Sample variance in the local measurements of the Hubble constant

Hao-Yi Wu[1]⋆ and Dragan Huterer[2]⋆

[1]*California Institute of Technology, MC 367–17, Pasadena, CA 91125, USA*
[2]*Department of Physics, University of Michigan, 450 Church Street, Ann Arbor, MI 48109, USA*

ABSTRACT

The current $>3\sigma$ tension between the Hubble constant H_0 measured from local distance indicators and from cosmic microwave background is one of the most highly debated issues in cosmology, as it possibly indicates new physics or unknown systematics. In this work, we explore whether this tension can be alleviated by the sample variance in the local measurements, which use a small fraction of the Hubble volume. We use a large-volume cosmological N-body simulation to model the local measurements and to quantify the variance due to local density fluctuations and sample selection. We explicitly take into account the inhomogeneous spatial distribution of type Ia supernovae. Despite the faithful modelling of the observations, our results confirm previous findings that sample variance in the local Hubble constant ($H_0^{\rm loc}$) measurements is small; we find $\sigma(H_0^{\rm loc}) = 0.31$ km s^{-1}Mpc^{-1}, a nearly negligible fraction of the ~ 6 km s^{-1}Mpc^{-1} necessary to explain the difference between the local and global H_0 measurements. While the H_0 tension could in principle be explained by our local neighbourhood being a underdense region of radius ~ 150 Mpc, the extreme required underdensity of such a void ($\delta \simeq -0.8$) makes it very unlikely in a ΛCDM universe, and it also violates existing observational constraints. Therefore, sample variance in a ΛCDM universe cannot appreciably alleviate the tension in H_0 measurements even after taking into account the inhomogeneous selection of type Ia supernovae.

Key words: methods: numerical – galaxies: haloes – cosmological parameters – large-scale structure of Universe – cosmology: theory.

1 INTRODUCTION

The Hubble constant H_0 – the current expansion rate of the Universe – has had a long history of increasingly precise measurements. It begins with Edwin Hubble's groundbreaking but startlingly inaccurate measurement of 500 km s^{-1}Mpc^{-1} (Hubble 1929), followed by subsequent decades with competing but mutually discrepant claims that H_0 is either about 50 km s^{-1}Mpc^{-1} (e.g. Sandage & Tammann 1982) or about 100 km s^{-1}Mpc^{-1} (e.g. de Vaucouleurs & Bollinger 1979). The situation was significantly clarified with measurements that employed data from the Hubble Space Telescope, which indicated $H_0 = 72 \pm 8$ km s^{-1}Mpc^{-1} (Freedman et al. 2001; also see e.g. Freedman & Madore 2010 for a review). Recent developments paved the way for truly precision-level measurements of the Hubble constant – one from the local distance ladder and standard candles (e.g, Freedman et al. 2012; Riess et al. 2016), and the other from the global constraints on cosmological parameters using the cosmic microwave background (CMB) measurements (e.g. Hinshaw et al. 2013; Planck Collaboration XIII 2016).

Interestingly, the latter two precise and physically very different types of measurements currently appear to be in tension. In particular, the H_0 derived from the local distance ladder and type Ia Supernovae (SNe, Riess et al. 2009, 2011, 2016) has been significantly higher than the H_0 derived from the acoustic peak scale of CMB (Planck Collaboration XVI 2014; Planck Collaboration XIII 2016). For example, Riess et al. (2016, R16 hereafter) presented $H_0^{\rm loc} = 73.24 \pm 1.74$ km s^{-1}Mpc^{-1}, while the Planck collaboration (Planck Collaboration XLVI 2016, P16 hereafter) presented $H_0^{\rm CMB} = 66.93 \pm 0.62$ km s^{-1}Mpc^{-1} (see their table 8). The R16 value is 3.4σ higher and the relative difference, $(H_0^{\rm loc} - H_0^{\rm CMB})/H_0^{\rm CMB} = 9$ per cent, is much larger than the 2.4 per cent error bar of R16 and the 0.9 per cent error bar of P16.

The H_0 discrepancy has caused much interest in the cosmology community, as it could be a harbinger of new physics. Essentially, the data indicate that the local measurement of the expansion rate, measured at distance scale $\lesssim 400$ Mpc, is higher than the globally inferred value from the Hubble volume $\sim (10\,{\rm Gpc})^3$. Various possibilities for this discrepancy have been discussed, including the systematic errors in CMB (e.g. Addison et al. 2016) or in the distance ladder (e.g. Efstathiou 2014), although the recent independent analyses of the local measurements (Cardona, Kunz & Pettorino 2017; Feeney,

⋆ E-mail: hywu@caltech.edu (H-YW); huterer@umich.edu (DH)

Mortlock & Dalmasso 2017; Follin & Knox 2017; Zhang et al. 2017) confirm the R16 results. Measurements using the time delay of gravitational lensing (e.g. Suyu et al. 2013; Bonvin et al. 2017) and the Tully–Fisher relation (e.g. Sorce, Tully & Courtois 2012) appear to be consistent with the R16 local results, while measurements using baryon acoustic oscillations (e.g. Aubourg et al. 2015) and tip of the red-giant branch distances (e.g. Tammann & Reindl 2013) are consistent with the P16 CMB results.

Local density fluctuations can make local measurements of the Hubble constant deviate from the global value (e.g. Turner, Cen & Ostriker 1992; Wang, Spergel & Turner 1998; Shi & Turner 1998; Cooray & Caldwell 2006; Hui & Greene 2006; Martinez-Vaquero et al. 2009; Sinclair, Davis & Haugbølle 2010; Courtois et al. 2013; Ben-Dayan et al. 2014; Fleury, Clarkson & Maartens 2017). In particular, if the Milky Way is located in an underdense region in the cosmic web (the so-called 'Hubble bubble'), nearby galaxies will tend to have positive peculiar velocities, which will bias the H_0^{loc} measurement high. Several lines of evidence have supported the idea of a local underdensity (e.g. Zehavi et al. 1998; Jha, Riess & Kirshner 2007; Keenan, Barger & Cowie 2013; Whitbourn & Shanks 2014). However, it is unclear whether such a local underdensity can fully account for the difference between H_0^{loc} and H_0^{CMB}.

Previous works have explored the statistical and systematic errors of H_0^{loc} using analytic models and N-body simulations. Marra et al. (2013) analytically calculated the systematic error of H_0^{loc} due to local density perturbations. They found that the local density perturbations can account for at most 2.4 per cent of the systematic errors of H_0^{loc} measured at $0.01 < z < 0.1$, and that a $6\,\mathrm{km\,s^{-1}Mpc^{-1}}$ deviation in H_0^{loc} is very rare. Using cosmological N-body simulations, Wojtak et al. (2014, W14 hereafter) explored observers located in voids and on dark matter haloes of different masses, quantifying the deviation of H_0^{loc} as a function of distance scales. They found that the discrepancy between H_0^{loc} and H_0^{CMB} cannot be accounted for using density fluctuations, unless the Milky Way is located at the centre of one of the 10 largest voids found in $6\ h^{-1}\mathrm{Gpc}$ volume. Odderskov, Hannestad & Haugbølle (2014) also reached the conclusion that the sample variance is too small to alleviate the tension, after taking into account the depth and sky coverage of SN observations in light-cone simulations.

The purpose of this work is to explicitly incorporate in the analysis the highly non-uniform spatial distribution of SNe used in the measurement of H_0^{loc}. We quantify two sources of sample variance:[1] local density contrast and inhomogeneous selection of SN sample. The aforementioned previous theoretical analyses of sample variance in H_0^{loc} relied on various schemes of weighting and averaging radial velocities of haloes isotropically selected from N-body simulations. However, even with weighting schemes that are judiciously chosen to mimic the actual observations, it is a priori far from obvious to us that the inhomogeneity of the SN sample would not increase the sample variance in the derived value of H_0^{loc}.

Our approach is to explicitly mimic the highly inhomogeneous selection function of SNe, using the R16 Supercal SN sample as a baseline. We incorporate variations in how the R16 SN sample is orientated relative to the simulation's coordinate system. We also study how haloes are assigned to physical SN locations in order to use them to calculate H_0^{loc} in simulations. We note that R16 explicitly corrected for the peculiar velocity of each SN using the density field from 2M++ (Carrick et al. 2015); however, such

corrections are susceptible to large systematic uncertainties in the local density field reconstruction. To be conservative, we do not take into account the fact that such corrections have been applied in the R16 analysis. Therefore, our calculation corresponds to an upper limit of how much sample variance contributes to the error budget; the sample variance will be smaller if the peculiar velocity correction is sufficiently accurate.

This paper is organized as follows. Section 2 describes our approach for quantifying the sample variance of H_0^{loc}. Section 3 details how we map observed SNe to dark matter haloes. Section 4 presents our main results, including the distribution of H_0^{loc} from various sample selections. We explore the effect of density fluctuations in Section 5. We discuss our results in Section 6 and conclude in Section 7. Throughout this work, we use the same cosmological parameters as the in the Dark Sky simulations (see Section 3.2).

2 SIMULATING THE H_0^{loc} MEASUREMENT

In this section, we review how the H_0^{loc} measurement is done, introduce our procedure for simulating this process and derive the estimator for H_0^{loc}.

2.1 Overview of our approach

We start by briefly reviewing how the Hubble constant is measured in R16. The R16 H_0^{loc} measurement first employs four distance 'anchors' with direct geometric distance measurements.[2]. The second rung of the distance ladder is a set of 19 galaxies ('calibrators') with both Cepheids and SNe; the four anchors set the absolute distance of Cepheids and the 19 calibrators provide distance calibration for SNe. The third and final rung of the distance ladder is a set of 217 SNe Ia at $0.023 < z < 0.15$. The global fit to the set of all standard candles – anchors, calibrators and SNe Ia – is used to determine the value of H_0^{loc}.

When we use the distance–redshift relation of standard candles to derive H_0^{loc}, the redshift can be modified by the local density fluctuations, and the derived H_0^{loc} is thus biased. The redshift range adopted in R16, $0.023 < z < 0.15$, is selected so as to avoid the effect of peculiar velocities at lower redshift and the effect of dark energy at higher redshift.

In this paper, we simulate the selection of the R16 SN distances in order to account for the sample variance of H_0^{loc}. Our approach is represented visually in Fig. 1 and is summarized as follows:

(i) Select a sub-volume in a cosmological simulation, within which each halo in some mass range is a possible host of an SN.

(ii) Orient the R16 SN coordinates with respect to the sub-volume.

(iii) Assign each of the SNe to the closest halo.

(iv) Calculate the deviation of the local Hubble constant, $\Delta H_0^{\mathrm{loc}} = H_0^{\mathrm{loc}} - H_0^{\mathrm{true}}$, from the radial velocities of these SN hosts.

(v) Go to (ii), repeat the measurements for many different orientations and obtain the histogram of $\Delta H_0^{\mathrm{loc}}$ of different orientations from a single sub-volume.

(vi) Go to (i), repeat the measurements for multiple, non-overlapping sub-volumes and obtain the distribution of $\Delta H_0^{\mathrm{loc}}$ from all sub-volumes and all orientations.

---

[1] In this work, we use the two terms 'sample variance' and 'cosmic variance' interchangeably.

[2] These anchors are: parallaxes of Cepheids in the Milky Way, water masers in NGC 4258 and detached eclipsing binaries in the Large Magellanic Cloud and in M31.
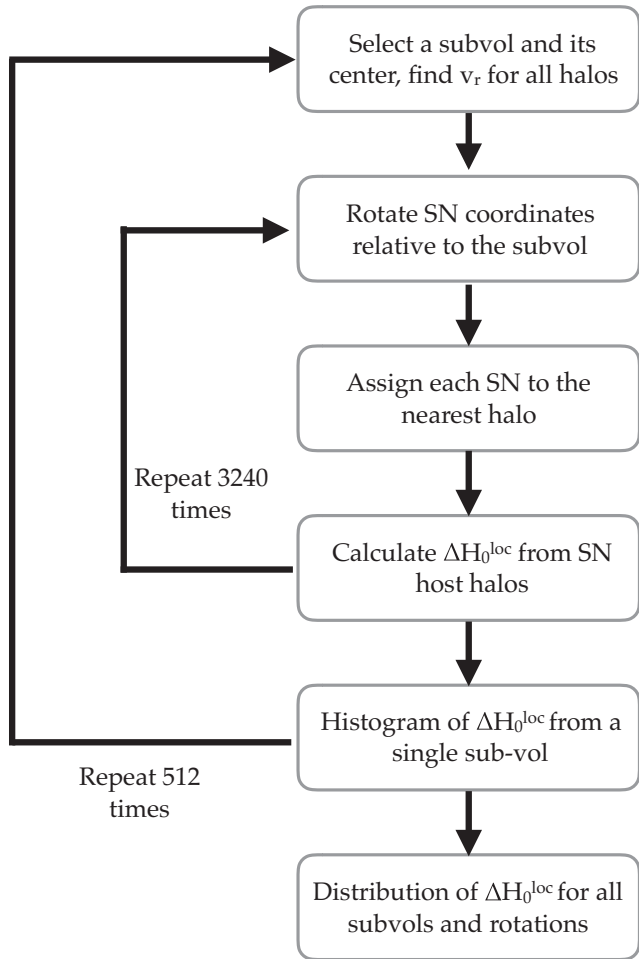
Figure 1. Flowchart depicting our simulation procedure. The inner loop runs over all mutual rotations between the SN and the simulation coordinate systems. The outer loop runs over all sub-volumes from the Dark Sky simulations, each sub-volume representing a realization of the local Universe.

2.2 From SN to $\Delta H_0^{\rm loc}$

Once the overall SN distance scale has been calibrated with nearby objects (Cepheids and the four anchors), the redshift and apparent magnitude of each SN provides a direct – but noisy – estimate of the Hubble constant. That is, the magnitude fluctuation, $\delta m = (5/\ln 10)(\delta r/r) \propto v_r/r$, is precisely an estimate of the deviation in the Hubble constant; here v_r is the peculiar velocity along the line of sight and r is the comoving distance. Below we adopt the formalism from R16 to drive the estimator of $H_0^{\rm loc}$.

In R16, the information from SNe is compressed into one number, a_x, which is defined as the intercept of the magnitude–redshift relation in the x-band and is related to $H_0^{\rm loc}$ via

$$\log_{10} H_0^{\rm loc} = \frac{M_x^0}{5} + a_x + 5 \,, \tag{1}$$

where M_x^0 is the absolute magnitude of SNe Ia derived from the distance anchors.

The value of a_x is determined from the Hubble diagram (see fig. 8 in R16). Let $m_{x,i}^0$ be the apparent magnitude of the $i^{\rm th}$ SN in

the x-band (fit by the SALT2 light-curve fitting algorithm) and z_i its measured redshift; a_x is then given by

$$a_x = \frac{1}{N} \sum_{i=1}^{N} \log_{10}(cz_i(1 + O(z_i))) - 0.2 m_{x,i}^0 \,. \tag{2}$$

The measured redshift is perturbed by the peculiar velocity along the line of sight, which leads to a perturbation in a_x,

$$\Delta a_x \simeq \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\ln 10} \frac{c \Delta z_i}{c z_i} \simeq \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\ln 10} \frac{v_{r,i}}{r_i H_0^{\rm loc}} \,. \tag{3}$$

Here, we adopt $\Delta z = v_r/c$ and $r = cz/H_0^{\rm loc}$, which hold for $z \ll 1$. From equations (1) and (3), we obtain

$$\Delta H_0^{\rm loc} = (H_0^{\rm loc} \ln 10) \, \Delta a_x = \frac{1}{N} \sum_{i=1}^{N} \frac{v_{r,i}}{r_i} \,. \tag{4}$$

equation (4) simply averages the contribution from each SN, and it makes sense to instead inverse-variance weight their contributions; this leads to the estimator we employ in this paper:

$$\Delta H_0^{\rm loc} = \frac{\displaystyle\sum_{i=1}^{N} \frac{1}{\sigma_i^2} \frac{v_{r,i}}{r_i}}{\displaystyle\sum_{i=1}^{N} \frac{1}{\sigma_i^2}} \qquad \text{(our estimator).} \tag{5}$$

Here, σ_i is the error bar of the magnitude of each SN, set to the square root of the diagonal of the noise covariance matrix of SNe (see Scolnic et al. 2014).

Since each SN gives a measurement of v_r/r, the minimum-variance estimator simply averages those measurements with the inverse-variance weighting. Our estimator in equation (5) is the appropriate one given our goal to simulate the use of SNe in obtaining the Hubble constant. We note that there is an alternative estimator based on least squares (LSQ), $\sum(w(r)rv_r)/\sum(w(r)r^2)$, where $w(r)$ is the weight (see e.g. W14 and Odderskov, Hannestad & Brandbyge 2017).[3] Our estimator is equivalent to the LSQ estimator with each object weighted by r^{-2} (see Appendix A for discussions on weighting). Therefore, our estimator gives more weight to objects at small r. For the distribution of SNe from R16 – shown in Fig. 2 – our estimator returns a *three times larger* standard deviation in $H_0^{\rm loc}$ than the LSQ estimator; see Appendix B. We will see below that, despite giving a larger variance when applied to the R16 set of SNe, the sample variance in $H_0^{\rm loc}$ remains small when compared to the difference between P16 and R16 measurements.

3 IMPLEMENTATION

In this section, we provide further details of the observation data and simulations used in our analysis.

3.1 Supercal SN data

We use the recent compilation of SN data 'Supercal' (Scolnic et al. 2014, 2015), which features a uniform photometric calibration based on the consistent photometric system of Pan-STARRS1 across 3π of the sky. This data set is based on the SALT2 light-curve

---

[3] We note that W14 used v_r averaged in a redshift bin and weighted each bin by r^2 (R. Wojtak, private communication). This procedure is equivalent to our equation (4).
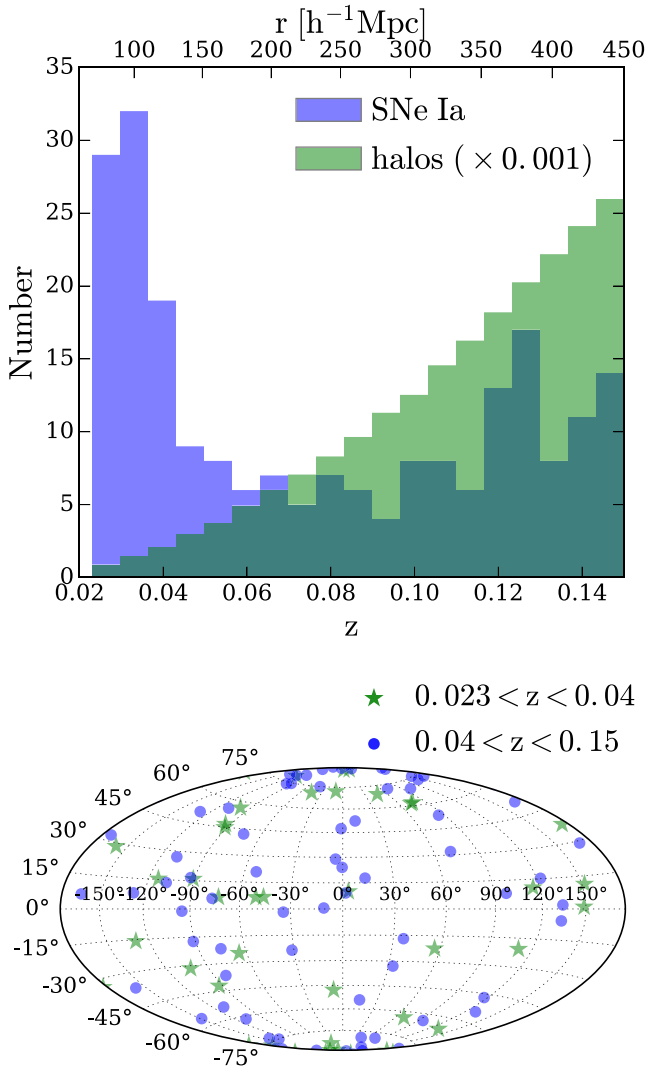
Figure 2. Top: redshift distributions of the Supercal SN sample (blue histogram) and the dark matter haloes in Dark Sky simulations (green histogram), for $0.023 < z < 0.15$, split into 20 bins. Bottom: angular distribution of the Supercal SN sample in the Galactic coordinates, shown with the Hammer (equal-area) projection. The two symbols correspond to two ranges of redshift as shown. Note that the angular distribution is highly sparse.

model (Guy et al. 2010) and includes a correction of distance bias due to intrinsic scatter of SN brightness and measurement noise (Scolnic & Kessler 2016). This data set is also used in the recent H_0^{loc} analysis of R16 (see their Section 4.3). In our analysis, we follow R16 and use 217 Supercal SNe Ia that lie in $0.023 < z < 0.15$. This choice of lower limit $z = 0.023$ corresponds to $\simeq 100$ Mpc, above which the effects of peculiar velocities in SN data are small. The upper panel of Fig. 2 shows the redshift distribution of these 217 SNe Ia; the radial distribution is skewed towards low redshift and is different from a volume-limited sample, for which the number per bin would be $\propto z^2$.

The lower panel of Fig. 2 shows the Hammer (equal-area) projection of the positions of the Supercal SN sample in the Galactic coordinates. While the SNe are found over the whole sky, their overall distribution is inhomogeneous. This becomes particularly obvious if we consider the full three-dimensional distribution of SNe in space.

3.2 Dark sky simulations

We use the public release of the Dark Sky simulations[4] (Skillman et al. 2014). In particular, we use the largest volume $ds14\_a$ with 1.07×10^{12} (10240^3) particles within a volume of $(8\ h^{-1}\text{Gpc})^3$. The mass resolution is $3.9 \times 10^{10}\ h^{-1}\text{M}_\odot$, and the Plummer equivalent softening length is $36.8\ h^{-1}$kpc. The cosmological parameters correspond to a flat ΛCDM model and are consistent with *Planck* and other probes (e.g. Planck Collaboration XIII 2016): $\Omega_M = 0.295$; $\Omega_b = 0.0468$; $\Omega_\Lambda = 0.705$; $h = 0.688$; $\sigma_8 = 0.835$.

The N-body simulation is performed using the adaptive tree code 2HOT (Warren 2013), and the dark matter haloes are identified using the halo finder ROCKSTAR (Behroozi, Wechsler & Wu 2013a). The data is accessible online using YT (Turk et al. 2011). We use dark matter haloes with virial mass $M_{vir} > 10^{12.3}\ \text{M}_\odot$ (35 particles). We have explicitly verified the completeness of haloes at this mass.

We divide this $(8\ h^{-1}\text{Gpc})^3$ volume into 512 sub-volumes of $(1\ h^{-1}\text{Gpc})^3$. We then choose the halo with virial mass $M_{vir} \in [10^{12.3}, 10^{12.4}]\ \text{M}_\odot$ that is closest to the centre of each sub-volume as our observer. This choice simulates 512 separate observers located on Milky Way mass haloes (see e.g. Bland-Hawthorn & Gerhard 2016, for constraints on Milky Way mass), and each observer has a separate sub-volume of the large-scale structure out to the distance of interest ($z_{max} = 0.15$). For the host haloes of SNe, we also use Milky Way mass haloes with $M_{vir} \in [10^{12.3}, 10^{12.4}]\ \text{M}_\odot$, and we have explicitly checked that this choice leads to the same results as using all haloes above $10^{13}\ \text{M}_\odot$.

Using N-body simulations, W14 found that if observers are located on dark matter haloes, H_0^{loc} will be biased low for $r_{max} \lesssim 100\ h^{-1}$Mpc, and the deviation is larger for higher halo mass. In Appendix A, we show that our choice of halo mass leads to no bias on H_0^{loc} at all scales. In addition, W14 showed that using the velocities with respect to the CMB rest frame is the correct choice, while using the velocities with respect to the observer's rest frame will lead to a unphysical bias. We use halo velocities with respect to the simulation's rest frame (which corresponds to the CMB rest frame).

3.3 Matching SNe to haloes

The top panel of Fig. 2 shows the redshift distribution of the SNe from R16 (blue histogram) and that of the dark matter haloes from the Dark Sky simulations (green histogram). While the number of haloes per redshift slice scales with z^2, the number of SNe peaks at $z \simeq 0.03$. Moreover, there are $\sim 200\,000$ haloes in each sub-volume in the aforementioned mass interval $[10^{12.3}, 10^{12.4}]\ \text{M}_\odot$ and redshift range $0.023 < z < 0.15$, but there are only 217 SNe in this redshift range. Below we show that the skewed and sparse distribution of SNe will lead to additional sample variance.

The matching of SNe to haloes is done as follows. First, for each Dark Sky sub-volume (1 h^{-1}Gpc on a side), we find the halo closest to the centre (a halo in our mass range can typically be found within $\sim 15\ h^{-1}$Mpc of any given point). We place the observer on this halo, which is roughly the Milky Way mass given our choice for halo masses.

We then consider how to orientate the SN coordinate system relative to the simulation sub-volume. While the distribution of SNe in redshift and angle is given, the orientation of their coordinate system relative to that of the simulation frame is arbitrary and, given

---

[4] http://darksky.slac.stanford.edu
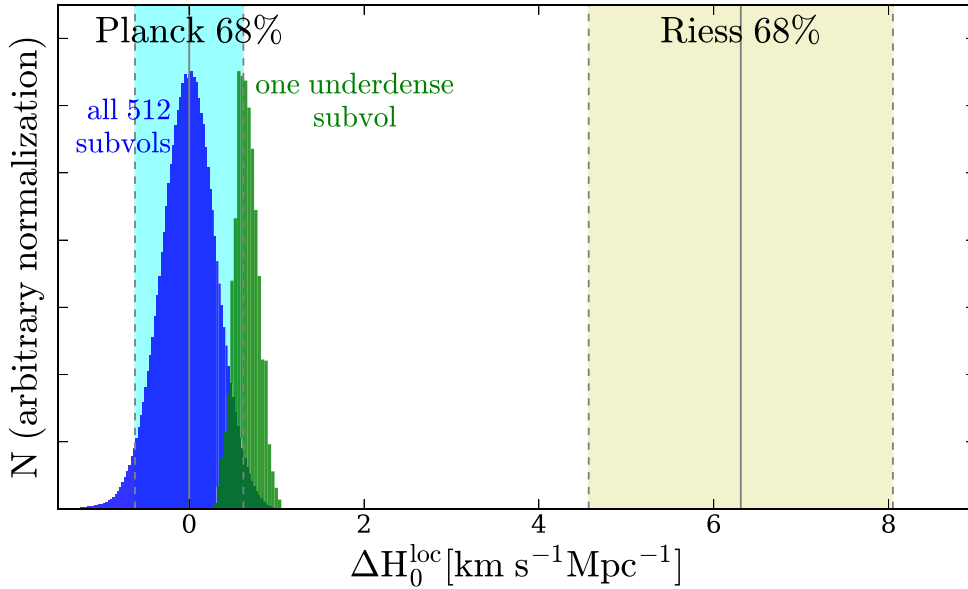
Figure 3. Sample variance in $\Delta H_0^{\mathrm{loc}}$ from our simulations, compared to P16 and R16 error bars (assuming P16 is the true global value). The blue histogram shows 3240 rotations of the SN coordinate system from 512 sub-volumes in the Dark Sky simulations, corresponding to ~1.5 million SN-to-halo coordinate system configurations. The green, slightly more jagged, histogram shows the results of a particularly underdense sub-volume with a high $\Delta H_0^{\mathrm{loc}}$ at the 2σ level relative to all sub-volumes. The two histograms are separately normalized. Note that the sample variance in H_0^{loc} is much smaller than the difference between R16 and P16 measurements.

the highly inhomogeneous distribution of SN in the volume, will surely lead to additional variance. To account for this previously unaccounted-for source of the sample variance, we explore many possible orientations of the SN frame within a fixed simulation frame. To vary over the orientations, we employ 3240 Euler angles; see Appendix C for details. We explicitly verified that the corresponding angular step is sufficiently small to provide converged results.

Now that the three-dimensional coordinates of each SN and halo have been calculated for each rotation, we assign the closest halo to each SN. The radial velocity of that SN is simply given by

$$v_{r,i} \equiv \boldsymbol{v}_i \cdot \frac{(\boldsymbol{r}_i - \boldsymbol{r}_{\mathrm{obs}})}{|\boldsymbol{r}_i - \boldsymbol{r}_{\mathrm{obs}}|}, \tag{6}$$

where the label i refers to the ith SN, and $\boldsymbol{r}_{\mathrm{obs}}$ is the location of the observer.

4 RESULTS

Our principal results are shown in Fig. 3 and Table 1.

Fig. 3 shows the distribution of $\Delta H_0^{\mathrm{loc}}$ from our simulated measurements, compared with the best-fitting values and 1σ measurement errors of R16 and P16. Both histograms are separately normalized for visual clarity. The blue histogram includes 3240 rotations of the SN coordinate system in each of the 512 sub-volumes, for a total of $\gtrsim 1.5$ million configurations. This histogram is approximately Gaussian with a standard deviation of 0.31 km s^{-1}Mpc^{-1}. In units of this sample variance-only error, the R16 measured value is about 19 standard deviations away from the P16 value. In addition, the green histogram shows an example of a very underdense sub-volume, with $\delta = -0.16$ (at the scale of $120 \ h^{-1}$Mpc) and $\Delta H_0^{\mathrm{loc}} = 0.65 \pm 0.13$ km s^{-1}Mpc^{-1} (~2σ upward fluctuation). None of our ~1.5 million configurations comes

even close to appreciably helping reconcile the P16 and R16 measurements.

Table 1 shows the cumulative value of the standard deviation in $\Delta H_0^{\mathrm{loc}}$ as various contributions to the sample variance are added to our simulations. Using all haloes (without weighting) in the range $0.023 < z < 0.15$, the mean of $\Delta H_0^{\mathrm{loc}}$ values among the 512 sub-volumes is less than 0.01 km s^{-1}Mpc^{-1}, and the standard deviation is 0.12 km s^{-1}Mpc^{-1}. Still calculating $\Delta H_0^{\mathrm{loc}}$ from all haloes, but now weighting each halo's contribution with $n^{\mathrm{SN}}(z)/n^{\mathrm{halo}}(z)$, produces $\sigma(\Delta H_0^{\mathrm{loc}}) = 0.38$ km s^{-1}Mpc^{-1}. Most of the increase in this quantity comes from the skewed redshift distribution of $n^{\mathrm{SN}}(z)$. We have checked that sampling 217 haloes randomly at $0.023 < z < 0.15$ barely increases the scatter (0.13 km s^{-1}Mpc^{-1}), while sampling 217 haloes according to $n^{\mathrm{SN}}(z)$ gives the same results as $n^{\mathrm{SN}}(z)/n^{\mathrm{halo}}(z)$ weighting (0.39 km s^{-1}Mpc^{-1}).

The penultimate column of Table 1 shows the results when the rotations of the SN coordinate system with respect to the simulation are taken into account, without weighting (using equation 4); the standard deviation rises slightly. The last column is similar to the previous one, with each SN weighted by the inverse of the square of the magnitude error (using equation 5). This weighting decreases the scatter slightly because the magnitude error is smaller (thus weight is larger) for high-redshift SNe. This last column contains our best estimate for the sample variance in $\Delta H_0^{\mathrm{loc}}$: approximately Gaussian distribution with $\sigma(\Delta H_0^{\mathrm{loc}}) = 0.31$ km s^{-1}Mpc^{-1}. We note that this value is similar to the 0.4 per cent uncertainty due to sample variance quoted in R16 (see their fig. 12).

5 IMPACT OF LOCAL DENSITY CONTRAST

We now investigate the relation between fluctuations in H_0^{loc} and the density contrast $\delta = (\rho - \bar{\rho})/\bar{\rho}$, which can also be measured in the simulation. A negative/positive density contrast implies a higher/lower local expansion rate than the average value. The

Table 1. Contributions to the total sample variance by various assumptions about the H_0^{loc} measurements.

	Contributions to rms variance in $\Delta H_0^{\mathrm{loc}}$ ($\mathrm{km\,s^{-1}Mpc^{-1}}$)			
Source of scatter:	Box-to-box, all haloes, no weighting	+ weighted by $n^{\mathrm{SN}}(z)/n^{\mathrm{halo}}(z)$	+ 3D matching, 3240 rotations, no weighting	+ weighted by mag err (equation 5)
Cumulative σ (0.023 < z < 0.15):	0.12	0.38	0.42	**0.31**

deviation in H_0^{loc} is directly related to the local density contrast δ via

$$\frac{\Delta H_0^{\mathrm{loc}}}{H_0^{\mathrm{loc}}} = -\frac{1}{3}\delta f(\Omega_M)\Theta(\delta, \Omega_M) , \qquad (7)$$

where Θ is the non-linear correction (Marra et al. 2013),

$$\Theta \simeq 1 - 0.0882\delta - \frac{0.123\sin\delta}{1.29 + \delta} . \qquad (8)$$

In a flat ΛCDM universe,

$$f(\Omega_M, z) \approx \left(\Omega_M(z)\right)^{\gamma} , \qquad (9)$$

where $\gamma \approx 0.55$. For the cosmology adopted by Dark Sky simulations, $f \approx 0.5$ at $z = 0$.

In this section, we put the $\Delta H_0^{\mathrm{loc}}$ and δ measured from simulations in the context of several observational results.

5.1 Observed local density contrast

We consider three observations of the density in the local Universe.

(i) Keenan et al. (2013) used the K-band (near-infrared) luminosity density to trace the matter density for $0.01 < z < 0.2$ over $600\,\mathrm{deg}^2$ of the sky. They found that the nearby Universe is underdense out to $\simeq 300\,\mathrm{Mpc}$. Specifically, at $\simeq 200\,\mathrm{Mpc}$, the density is $\simeq 30$ per cent lower than at $\simeq 600\,\mathrm{Mpc}$ ($\delta = -0.3$, see their fig. 11).

(ii) Whitbourn & Shanks (2014) used the K-band and r-band galaxy number densities out to $\simeq 300\,h^{-1}\mathrm{Mpc}$. They included almost full-sky data but the depth is not uniform. They found $\simeq 5$, 15 and 40 per cent underdensity ($\delta = -0.05, -0.15$ and -0.4) out to $150\,h^{-1}\mathrm{Mpc}$, depending on the survey field (see their figs 3, 8 and 10).

(iii) Carrick et al. (2015) used the galaxy luminosity density from 2M++ to reconstruct the density of galaxies out to $200\,h^{-1}\mathrm{Mpc}$. They included full-sky results out to $\simeq 120\,h^{-1}\mathrm{Mpc}$. They found no evidence for local underdensity ($\delta = 0$); in fact, their results suggested slight overdensity (see their fig. 10).[5]

We emphasize that there are substantial sources of systematic uncertainty in mapping out the mass density contrast from the luminosity and number density of galaxies. Hence, we only include the observational constraints above as a rough guideline on what the current data indicate.

5.2 Comparing simulations and observations

We next turn to the relation between density contrast and local measurements of the Hubble constant in the simulations. We first verify that, as expected, $\Delta H_0^{\mathrm{loc}}$ and δ measured over the same volume obey

---

[5] Their luminosity-weighted density contrast and peculiar velocity are publicly available at http://cosmicflows.iap.fr

the linear-theory relation in equation (7) if all haloes in the volume are used in the H_0^{loc} measurement (as long as very nearby objects are excised to avoid non-linear effects). In practice, however, the observational constraints of δ are limited to relatively low redshift (e.g. $\lesssim 120\,h^{-1}\mathrm{Mpc}$ or $z \lesssim 0.04$ for Carrick et al. 2015), while the SN sample extends to $z = 0.15$.

Despite this mismatch in redshift range, we have found that $\Delta H_0^{\mathrm{loc}}$ measured for $0.023 < z < 0.15$ (with the R16 selection) correlates strongly with δ measured for $z < 0.04$. This is unsurprising because the R16 SN selection peaks at $z < 0.04$, and thus $\Delta H_0^{\mathrm{loc}}$ measured for $z < 0.15$ and that for $z < 0.04$ correlate strongly with each other. This correlation indicates that we can infer $\Delta H_0^{\mathrm{loc}}$ from the δ measured at relatively low redshift.

The left-hand panel of Fig. 4 shows the correlation between δ and $\Delta H_0^{\mathrm{loc}}$. The blue points show the $\Delta H_0^{\mathrm{loc}}$ measured using all haloes in sub-volumes, while the green points with error bars show $\Delta H_0^{\mathrm{loc}}$ from 3240 rotations of each sub-volume. Note that the comparison to the standard theory expectation for the $\Delta H_0^{\mathrm{loc}}$–$\delta$ relation is not applicable here, since the two quantities are evaluated at different redshifts. We fit a straight line to the points and extend it to low δ.

The right-hand panel of Fig. 4 is the zoomed-out version of the left-hand panel, with some additional information. The green points with errors again show the measurements from our simulation. The solid line is the linear fit, and the dash curve has the non-linear correction (equation 8) applied to the linear fit. The three vertical arrows denote the observational estimates of $\delta = -0.3, -0.15$ and -0.05, outlined in Section 5.1. As can be seen, $\delta < -0.3$ is expected to be extremely rare, as we do not find one in 512 sub-volumes. Moreover, even such a low δ could not account for the $6\,\mathrm{km\,s^{-1}Mpc^{-1}}$ difference between R16 and P16; the $6\,\mathrm{km\,s^{-1}Mpc^{-1}}$ difference corresponds to $\delta \simeq -0.8$.

We conclude that, while $\Delta H_0^{\mathrm{loc}}$ follows the expected trend with the local mass density contrast δ, the underdensity required to ameliorate the discrepancy between the local and global measurements of H_0 far exceeds the values of δ seen in simulations or in direct observations.

6 DISCUSSION

We now discuss whether there are sources of sample variance not captured in our simulations that could significantly increase the sample variance estimated in Section 4 ($\sigma(\Delta H_0^{\mathrm{loc}}) \simeq 0.31\,\mathrm{km\,s^{-1}Mpc^{-1}}$ for the R16 SN selection).

The first question is whether we have sufficient statistics in our simulation. The total simulation volume, $(8\,h^{-1}\mathrm{Gpc})^3$, captures the effect of sufficiently large modes and is subject to negligible variance due to modes larger than the simulation volume. In addition, we explicitly checked that our results (e.g. the total histogram in Fig. 3) are well converged with respect to the number of sub-volumes, and our choice of 512 sub-volumes is sufficient. This also guarantees that choosing the halo closest to the centre of each sub-volume as
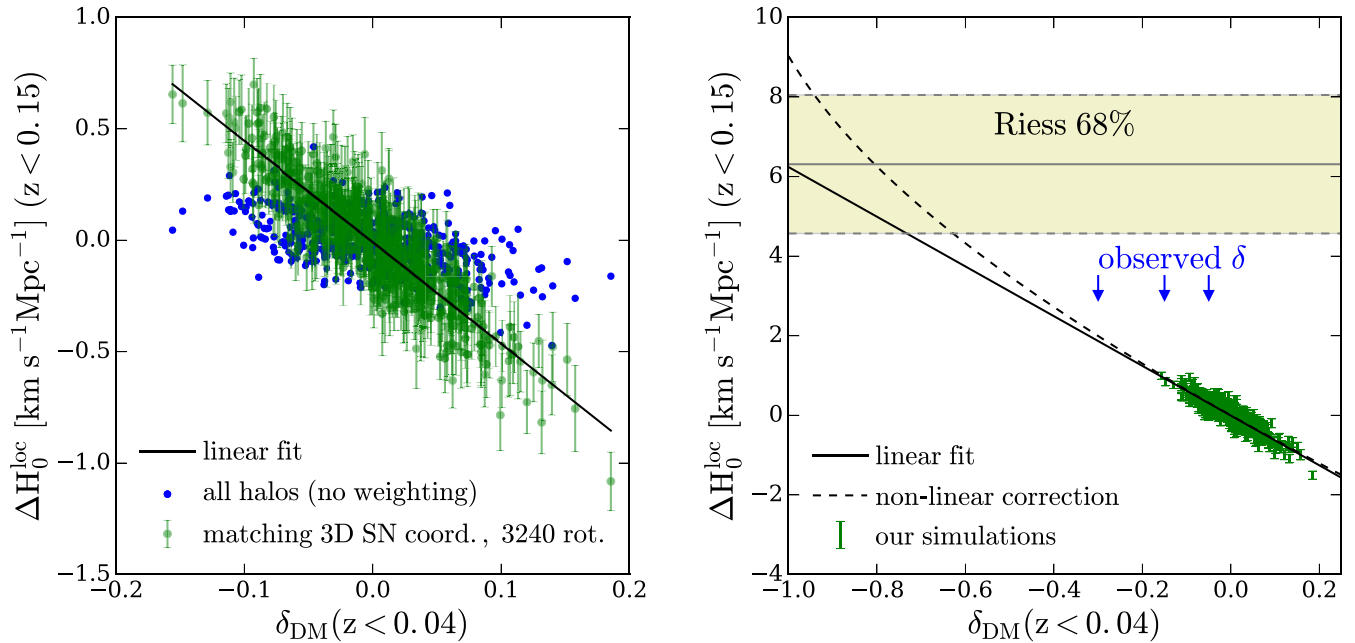
Figure 4. Correlation between ΔH_0^{loc} for $z_{max} = 0.15$ (corresponding to the SN sample) and dark matter density contrast δ for $z_{max} = 0.04$ (corresponding to the distance scale for local density measurements); both are measured from 512 sub-volumes of the Dark Sky simulations. Left: ΔH_0^{loc} measurements from matching the 3D coordinates of SNe and haloes in sub-volumes (green points with error bars), compared to inference from all haloes in sub-volumes (blue points). The error bars on the green points reflect the variances from rotations of the SN coordinate system within each sub-volume. The solid line shows the linear fit to the green points. Right: zoomed-out version of the left-hand panel. We additionally mark the location of several δ values from observations, as well as the 1σ range favoured by the R16 analysis. We note that none of the observations of δ can account for the 6 km s^{-1}Mpc^{-1} difference between H_0^{loc} and H_0^{CMB}.

the location of the observer is sufficient; varying the observer's location prescription within sub-volumes would produce additional volume samples, which are unnecessary given that we already have sufficient statistics. Obtaining more sub-volumes populated by SN, either by choosing alternate locations of the observer in existing sub-volumes or by simply using additional sub-volumes from, e.g. additional runs of Dark Sky simulations, would help populate the far tails of the distribution of the values of H_0^{loc}, and produce a quantitative answer to the extremum-value likelihood of the Hubble constant being greater than some threshold. However, these additional runs would not change anything in our conclusion that sample variance alone falls far short of reconciling the values of the global and local measurements of H_0. In summary, we have sufficiently many sub-volumes for simulating the H_0^{loc} measurements.

Next, we discuss the impact of velocity bias. In our analysis, we assume that the velocity of an SN is equal to the velocity of its host halo and that the velocity of a halo traces the velocity field of the surrounding dark matter. The former assumption holds for Milky Way mass haloes; even if an SN has the maximum circular velocity of its host halo, its motion within the halo is negligible compared with the peculiar velocity of the host halo. The latter assumption may break down due to dynamic friction or tidal stripping (see e.g. Wu et al. 2013); in this work, we explicitly select isolated haloes (haloes that are not within the virial radius of other haloes), and therefore we expect the effect of velocity bias to be negligible. Moreover, an $\sim$5 per cent velocity bias is unlikely to be coherent over scales of hundreds of Mpc, which is required in order to increase the locally inferred value of H_0^{loc}.

Another issue of interest is the choice of the mass of host haloes. In this work, we use haloes with $M_{vir} \in [10^{12.3}, 10^{12.4}]$ M$_\odot$ as

SN hosts. The SN host stellar mass information is partially available for the Supercal sample (Scolnic et al. 2014); it has a fairly wide distribution that peaks around $M_{stellar} \simeq 10^{11}$ M$_\odot$. Given that $M_{vir}/M_{stellar} \sim O(100)$ for this mass range (see e.g. Behroozi, Wechsler & Conroy 2013b), our assumption that the SN host haloes lie in the range $M_{vir} \in [10^{12.3}, 10^{12.4}]$ M$_\odot$ is reasonable. We have nevertheless verified that we get essentially the same results if we match the SNe to all haloes above 10^{13} M$_\odot$.

On the whole, we conclude that our analysis is robust with respect to the statistics of simulated volumes and the halo mass selection. More specifically, our assignment of haloes to SN hosts is faithful to observations, and we do not expect that other choices for the assignment – or in any other step of our calculation – would substantially change our quantitative conclusions about the size of the sample variance in H_0^{loc} measurements.

7 CONCLUSIONS

In this paper, we have studied the sample variance in the local measurements of the Hubble constant. A statistical fluctuation in this measurement, due to the fact that the volume probed is a small fraction of the volume enclosed by the surface of last scattering, is arguably the simplest explanation why the local measurements, H_0^{loc}, would be biased relative to the global value inferred from the CMB, H_0^{CMB}. We have been particularly motivated by the possibility that the spatially inhomogeneous distribution of SNe, from which the recession velocities and thus the Hubble constant are inferred, could further increase the sample variance on top of what has been estimated in the literature. To account for the SN selection, we,

for the first time, explicitly model the exact, inhomogeneous three-dimensional distribution of SNe used in the H_0^{loc} measurements, and we comb through more than one million possible ways in which these SNe could sample the velocity field given by haloes in the 8 h^{-1}Gpc *N*-body Dark Sky simulations.

For observers centred on haloes of virial mass $\sim 10^{12.3}$ M$_\odot$, we find that the measurement of H_0^{loc} is on average unbiased for all scales, while the statistical error depends on the redshift distribution of SNe. Employing the skewed redshift distribution and sparse angular distribution of SNe utilized in R16, we find a final statistical error of $\sigma(\Delta H_0^{\mathrm{loc}}) = 0.31$ km s^{-1}Mpc^{-1}, much smaller than the ~ 6 km s^{-1}Mpc^{-1} discrepancy between the local and global H_0 measurements. Our results are robust with respect to the halo mass range chosen.

We next turn to the relation between the local measurements of the Hubble constant and density contrast. As expected, $\Delta H_0^{\mathrm{loc}}$ and δ measured over the same volume obey the relation from linear perturbation theory if all haloes in the volume are used in the $\Delta H_0^{\mathrm{loc}}$ measurement. However, once the SN-matched haloes are used to mimic the actual measurement, the resulting non-uniform sampling of the volume (and hence the velocity field) spoils the expected $\Delta H_0^{\mathrm{loc}}$–$\delta$ relation. Nevertheless, we find a linear correlation between $\Delta H_0^{\mathrm{loc}}(z_{\mathrm{max}} = 0.15)$ and density measured in a much smaller volume effectively sampled by SN, $\delta(z_{\mathrm{max}} = 0.04)$. We find that resolving the H_0 measurement discrepancy by postulating a local void requires an underdensity of $\delta \simeq -0.8$ with a radius of 120 h^{-1}Mpc, which is extremely unlikely in a ΛCDM universe. Existing direct observational constraints on the mass density at this scale, although highly uncertain, also exclude such a low density.

It is worth noting that R16 explicitly corrected for the peculiar velocity of each SN using a reconstruction of the local density field. This correction, if accurate, explicitly removes the effects of the local over/underdensity, and hence sample variance. Density field reconstructions are subject to a range of systematic errors, and therefore the sample variance correction is necessarily expected to be only partial. Overall, our results are conservative: the sample variance in the measured H_0^{loc} is *at most* the value found in our analysis and likely just a fraction thereof due to explicit peculiar velocity corrections.

In summary, assuming the standard ΛCDM cosmological model, the current tension between H_0^{loc} and H_0^{CMB} is unlikely to be resolved or even appreciably ameliorated by invoking sample variance. Given that the tension between the two measurements cannot be fully attributed to the sample variance of the local measurements, is imperative to understand the systematics in local and CMB experiments, as well as the possible new physics behind the tension.

ACKNOWLEDGEMENTS

REFERENCES

Addison G. E., Huang Y., Watts D. J., Bennett C. L., Halpern M., Hinshaw G., Weiland J. L., 2016, ApJ, 818, 132
Aubourg É. et al., 2015, Phys. Rev. D, 92, 123516
Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013a, ApJ, 762, 109
Behroozi P. S., Wechsler R. H., Conroy C., 2013b, ApJ, 770, 57
Ben-Dayan I., Durrer R., Marozzi G., Schwarz D. J., 2014, Phys. Rev. Lett., 112, 221301
Bland-Hawthorn J., Gerhard O., 2016, ARA&A, 54, 529
Bonvin V. et al., 2017, MNRAS, 465, 4914
Cardona W., Kunz M., Pettorino V., 2017, J. Cosmol. Astropart. Phys., 3, 056
Carrick J., Turnbull S. J., Lavaux G., Hudson M. J., 2015, MNRAS, 450, 317
Cooray A., Caldwell R. R., 2006, Phys. Rev. D, 73, 103002
Courtois H. M., Pomarède D., Tully R. B., Hoffman Y., Courtois D., 2013, AJ, 146, 69
de Vaucouleurs G., Bollinger G., 1979, ApJ, 233, 433
Efstathiou G., 2014, MNRAS, 440, 1138
Feeney S. M., Mortlock D. J., Dalmasso N., 2017, preprint (arXiv:1707.00007)
Fleury P., Clarkson C., Maartens R., 2017, J. Cosmol. Astropart. Phys., 3, 062
Follin B., Knox L., 2017, preprint (arXiv:1707.01175)
Freedman W. L., Madore B. F., 2010, ARA&A, 48, 673
Freedman W. L. et al., 2001, ApJ, 553, 47
Freedman W. L., Madore B. F., Scowcroft V., Burns C., Monson A., Persson S. E., Seibert M., Rigby J., 2012, ApJ, 758, 24
Guy J. et al., 2010, A&A, 523, A7
Hinshaw G. et al., 2013, ApJS, 208, 19
Hubble E., 1929, Proc. Natl. Acad. Sci., 15, 168
Hui L., Greene P. B., 2006, Phys. Rev. D, 73, 123526
Jha S., Riess A. G., Kirshner R. P., 2007, ApJ, 659, 122
Keenan R. C., Barger A. J., Cowie L. L., 2013, ApJ, 775, 62
Marra V., Amendola L., Sawicki I., Valkenburg W., 2013, Phys. Rev. Lett., 110, 241305
Martinez-Vaquero L. A., Yepes G., Hoffman Y., Gottlöber S., Sivan M., 2009, MNRAS, 397, 2070
Odderskov I., Hannestad S., Haugbølle T., 2014, J. Cosmol. Astropart. Phys., 10, 028
Odderskov I., Hannestad S., Brandbyge J., 2017, J. Cosmol. Astropart. Phys., 3, 022
Planck Collaboration XLVI, 2016, A&A, 596, A107
Planck Collaboration XIII, 2016, A&A, 594, A13
Planck Collaboration XVI, 2014, A&A, 571, A16
Riess A. G. et al., 2009, ApJ, 699, 539
Riess A. G. et al., 2011, ApJ, 730, 119
Riess A. G. et al., 2016, ApJ, 826, 56
Sandage A., Tammann G. A., 1982, ApJ, 256, 339
Scolnic D., Kessler R., 2016, ApJ, 822, L35
Scolnic D. et al., 2014, ApJ, 795, 45
Scolnic D. et al., 2015, ApJ, 815, 117
Shi X., Turner M. S., 1998, ApJ, 493, 519
Sinclair B., Davis T. M., Haugbølle T., 2010, ApJ, 718, 1445
Skillman S. W., Warren M. S., Turk M. J., Wechsler R. H., Holz D. E., Sutter P. M., 2014, preprint (arXiv:1407.2600)
Sorce J. G., Tully R. B., Courtois H. M., 2012, ApJ, 758, L12
Suyu S. H. et al., 2013, ApJ, 766, 70
Tammann G. A., Reindl B., 2013, A&A, 549, A136
Turk M. J., Smith B. D., Oishi J. S., Skory S., Skillman S. W., Abel T., Norman M. L., 2011, ApJS, 192, 9
Turner E. L., Cen R., Ostriker J. P., 1992, AJ, 103, 1427
Wang Y., Spergel D. N., Turner E. L., 1998, ApJ, 498, 1
Warren M. S., 2013, preprint (arXiv:1310.4502)
Whitbourn J. R., Shanks T., 2014, MNRAS, 437, 2146
Wojtak R., Knebe A., Watson W. A., Iliev I. T., Heß S., Rapetti D., Yepes G., Gottlöber S., 2014, MNRAS, 438, 1805

Wu H.-Y., Hahn O., Evrard A. E., Wechsler R. H., Dolag K., 2013, MNRAS, 436, 460

Zehavi I., Riess A. G., Kirshner R. P., Dekel A., 1998, ApJ, 503, 483

Zhang B. R., Childress M. J., Davis T. M., Karpenka N. V., Lidman C., Schmidt B. P., Smith M., 2017, preprint (arXiv:1706.07573)

APPENDIX A: IMPACT OF MAXIMUM REDSHIFT AND WEIGHTING

In this appendix, we discuss the weighting of haloes based on $n^{SN}(z)/n^{halo}(z)$. We first demonstrate that, as mentioned in Section 3.2, our choice of halo mass, $[10^{12.3}, 10^{12.4}]$ $M_\odot$, leads to unbiased H_0^{loc} at all scales. Fig. A1 shows how the deviation of the locally measured Hubble constant (ΔH_0^{loc}) depends on the maximum redshift and distance of the SN sample. The blue curves and shades correspond to using all haloes between $z_{min} = 0.023$ and z_{max} (varies along the x-axis), and different curves correspond to the median and the 68 per cent and 95 per cent intervals. Clearly,
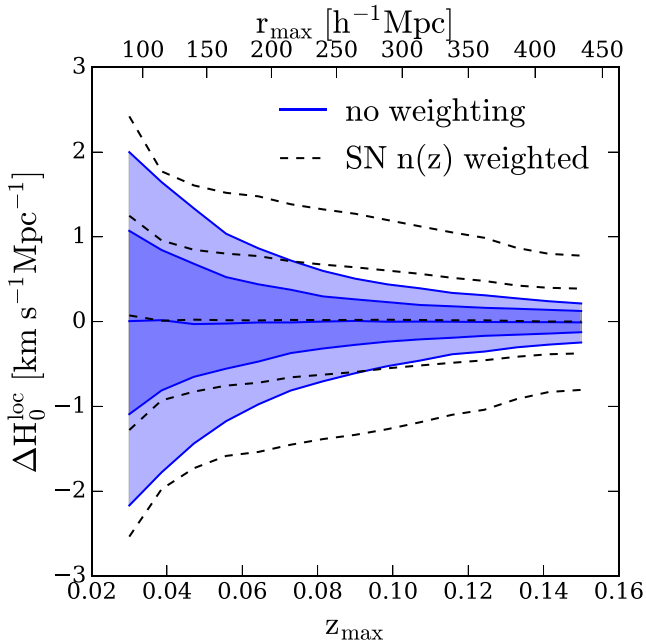
Figure A1. Deviations of the locally measured Hubble constant as a function of maximum redshift (bottom x-axis) and distance (top x-axis). Observers and SNe are centred on haloes within $10^{12.3} < M_{vir}/ M_\odot < 10^{12.4}$. The blue curves and shades correspond to no weighting, and the black dash curves correspond to weighting with the redshift distribution of SNe in R16. The curves correspond to the median and the 68 per cent and 95 per cent intervals. We see that, at this halo mass, H_0^{loc} is essentially unbiased on all scales. Using the realistic redshift distribution leads to a larger statistical error.

when observers are located on haloes of this mass, the median of H_0^{loc} is unbiased for all scales because haloes of this mass occupy approximately unbiased environments. The statistical error of H_0^{loc} drops precipitously with redshift due to the rapidly increasing number of haloes. We note that W14 found $\Delta H_0^{loc} < 0$ because they put observers on more massive haloes, which tend to be in overdense environments.

The black dash curves in Fig. A1 show how ΔH_0^{loc} depends on z_{max} when we weight haloes by $n^{SN}(z)/n^{halo}(z)$. We see that the median remains unbiased, while the statistical error becomes significantly larger than the case of all haloes. The statistical errors decrease slowly with redshift, reflecting the rarity of SN at higher redshift. For $z_{max} = 0.15$, the 68 per cent statistical error is 0.5 per cent.

APPENDIX B: ESTIMATORS OF ΔH_0^{loc}

In Section 2.2, we have derived the estimator for ΔH_0^{loc},

$$\Delta H_0^{loc} = \frac{\sum w_i \mathbf{v_i} \cdot \mathbf{r_i}/r_i^2}{\sum w_i} , \quad (B1)$$

which is relevant for the SN analysis in R16. In this appendix, we discuss the LSQ estimator, which is sometimes seen in the literature.

For $z \ll 1$, the peculiar velocity affects the local measurements of the Hubble constant H_0^{loc} via

$$H_0 r + v_r = H_0^{loc} r . \quad (B2)$$

Let us assume that the ith SN in our sample has a position vector $\mathbf{r}_i$ with respect to the Milky Way, a peculiar velocity $\mathbf{v}_i$ with respect to the CMB rest frame, and weighting w_i. The problem is equivalent to a weighted linear regression with zero intercept; thus, ΔH_0^{loc} is given by the LSQ estimator

$$\Delta H_0^{loc} = H_0^{loc} - H_0 = \frac{\sum w_i \boldsymbol{v}_i \cdot \boldsymbol{r}_i}{\sum w_i r_i^2} \quad (LSQ) . \quad (B3)$$

While the LSQ estimator is the correct one for estimating the slope of a relation (i.e. using v_r and r to estimate ΔH_0^{loc}), it does not apply to SN measurements, given that v_r and r are not separately available. Rather, SNe provide measurements of the specific combination v_r/r, which corresponds to the deviation from the global Hubble constant, and hence the appropriate estimator is given in equation (B1).

Fig. B1 shows the comparison between our estimator and the LSQ estimator for ΔH_0^{loc}. We weight the haloes by $n^{SN}(z)/n^{halo}(z)$ and show ΔH_0^{loc} as a function of the maximum redshift of the survey. Our estimator produces a larger scatter of ΔH_0^{loc} than the LSQ estimator, and the difference increases with redshift. At $z_{max} = 0.15$, our estimator produces approximately three times larger scatter than the LSQ estimator.
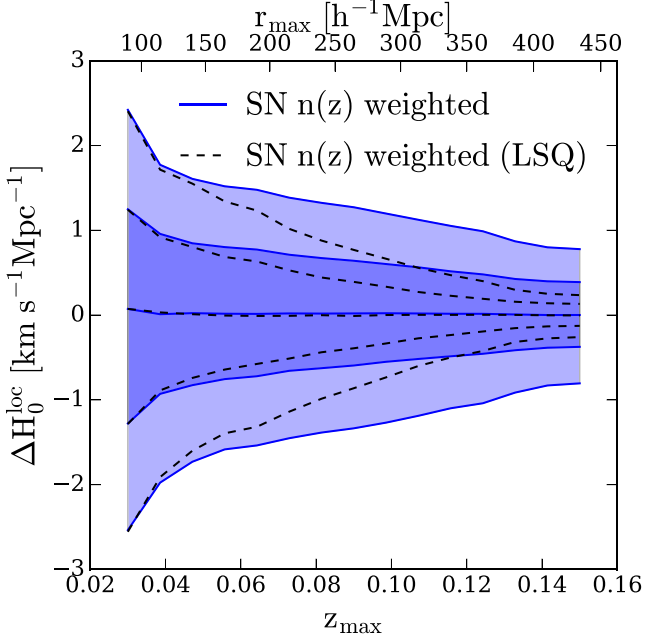
Figure B1. Comparison between our estimator (equation B1) and the LSQ estimator (equation B3), weighted with the redshift distribution of SNe in R16. As can be seen, the LSQ estimator systematically gives smaller ΔH_0^{loc}, and the difference is larger for larger z_{max}.

APPENDIX C: SAMPLING DIFFERENT ORIENTATIONS: EULER ANGLES

Here, we provide more details about varying the orientation of the SN coordinate system relative to the simulation coordinate system. We apply Euler angles to uniformly rotate the coordinates of SNe. Specifically, we use the 'x–z–x' convention; that is, we first rotate the coordinate system around the x-axis by α, second rotate around the z'-axis by β and third rotate the x''-axis by γ.

The ranges of the angles are given by

$$\alpha \in [0, 2\pi)$$
$$\cos \beta \in [1, -1]$$
$$\gamma \in [0, 2\pi) \,. \tag{C1}$$

For α and γ, we use 18 angles spaced by $20°$ ($0°$–$340°$). For β, we use 10 angles with equally spaced $\cos \beta$ (-1 to 1). This leads to a total of $(360/20)^2(180/20 + 1) = 3240$ possible orientations.

This paper has been typeset from a TeX/LaTeX file prepared by the author.