Leclercq et al. 2017; J. Jasche & G. Lavaux 2019; G. Lavaux et al. 2019; N. Porqueres et al. 2019; S. Stopyra et al. 2024; S. McAlpine et al. 2025), which infers a posterior distribution of voxel-by-voxel densities at $z = 1000$ (and hence at all lower redshifts) by employing a gravity forward model, while accounting for redshift-space distortions, observational selection effects and a galaxy bias model, and comparing the resulting predictions to the observed galaxy number density field with a Poisson likelihood. In particular, we use the `Manticore-Local` run, which constrains the initial conditions to match the matter distribution of the local Universe applied to the 2M++ data.

To increase the resolution beyond the native BORG inference, we use the posterior resimulation suite `Manticore-Local` (S. McAlpine et al. 2025). This comprises 30 N-body simulations with the SWIFT code (M. Schaller et al. 2024), each based on independent posterior initial conditions inferred on a 256^3 grid (spatial resolution 2.65 h^{-1} Mpc). The simulations use 1024^3 dark matter particles (no baryonic physics), yielding a mean interparticle separation of $0.7\,h^{-1}$ Mpc and a particle mass of $2 \times 10^{10} h^{-1}\,\mathrm{M}_\odot$. Each box has a size of $681.1\,h^{-1}$ Mpc, while the constrained volume extends to $\sim 155\,h^{-1}$ Mpc, set by the completeness of the 2M++ survey; regions beyond this are unconstrained but provide the correct long-range forces in a Λ-CDM framework. From the $z = 0$ particle snapshots, we construct density and velocity fields using the smooth particle hydrodynamics technique (J. J. Monaghan 1992; S. Colombi, M. J. Chodorowski & R. Teyssier 2007), with a minimum of 32 smoothing neighbours as in section IV.B.1 of D. J. Bartlett et al. (2022). These fields are rendered on a 1024^3 grid with a voxel size of $0.7\,h^{-1}$ Mpc. In Fig. B1 ,we compare the LOS velocities for a sample of Cepheid host galaxies between the `Manticore-Local` and C15 reconstructions, finding very good agreement.

`Manticore-Local` assumes cosmological parameters from the Dark Energy Survey Year 3 (DES Y3; T. M. C. Abbott et al. 2022) '3×2pt + All Ext'. Λ-CDM cosmology: $h = 0.681$, $\Omega_\mathrm{m} = 0.306$, $\Omega_\Lambda = 0.694$, $\Omega_\mathrm{b} = 0.0486$, $A_\mathrm{s} = 2.099 \times 10^{-9}$, $n_\mathrm{s} = 0.967$, and $\sigma_8 = 0.807$. Thus, unlike C15, which is derived independently of the growth factor and consequently of σ_8, these are assumed in `Manticore-Local`. However, by introducing a free parameter β (see Appendix B2) to scale the predicted peculiar velocities, we

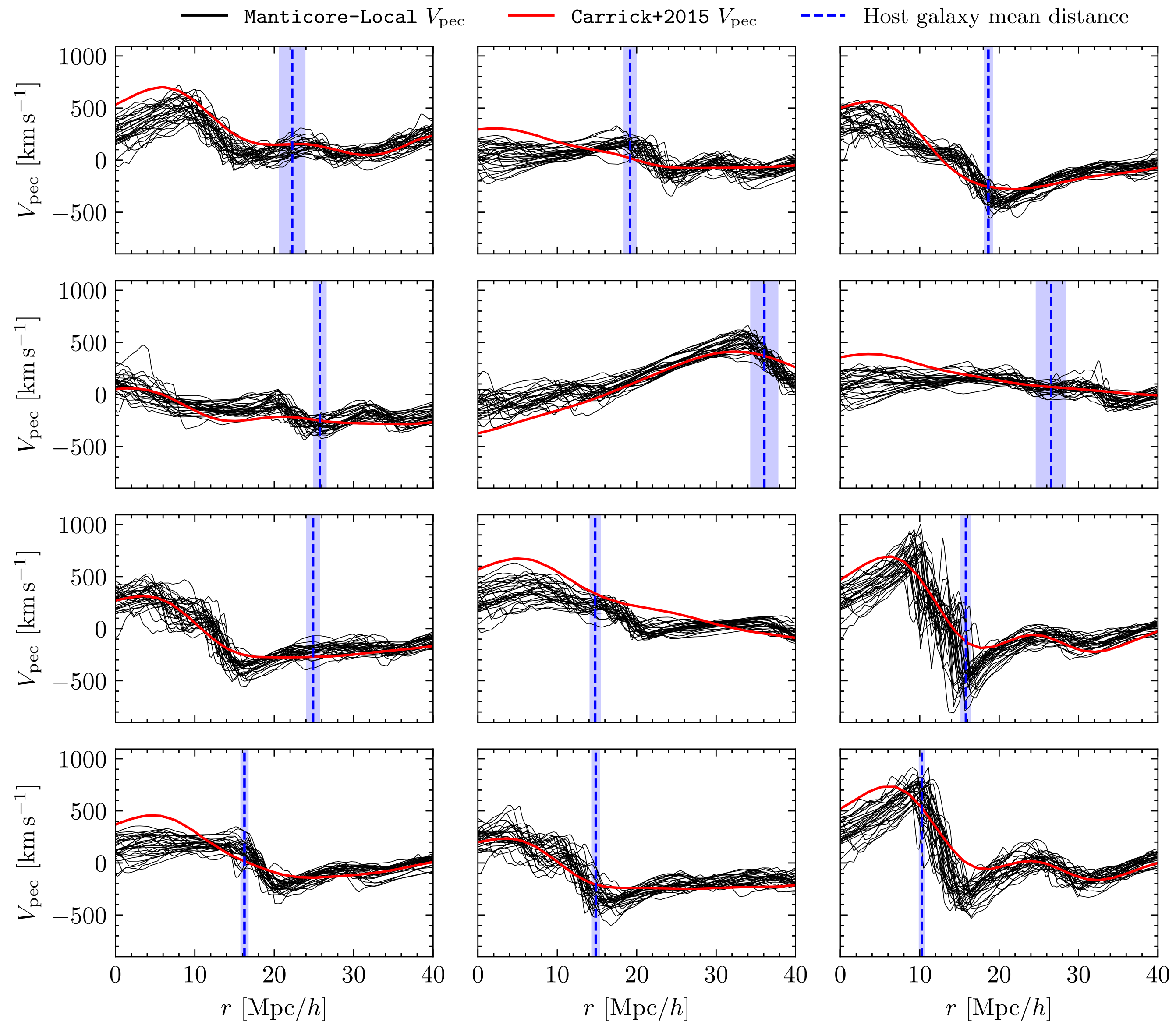

Figure B1. Comparison of radial peculiar velocities, $V_{\rm pec}$, between `Manticore-Local` (black) and J. Carrick et al. (2015) (red) along the LOS to 12 randomly selected Cepheid host galaxies, whose mean inferred distances are shown by the blue vertical lines with shaded bands indicating 1σ uncertainty. Each panel shows 30 `Manticore-Local` realizations, representing independent draws from the BORG posterior, while J. Carrick et al. (2015) provide only a mean field estimate. The two reconstructions are generally in good agreement. In several cases, the host galaxies reside in clusters, as indicated by the characteristic infall signature in $V_{\rm pec}$.

can (approximately) assess the impact of varying σ_8 – we find it to have no impact on the inferred H_0.

Moreover, while `Manticore-Local` assumes $H_0 = 68.1\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, we assume our inference to be insensitive to using it while varying H_0. This is because the deviations we probe from the fiducial H_0 are small and because the hosts typically reside in relatively low-density regions (see E. Tsaprazi & A. F. Heavens 2025 and table 2 of E. Tsaprazi et al. 2022). Moreover, when querying `Manticore-Local` for either density or velocity, we consistently convert sampled distances from Mpc to h^{-1} Mpc using the sampled value of H_0. This removes the first-order effect of varying H_0.

To model the galaxy bias $n_g(r)$ we adopt a phenomenological double power law,

$$n_g(r,\,\boldsymbol{\theta}) = \left(\frac{\rho}{\rho_t}\right)^{\alpha_{\rm low}} \left(1 + \frac{\rho}{\rho_t}\right)^{\alpha_{\rm high}-\alpha_{\rm low}}, \tag{B3}$$

where $\rho \equiv \rho(r)$ is the density at distance r along the LOS to the host, $\alpha_{\rm low}$ and $\alpha_{\rm high}$ are the two slopes, and ρ_t is the transition density, all of which we infer jointly with the other parameters. We place a truncated Gaussian prior on $\alpha_{\rm low}$, centred at 1 with unit standard deviation and restricted to positive values, and on $\alpha_{\rm high}$, centred at 0.5 with unit standard deviation and likewise truncated at zero. For

$\ln \rho_t$ we adopt a zero-centred Gaussian prior with standard deviation of two.

We also consider a model in which the small-scale velocity dispersion depends on local density to capture higher unresolved dispersion in clusters,

$$\sigma_v(\delta) = \sigma_{v,\mathrm{low}} + \frac{\sigma_{v,\mathrm{high}} - \sigma_{v,\mathrm{low}}}{1 + \left(\frac{1+\delta}{\rho_{\sigma_v}}\right)^{-k_{\sigma_v}}}, \tag{B4}$$

where δ is the matter overdensity, $\sigma_{v,\mathrm{low}}$ and $\sigma_{v,\mathrm{high}}$ set the asymptotic dispersions in low- and high-density regions, ρ_{σ_v} is the transition density, and k_{σ_v} controls the steepness of the sigmoid. We adopt the following priors: $\sigma_{v,\mathrm{low}}$ and $\sigma_{v,\mathrm{high}}$ each follow a Maxwell distribution with scale $200\,\mathrm{km\,s^{-1}}$ (mean of approximately $320\,\mathrm{km\,s^{-1}}$), $\ln \rho_{\sigma_v}$ follows a normal distribution with mean of unity and standard deviation of five, and k_{σ_v} follows a truncated normal with mean of unity, standard deviation of unity, and lower bound at zero. Our fiducial analysis retains the simpler assumption of a constant σ_v, but Table 3 shows that this density-dependent model yields almost identical results.

Instead of using a single field, for `Manticore-Local` we employ a set of 30 posterior samples (30 independent density and velocity fields from N-body simulations with constrained initial conditions) over which we marginalize. Following R. Stiskalek et al. (2025), we do this by averaging the likelihood over the `Manticore-Local` posterior samples.

B2 The Carrick et al. (2015) reconstruction

C15 reconstructs the luminosity-weighted density field from redshift-space galaxy positions in the 2M++ catalogue (G. Lavaux & M. J. Hudson 2011) using the iterative scheme of A. Yahil et al. (1991). The velocity field is derived from the density field using linear theory and scaled by a free parameter β, defined as

$$\beta \equiv \frac{f\sigma_{8,\mathrm{NL}}}{\sigma_{8,g}}, \tag{B5}$$

where $f \approx \Omega_\mathrm{m}^{0.55}$ is the dimensionless growth rate in Λ -CDM (F. R. Bouchet et al. 1995; L. Wang & P. J. Steinhardt 1998). Here, $\sigma_{8,g}$ denotes the fluctuation amplitude in the galaxy field, while $\sigma_{8,\mathrm{NL}}$ corresponds to the non-linear matter field. In the 2M++ data, $\sigma_{8,g}$ is estimated to be 0.98 ± 0.07 (M. Westover 2007) or 0.99 ± 0.04 (J. Carrick et al. 2015). The velocity field is computed on a 256^3 grid within a $400\,h^{-1}$ Mpc box, assuming $\Omega_\mathrm{m} = 0.3$. This reconstruction has been used to constrain structure growth and the S_8 parameter using peculiar velocity data (e.g. S. S. Boruah, M. J. Hudson & G. Lavaux 2020; K. Said et al. 2020; P. Boubel et al. 2024a; R. Stiskalek et al. 2025), and is a standard tool to correct for peculiar velocities in H_0 inferences (e.g. C. Guidorzi et al. 2017; S. S. Boruah, M. J. Hudson & G. Lavaux 2021; D. Brout et al. 2022; E. R. Peterson et al. 2022; P. Boubel et al. 2024b; S. S. Boruah, G. Geshnizjani & G. Lavaux 2025). Since the C15 reconstruction is linear, it is appropriate to adopt a simple linear bias model of the form

$$n_g(r) = 1 + b_1\delta(r), \tag{B6}$$

where b_1 is the linear bias parameter. To maintain consistency with the earlier definition of β, we set $b_1 \equiv f/\beta$.

APPENDIX C: THEORY PECULIAR VELOCITY COVARIANCE

The peculiar velocity covariance matrix quantifies correlations in LOS velocities induced by large-scale structure (N. Kaiser 1988; L. Hui & P. B. Greene 2006; T. M. Davis et al. 2011). For galaxies indexed by i and j, it is defined as

$$\xi_{ij} = \left\langle (\boldsymbol{v}_i \cdot \hat{\boldsymbol{r}}_i)(\boldsymbol{v}_j \cdot \hat{\boldsymbol{r}}_j) \right\rangle, \tag{C1}$$

and is computed from the Λ -CDM matter power spectrum as

$$\xi_{ij} = \frac{\mathrm{d}D_i}{\mathrm{d}\tau}\frac{\mathrm{d}D_j}{\mathrm{d}\tau}\int \frac{\mathrm{d}k}{2\pi^2} P(k) F(kr_i,\, kr_j,\, \hat{\boldsymbol{r}}_i \cdot \hat{\boldsymbol{r}}_j), \tag{C2}$$

where D_i is the linear growth factor at the redshift of the ith galaxy, and τ is the conformal time with $d\tau \equiv dt/a$. F is defined byF(u, v, $\cos\theta) = \sum_{\ell=0}^{\infty}(2\ell+1)j_\ell'(u)j_\ell'(v)P_\ell(\cos\theta)$where $j_\ell'(x)$ is the derivative of the spherical Bessel function of order ℓ with respect to its argument and $P_\ell(x)$ is the Legendre polynomial of order ℓ. We evaluate equation (C2) using the non-linear matter power spectrum $P(k)$ from CAMB (A. Lewis & A. Challinor 2011) over the interval $k \in [10^{-4},\, 20]$ h/Mpc, assuming the fiducial *Planck* cosmology (Planck Collaboration VI 2020). We verify that changing the background cosmology, including variations in the assumed value of H_0 when computing $P(k)$ (e.g. switching from the *Planck* to the SH0ES value), does not affect the inferred value of H_0. We truncate the sum in equation (C3) at $\ell = 2000$, which is sufficient for convergence.

For the special case of considering the same object so that $\cos\theta = 1$ and $r_i = r_j = r$, we have that $F(kr,\, kr,\, 1) = 1/3$, yielding the diagonal elements

$$\xi_{ii} = \left(\frac{\mathrm{d}D_i}{\mathrm{d}\tau}\right)^2 \int \frac{\mathrm{d}k}{6\pi^2} P(k). \tag{C3}$$

We find a value of approximately $255\,\mathrm{km\,s^{-1}}$ for our choice of the non-linear matter power spectrum $P(k)$. Adopting a linear $P(k)$ instead yields diagonal elements of approximately $200\,\mathrm{km\,s^{-1}}$. This also contrasts with K22, who incorrectly claim that the diagonal velocity dispersion is approximately $380\,\mathrm{km\,s^{-1}}$ when computed using a non-linear matter power spectrum, leading to an overestimation of their uncertainty. We have managed to reproduce this number by using inconsistent units for $P(k)$ and k.

APPENDIX D: CEPHEID HOST PROPERTIES

In Table D1, we list the Cepheid host galaxy Galactic coordinates, CMB-frame redshifts, distance moduli, and peculiar velocities as inferred from both `Manticore-Local` and C15. For both inferences, we assume SN magnitude selection. Comparing their peculiar velocities, we find the two models to be generally consistent. The `Manticore-Local` peculiar velocities assume $\beta = 1$, include the sampled $\boldsymbol{V}_\mathrm{ext}$, and are computed by stacking the velocities from each Markov Chain Monte Carlo step and across the 30 `Manticore-Local` realizations. In contrast, the peculiar velocities from C15 are computed from the mean reconstructed velocity field while jointly sampling β and $\boldsymbol{V}_\mathrm{ext}$; their quoted uncertainties therefore represent only the posterior width in those parameters, not of the large-scale structure reconstruction uncertainty. As a comparison, the `Manticore-Local` model yields a residual velocity scatter of $\sigma_v = 148.2 \pm 31.7\,\mathrm{km\,s^{-1}}$, whereas the C15 model gives $\sigma_v = 221.1 \pm 33.0\,\mathrm{km\,s^{-1}}$.

Table D1. Summary of the 35 Cepheid host galaxies used in our analysis.

Name	ℓ [deg]	b [deg]	$cz_{\rm CMB}$ [km s^{-1}]	$\mu_{\rm host}$ [mag]	$V_{\rm pec}^{\texttt{Manticore-Local}}$ [km s^{-1}]	$V_{\rm pec}^{\rm C15}$ [km s^{-1}]
M101	102	60	366	$29.17^{+0.04}_{-0.04}$	-81^{+99}_{-84}	-14^{+73}_{-66}
M1337	303	53	2896	$32.91^{+0.07}_{-0.07}$	155^{+73}_{-73}	105^{+59}_{-55}
N0691	141	-39	2581	$32.76^{+0.07}_{-0.08}$	147^{+82}_{-72}	191^{+64}_{-67}
N1015	172	-54	2401	$32.55^{+0.06}_{-0.06}$	29^{+76}_{-66}	-2^{+55}_{-57}
N1309	202	-53	2003	$32.51^{+0.05}_{-0.05}$	-180^{+99}_{-105}	-155^{+59}_{-56}
N1365	238	-55	1379	$31.36^{+0.05}_{-0.05}$	-17^{+150}_{-108}	-46^{+65}_{-62}
N1448	252	-51	1097	$31.28^{+0.04}_{-0.04}$	-46^{+132}_{-129}	-50^{+68}_{-67}
N1559	274	-41	1304	$31.46^{+0.05}_{-0.05}$	-93^{+108}_{-94}	-76^{+75}_{-78}
N2442	281	-22	1544	$31.46^{+0.06}_{-0.06}$	166^{+80}_{-88}	140^{+80}_{-82}
N2525	232	11	1787	$32.00^{+0.07}_{-0.07}$	25^{+81}_{-79}	-65^{+74}_{-78}
N2608	195	34	2386	$32.70^{+0.06}_{-0.06}$	-191^{+78}_{-78}	-298^{+79}_{-81}
N3021	192	51	1775	$32.37^{+0.10}_{-0.09}$	-355^{+102}_{-91}	-394^{+78}_{-75}
N3147	136	39	2977	$33.18^{+0.07}_{-0.07}$	-294^{+94}_{-89}	-319^{+89}_{-78}
N3254	200	59	1649	$32.34^{+0.06}_{-0.06}$	-430^{+83}_{-87}	-405^{+71}_{-70}
N3370	225	60	1619	$32.13^{+0.05}_{-0.05}$	-353^{+105}_{-97}	-329^{+64}_{-67}
N3447	228	61	1394	$31.93^{+0.04}_{-0.03}$	-205^{+98}_{-104}	-296^{+64}_{-68}
N3583	158	62	2317	$32.80^{+0.06}_{-0.06}$	-313^{+68}_{-74}	-403^{+75}_{-69}
N3972	139	60	1106	$31.69^{+0.06}_{-0.07}$	-465^{+120}_{-99}	-364^{+79}_{-71}
N3982	139	60	1106	$31.65^{+0.06}_{-0.06}$	-453^{+129}_{-107}	-359^{+78}_{-70}
N4038	287	42	1979	$31.66^{+0.10}_{-0.09}$	340^{+75}_{-86}	350^{+66}_{-59}
N4424	284	71	767	$30.83^{+0.12}_{-0.10}$	452^{+340}_{-298}	401^{+82}_{-84}
N4536	293	65	1049	$30.84^{+0.05}_{-0.05}$	646^{+145}_{-243}	467^{+63}_{-64}
N4639	294	76	1385	$31.79^{+0.08}_{-0.10}$	-414^{+199}_{-125}	-182^{+57}_{-58}
N4680	301	51	2791	$32.73^{+0.11}_{-0.12}$	207^{+77}_{-79}	169^{+58}_{-54}
N5468	335	53	2992	$33.10^{+0.06}_{-0.05}$	28^{+75}_{-68}	-30^{+57}_{-52}
N5584	345	55	1904	$31.81^{+0.05}_{-0.05}$	142^{+99}_{-94}	22^{+56}_{-55}
N5643	321	15	1433	$30.52^{+0.05}_{-0.05}$	351^{+91}_{-105}	451^{+78}_{-79}
N5728	337	38	3148	$33.02^{+0.10}_{-0.07}$	150^{+73}_{-74}	121^{+69}_{-63}
N5861	349	39	2179	$32.28^{+0.08}_{-0.09}$	175^{+84}_{-97}	51^{+70}_{-63}
N5917	355	40	2108	$32.36^{+0.07}_{-0.07}$	1^{+95}_{-100}	-55^{+70}_{-63}
N7250	94	-14	878	$31.52^{+0.10}_{-0.10}$	-375^{+73}_{-70}	-398^{+69}_{-74}
N7329	321	-46	3124	$33.26^{+0.07}_{-0.07}$	79^{+84}_{-89}	11^{+79}_{-90}
N7541	83	-51	2305	$32.60^{+0.07}_{-0.07}$	-18^{+80}_{-94}	-70^{+58}_{-57}
N7678	99	-37	3145	$33.23^{+0.07}_{-0.07}$	75^{+84}_{-74}	88^{+57}_{-64}
U9391	101	53	1991	$32.79^{+0.06}_{-0.06}$	-446^{+81}_{-82}	-385^{+76}_{-71}

Notes. Columns list galaxy name, Galactic longitude, latitude, CMB-frame redshift, host distance modulus inferred in the `Manticore-Local` analysis, and the predicted LOS peculiar velocities from `Manticore-Local` and C15. Uncertainties are 1σ. All values assume SN magnitude selection.

This paper has been typeset from a TEX/LATEX file prepared by the author.