Intrinsic tension in the supernova sector of the local Hubble constant measurement and its implications

Wojtak, Radoslaw; Hjorth, Jens

Published in:
Monthly Notices of the Royal Astronomical Society

DOI:
10.1093/mnras/stac1878

Publication date:
2022

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):
Intrinsic tension in the supernova sector of the local Hubble constant measurement and its implications

Radosław Wojtak and Jens Hjorth

DARK, Niels Bohr Institute, University of Copenhagen, Jagtvej 128, DK-2200 Copenhagen, Denmark

Accepted 2022 June 30. Received 2022 June 23; in original form 2022 May 20

ABSTRACT

We reanalyse observations of Type Ia supernovae (SNe) and Cepheids used in the local determination of the Hubble constant and find strong evidence that SN standardization in the calibration sample (galaxies with observed Cepheids) requires a steeper slope of the colour correction than in the cosmological sample (galaxies in the Hubble flow). The colour correction in the calibration sample is consistent with being entirely due to an extinction correction due to dust with properties similar to those of the Milky Way ($R_V \approx 4.6 \pm 0.4$) and there is no evidence for intrinsic scatter in the SN peak magnitudes. An immediate consequence of this finding is that the local measurement of the Hubble constant becomes dependent on the choice of SN reference colour, i.e. the colour of an unreddened SN. Specifically, the Hubble constant inferred from the same observations decreases gradually with the reference colour assumed in the SN standardization. We recover the Hubble constant measured by SH0ES for the standard choice of reference colour (SALT2 colour parameter $c = 0$), while for a reference colour that coincides with the blue end of the observed SN colour distribution ($c \approx -0.13$), the Hubble constant from Planck observations of the cosmic microwave background (CMB) [assuming a flat Lambda cold dark matter ($\Lambda$CDM) cosmological model] is recovered. These results are intriguing in that they may provide an avenue for resolving the Hubble tension. However, since there is no obvious physical basis for the differences in colour corrections in the two SN samples, the origin of these requires further investigation.

Key words: methods: statistical – cosmological parameters – cosmology: observations – distance scale.

1 INTRODUCTION

The discrepancy between the Hubble constant measured from observations of Type Ia supernovae (SNe) and Cepheids with geometric distance calibrations (Riess et al. 2019, 2021a, b) and from Planck observations of the cosmic microwave background (CMB) assuming a flat Lambda cold dark matter ($\Lambda$CDM) cosmological model (Planck Collaboration VI 2020) has recently drawn much attention. With $H_0 = 73.04 \pm 1.04$ km s$^{-1}$ Mpc$^{-1}$ derived from the most recent advances in Cepheid and SN observations of the SH0ES collaboration (for supernova $H_0$ for the equation of state, Riess et al. 2021b) and $H_0 = 67.36 \pm 0.54$ km s$^{-1}$ Mpc$^{-1}$ inferred from the Planck data (Planck Collaboration VI 2020), it is currently the strongest divergence in cosmological measurements with a confidence level reaching 5$\sigma$. The nature of this tension is unknown and its origin may lie in either unaccounted systematic effects or an incomplete theoretical framework of the standard $\Lambda$CDM cosmological model.

Multiple extensions or modifications of the standard $\Lambda$CDM cosmological model were recently proposed as potential solutions to the Hubble constant tension (Di Valentino et al. 2021). However, none of the current proposals give a satisfying solution without affecting a wide range of other cosmological measurements. Arguably, the most promising scenario involves early dark energy, which is a hypothesized extra energy component manifesting itself before photon decoupling (Poulin et al. 2019). This model is theoretically designed to shorten the sound horizon scale and thus elevate the Hubble constant value derived from the CMB while keeping the observed baryon acoustic oscillation (BAO) angular scale unchanged (Knox & Millea 2020). The main drawback of this proposal is that the Planck data do not provide any statistically significant evidence for early dark energy (Arendse et al. 2020; Vagnozzi 2021; Fondi, Melchiorri & Pagano 2022). Furthermore, early dark energy has also been proven to affect the power spectrum derived from the CMB in a way that spoils a fair consistency between the Planck cosmology and constraints from observations probing large-scale structure (Hill et al. 2020). At the opposite end of the spectrum of cosmological scenarios lie models that attempt to reconcile the local and Planck measurements of the Hubble constant by ad hoc modifications of the very recent expansion history by means of tuning a time-dependent equation of state for dark energy (Di Valentino et al. 2021). This approach, however, ignores the fact that the Hubble constant measured locally is not a direct observable that can be used as a prior in cosmological analyses with Type Ia SN data. Finding an expansion history that interpolates between the local value of the Hubble constant and the Planck cosmology at high redshifts does not resolve the Hubble constant tension, but rather shifts the problem of discordant distance scales to a discrepancy between the absolute luminosity of Type Ia SNe calibrated with Cepheids and its analogue obtained in the so-called inverse distance ladder method based on distance scales calibrated with the Planck data (Camarena & Marra 2021; Efstathiou 2021).

Independent measurements of the Hubble constant on intermediate cosmic distance scales can potentially provide decisive arguments
supporting or ruling out the interpretation of the Hubble constant tension as a cosmological anomaly. Arguably, the best technique operating on these scales is using time delays of gravitationally lensed and multiply imaged variable sources. Although substantial progress has been made in this field, present estimates of the Hubble constant are limited by the accuracy of lens models and range between the SH0ES (Wong et al. 2020) and Planck values (Birrer et al. 2020).

Despite a growing conviction that the Hubble constant tension is a cosmological anomaly, alternative scenarios involving hidden and currently unaccounted systematic effects are not completely ruled out. The majority of studies in this area were undertaken to test the robustness of the local measurement of the Hubble constant with respect to possible changes in modelling the Cepheid data. A wide range of possible systematic effects, including non-standard colour correction (Mortsell et al. 2021), the impact of outliers (Efstathiou 2014), blending, and many other effects (for an exhausting list of tests, see Riess et al. 2021b), were shown to have a negligible impact on the Hubble constant determination (Riess et al. 2021b). The ultimate robustness test of the Cepheid data sector should involve an alternative distance calibration that could replace entirely the Cepheid rod of the distance ladder. Cepheid-independent measurements of the Hubble constant were carried out recently following advances in calibrating cosmological distances with the tip of red giant branch (TRGB; Freedman et al. 2019) or surface brightness fluctuations (SBFs) (Khetan et al. 2021; Garnavich et al. 2022). The results are broadly consistent with the Hubble constant value inferred from Cepheid calibration. However, none of the measurements are currently precise enough to discriminate decisively between the Planck and Cepheid-based values of the Hubble constant.

Compared to the quite exhausting robustness tests of the Cepheid sector, rather little attention has been drawn to Type Ia SNe as a possible source of unknown systematic effects in the local $H_0$ determination. This is worth pursuing considering the purely phenomenological nature of the model used to standardize SN peak magnitudes for distance measurements (Tripp 1998). Moreover, the fact that the apparent intrinsic scatter of $\approx 0.1$ mag in the Hubble diagram with Type Ia SNe found consistently in all independent studies (see e.g. Scolnic et al. 2018; Jones et al. 2019) is comparable to the difference between the local and Planck values of the Hubble constant expressed in distance moduli, i.e. $\Delta \mu \approx 0.16$, is intriguing. The robustness of the Hubble constant measurement based on observations of Type Ia SNe directly relies on accurate distance propagation between SN host galaxies with observed Cepheids (calibration sample) and SN host galaxies in the Hubble flow (cosmological sample). If the currently used SN standardization is not equally accurate in both SN samples, biases may potentially affect the Hubble constant measurement.

The main goal of this study is to quantify to what extent the calibration and cosmological SN samples are consistent with the same universal colour correction, which, alongside the correction due to differences in light-curve shape (Phillips 1993), constitutes the commonly used SN standardization model (Tripp 1998). Our study involves reanalysis of existing observations of Cepheids and Type Ia SNe employing observationally motivated extensions to the standard approach adopted in Riess et al. (2021b). Based on a revised SN colour correction resulting from our analysis, we rederive the Hubble constant and discuss the conditions for resolving the Hubble constant tension.

The outline of the paper is as follows: In Section 2, we describe the data, models, and methods used in our study. The main results are presented in Section 3. This includes detection of an anomaly in the SN data sector of the local Hubble constant determination (Section 3.1), the evidence for discrepant colour corrections in the calibration and cosmological SN samples (Section 3.2), and its impact on the Hubble constant determination (Section 3.3). We discuss the results in Section 4 and summarize our findings in Section 5.

## 2 Data and Model

We use the complete data set, which was the basis for the recent measurements of the Hubble constant presented in Riess et al. (2019, 2021a). The data comprise observations of Cepheids and Type Ia SNe, as well as a range of geometric distance estimates. For the sake of better representation of the data structure in relation to the model, we split the data into seven independent blocks, each described by its own likelihood and the corresponding set of parameters. Table 1 provides a concise description of each data block in terms of data, likelihood formula, and model parameters.

### 2.1 Data blocks

The first three data blocks listed in Table 1 comprise measurements of reddening-free Wesenheit apparent magnitudes, $m_{F160w}^W$, and pulsation periods, $P$, of Cepheids observed in the Milky Way (MW) ($L_{MW}$; Riess et al. 2021a), the Large Magellanic Cloud ($L_{LMC}$; Riess et al. 2019), and 20 galaxies ($L_{cal}$; Riess et al. 2016), including 19 Type Ia SNe host galaxies and the magemaser galaxy NGC 4258. These measurements constrain distance moduli $\mu$ via the following equation:

$$m_{F160w}^W = \mu + M_{F160w} + b_W(\log_{10} P - 1) + z_W \Delta [O/H].$$

(1)

where the Hubble Space Telescope F160W-band absolute magnitude, $M_{F160w}$, and the free coefficients $b_W$ and $z_W$ are measured directly from the data. The additional term in equation (1) incorporates corrections due to metallicity $\Delta [O/H]$, which is directly measured for all MW Cepheids and local environments of Cepheids in galaxies of the calibration sample. Following Riess et al. (2019), we assume that all Cepheids observed in the LMC have the same metallicity equal to the mean $\Delta [O/H] = -0.30$ dex found in the LMC. For metallicity $[O/H]$ measurements provided by Riess et al. (2016), we derive $\Delta [O/H]$ assuming the solar metallicity given by the calibration of these metallicity estimates, i.e. 8.93 (Anders & Greves 1989). Wesenheit magnitudes of Cepheids in the 20 calibration galaxies are calculated from magnitudes $m_{F160w}$ and colours $m_{F555w} - m_{F814w}$, provided by Riess et al. (2016) using the relation

$$m_{F160w}^W = m_{F160w} + R(m_{F555w} - m_{F814w})$$

(2)

with $R = -0.386$ corresponding to an extinction coefficient $R_B = 4.3$ in the reddening law of Fitzpatrick (1999), as assumed in Riess et al. (2019) and Riess et al. (2021a).

Distances to Cepheids in the MW are constrained by Gaia measurements of their parallaxes, $\pi$, (Gaia Collaboration 2021). Following Riess et al. (2021a), we include zero-point $zp$ as a free parameter in order to obtain an unbiased relation between observed parallaxes, $\pi$, and distance moduli using the following equation:

$$\pi + zp = 10^{-0.2(\mu - 10)}$$

(3)

As described explicitly in Table 1, the likelihood for the MW Cepheids accounts for uncertainties in the parallax and magnitude measurements. Here, we also include an error of 0.01 dex in the metallicities. The errors in likelihoods $L_{LMC}$ and $L_{cal}$ are given by the measurement uncertainties in $m_{F160w}$ provided by Riess et al.
Table 1. Summary of data sets, likelihood functions, and parameters used in this study. The Cepheid data block (Cepheids) comprises observations of Cepheids in the Large Magellanic Cloud (LMC), MW, and 20 galaxies, including 19 SN host galaxies and the NGC 4258 megamaser galaxy (cal). The distance anchor data block (Anchors) contains measurements of geometric distances to the LMC and NGC 4258. The SN data block (SNe) comprises light-curve parameters of Type Ia SNe in the 19 calibration galaxies (SN cal) and in the cosmological sample at redshifts 0.023 < z < 0.15 (SN). The notation σx,y denotes measurement uncertainty for quantity x and COVx,y to denote the element of a covariance matrix for variables x and y. The model adopted in this study includes 11 primary parameters: (MF160W, bw, zw) describing Cepheid calibration given by equation (1); zero-pointzp of the Cepheid parallaxes from Gaia; (Mα, α, β, βcal) describing Type Ia SN calibration given by equation (5) (where βcal refers to the calibration sample); (σint, σint cal) describing intrinsic scatter in the corrected magnitudes of SNe, respectively, in the cosmological and calibration samples; and the Hubble constant H0. In addition, the model includes 21 latent parameters that quantify the distance moduli of 19 calibration galaxies (μ1,...,μ19), the NGC 4258 megamaser galaxy (μ20 ≡ μ4258), and the LMC (μLMC).

<table>
<thead>
<tr>
<th>Label</th>
<th>Data</th>
<th>In L</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>LMC</td>
<td>[mF160W, P, σmF160W]μ</td>
<td>−1/2 ∑ (wF160W/μ, P, σmF160W)</td>
<td>MF160W, bw, zw</td>
</tr>
<tr>
<td></td>
<td>Δ[O/H] = −0.3</td>
<td></td>
<td>μLMC</td>
</tr>
<tr>
<td>MW</td>
<td>[μ, σμ, mF160W, σmF160W, P],</td>
<td>−1/2 ∑ (μ, P, σμ, mF160W, σmF160W)</td>
<td>MF160W, bw, zw</td>
</tr>
<tr>
<td>cal</td>
<td>Δ[O/H] = 0.01</td>
<td></td>
<td>zw</td>
</tr>
<tr>
<td>4258</td>
<td>μ4258 = 29.398°</td>
<td>−1/2 [μ(4258) − μ4258]/σμ</td>
<td>μ4258 ≡ μ20</td>
</tr>
<tr>
<td>LMC dist</td>
<td>σμ = 0.032</td>
<td></td>
<td>μLMC</td>
</tr>
<tr>
<td>SNe</td>
<td>{mb, x1, c, COV}μ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SN cal</td>
<td>{mb, x1, c, COV, z}μ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SN</td>
<td>{mb, x1, c, COV, z}μ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COVz = 200 km s−1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. References for data sources: ¹Riess et al. (2019), ²Riess et al. (2021a), ³Gaia Collaboration (2021), ⁴Riess et al. (2016), ⁵Reid et al. (2019), ⁶Pietrzyński et al. (2019), and ⁷Scolnic et al. (2015).

(2016) and Riess et al. (2019). Distance moduli of the LMC and the 20 calibrator galaxies are described by latent variables for which constraints are obtained as a byproduct of the entire analysis combining all likelihoods.

Independent measurements of geometric distances to the LMC from detached eclipsing binaries (Pietrzyński et al. 2019) and to NGC 4258 from megamasers (Reid, Pesce & Riess 2019) are included in likelihoods LLMC and Lcal. Together with parallax distances to Cepheids in the MW, these measurements serve as the anchors of the cosmological distance scale entering the Hubble constant determination.

The last two data blocks in Table 1 include observations of Type Ia SNe. We use light-curve parameters obtained by Scolnic et al. (2015) using the SALT2 fitting methodology (Guy et al. 2005; Scolnic & Kessler 2016). Distance moduli are derived from peak apparent magnitudes, mb, by applying corrections related to the light-curve shape quantified by x1 and colour parameter c, where the latter is thought to combine effectively two physical effects: extinction correction due to intervening dust in SN host galaxies and a possible correlation between the absolute magnitude and the SN intrinsic colour. We employ the standard correction model formulated by Tripp (1998),

\[ m_B = μ + M_B - α x_1 + β c, \]

where B-band absolute magnitude, MB, and correction coefficients α and β are measured directly from the SN data. Errors in distance moduli adopted in the SN likelihoods are given explicitly in Table 1. They include all elements of covariance matrices obtained for individual SNe, intrinsic scatter σint and extra error due to unconstrained peculiar velocities (included only in LSN) with σvel = 200 km s−1 (Carrick et al. 2015). For the cosmological data block (LSN), we selected SNe using the same criteria as those adopted by Riess et al. (2016). The primary selection condition is given by redshift range 0.023 < z < 0.15 and cuts in light-curve parameters: |c| < 0.3 and |x1| ≤ 3. In addition, SNe with low-quality fits are rejected. SNe passing the fit quality check (223) are characterized by sufficient goodness of fit (fitprob > 0.001) and well-constrained light-curve parameters with errors < 1.5 for x1, < 2 d for the peak time and < 0.2 mag for the corrected peak magnitude.

Finally, as a novel approach, in this study, we will allow for SNe in the calibration block to be characterized by an independent slope.
\( \beta_{\text{cal}} \), of the colour correction, i.e.

\[
m_B = \mu + M_B - \alpha x_1 + \beta_{\text{cal}} c,
\]

and associated intrinsic scatter, \( \sigma_{\text{int,cal}} \), in \( m_B \). These extra parameters will enable us to test the standard approach of assuming universality of the colour correction (\( \beta_{\text{cal}} = \beta \)) and intrinsic scatter (\( \sigma_{\text{int}} \approx \sigma_{\text{int,cal}} \)) and explore its impact on the local \( H_0 \) determination.

### 2.2 Parameter estimation

We constrain model parameters using the likelihood function, which is a product of likelihoods from all seven data blocks, i.e.

\[
L \propto L_{\text{LMC}} \times L_{\text{MW}} \times L_{\text{c,1528}} \times L_{\text{SN,cal}} \times L_{\text{SN}}.
\]

Our model is described by 11 primary parameters (\( H_0, M_{\text{F160W}}, b_W, z_W, \beta, M_B, \alpha, \beta_{\text{cal}}, \sigma_{\text{int,cal}} \)) out of which two (\( \beta_{\text{cal}} \) and \( \sigma_{\text{int,cal}} \) — the colour correction slope and intrinsic scatter in the SN calibration sample) are optional. In addition, 21 latent variables are constrained as byproduct of fitting. These are the distance moduli of the 19 calibration galaxies, the LMC, and the megamaser galaxy NGC 4258. Including these latent variables as extra nuisance parameters in the analysis is not relevant for constraining the primary parameters and one can alternatively employ a compressed likelihood obtained by analytical integration of \( L \) over all latent variables. However, having access to the latent variables enables a range of sanity tests aimed at checking consistencies between different data blocks given the best-fitting model. In particular, in our study, we scrutinize the intrinsic consistency between the calibration (\( L_{\text{SN,cal}} \)) and cosmological (\( L_{\text{SN}} \)) SN data blocks.

We use a Markov Chain Monte Carlo technique to integrate the posterior probability function and find best-fitting model parameters. The chains are computed with the emcee code (Foreman-Mackey et al. 2013). Unless explicitly stated, best-fitting parameters are provided as the posterior mean values and errors given by 16th and 84th percentiles of the marginalized probability distributions. The 1σ and 2σ confidence contours shown in all figures contain 68 and 95 per cent, respectively, of the corresponding two-dimensional (2D) marginalized probability distributions.

The Hubble constant is derived assuming that the shape of the distance modulus as a function of redshift, which is independent of \( H_0 \), is given by the Planck cosmological model (Planck Collaboration VI 2020). The respective formula for distance modulus reads

\[
\mu(z) = \mu_{\text{Planck}}(z) + 5 \log g_0 (H_0_{\text{Planck}} / H_0),
\]

where all quantities with a Planck subscript are given by the Planck model, while \( H_0 \) is a free parameter in our analysis (see also the SN block in Table 1). For the adopted SN redshift range in the cosmological data block, this strategy is in practice equivalent to the commonly used third-order approximation of cosmological distance as a function of redshift with fixed values of the deceleration parameter \( q_0 \) and the jerk parameter \( j_0 \) (Visser 2004). Small variations in \( q_0 \) and \( j_0 \) within a wide range of dark matter and dark energy density parameters have been shown to have a negligible impact on the current estimation of the Hubble constant (Riess et al. 2021b).

### 3 ANALYSIS

#### 3.1 Baseline model

To validate our approach, we begin by fitting a model that resembles closely the fitting strategy adopted by Riess et al. (2016) or the baseline model in Riess et al. (2021b). The model includes the minimum number of primary parameters which are necessary to fit the Cepheid data (\( M_{\text{F160W}}, b_W, z_W, \beta_p \) and SN data (\( M_B, \alpha, \beta, \sigma_{\text{int}}, H_0 \)). Here, we assume that the colour correction coefficient and intrinsic scatter of SN corrected magnitudes are the same in the SN calibration (\( L_{\text{SN,cal}} \)) and cosmological (\( L_{\text{SN}} \)) data blocks, i.e. \( \beta = \beta_{\text{cal}} \) and \( \sigma_{\text{int}} = \sigma_{\text{int,cal}} \).

Our best-fitting baseline model recovers all essential results obtained for the same data compilation in the original studies. In particular, we find excellent agreement with the Cepheid parameters (\( M_{\text{F160W}}, b_W, z_W \) measured by Riess et al. 2021b), the zero-point \( z_p \) of the Cepheid parallaxes from Gaia determined by Riess et al. (2021a), and the SN calibration parameters (\( M_B, \alpha, \beta, \sigma_{\text{int}} \)) found for similar low-redshift \( z \leq 0.1 \) SN samples (e.g. Jones et al. 2019). The best-fitting Hubble constant \( H_0 = (73.14 \pm 1.31) \text{ km s}^{-1} \text{ Mpc}^{-1} \) is fully consistent with the original measurements (Riess et al. 2019, 2021a) based on the same calibration sample and the most recent updates in distance anchors as summarized in Riess et al. (2021b). Our measurement is at a 1σ discrepancy with the Hubble constant determination from Planck observations assuming a flat \( \Lambda \)CDM cosmological model (Planck Collaboration VI 2020). This tension increases to a 3σ level in Riess et al. (2021b) primarily due to ~30 per cent smaller errors resulting from the twice as large calibration sample.

We use the byproduct constraints on the distance moduli of the calibration galaxies (\( \mu_1, \cdots, \mu_{11} \) in Table 1) to check if the distribution of residuals in the SN block confirms that the baseline model provides a complete and unbiased description of the data. Our study reveals what appears to be an anomalous relation between residual distance moduli of the SNe in the calibration sample and the SN colour parameter (see the left-hand panel in Fig. 1). This trend suggests that the universality of colour corrections assumed in the baseline model is inconsistent with the SN data. The apparent overestimation of the distances of red SNe \( (c > 0) \) and the corresponding distance underestimate of blue SNe \( (c < 0) \) appear to be comparable to the intrinsic scatter found in cosmological SN samples. This means that the anomaly can be concealed in the intrinsic scatter of the baseline model. This made it particularly difficult to detect in previous analyses. Fig. 1 also shows that observed SN colours are distributed similarly in the calibration and cosmological samples and are consistent with the corresponding colour distribution from the most recent SN compilation (Pantheon+; Brout et al. 2022).

We quantify the statistical significance of the trend shown in Fig. 1 by fitting a linear model. Keeping in mind that the apparent discrepancy between the SN colour correction in the calibration sample and the cosmological sample can also worsen the fit in the SN cosmological sample, although to a lesser extent, we expect that this approach gives us a lower limit of the actual significance of the anomaly. A complete and rigorous analysis is presented in the following subsection, where we compare the baseline model to its minimum extensions motivated by the new trend found in the data. Taking into account uncertainties in both variables and assuming that they are uncorrelated, we find a positive correlation between residual distance moduli \( \Delta \mu \) and SN colours \( c \) (Fig. 1) at the ~3σ significance level, with a linear slope of \( 1.10^{+0.32}_{-0.30} \). The slope is \( 0.93^{+0.41}_{-0.40} \) if intrinsic scatter is included as an extra free nuisance parameter.

#### 3.2 Extensions to the baseline model

We now consider three models which allow for independent colour corrections in the calibration and cosmological samples. The intrinsic scatter in the calibration sample is assumed to be either an extra
The independent parameter (model A), vanishing (σ_{int,cal} ≡ 0, model B), or equal to the analogous scatter in the cosmological sample (σ_{int,cal} ≡ σ_{int,mod C}). We summarize the best-fitting SN parameters in Table 2. All modifications with respect to the baseline model occur for parameters which are relevant for both SN data blocks, while Cepheid calibration parameters remain virtually unchanged and are provided in Table A1 (appendix) for the sake of completeness. The constraints on SN parameters (M_B, α, β, σ_{int,cal}, β_{cal,σ_{int,cal}}) and the Hubble constant in model A are shown in Fig. 2. The red lines indicate combinations of parameters reducing model A to the baseline model, i.e., β = β_{cal} and σ_{int} = σ_{int,cal}. The right-hand panel of Fig. 1 demonstrates explicitly how the anomalous trend in distance modulus residuals obtained in the baseline model vanishes in model B (the most preferred by the data, as explained below).

While the three models (A, B, and C) treat the intrinsic scatter in the peak magnitudes in the calibration sample differently, they all consistently show that the slope of the SN colour correction, β_{cal}, in the calibration sample is larger than in the cosmological sample β. We find β_{cal} − β = 1.23 ± 0.54 for model A, β_{cal} − β = 1.51 ± 0.40 for model B, and β_{cal} − β = 1.05 ± 0.56 for model C. Fig. 3 shows the marginalized posterior distribution for β_{cal} − β obtained for the three models. Numerical estimation of the probability that β_{cal} ≤ β yields 0.017 (2.4σ) for model A, 10^{-4} (3.8σ) for model B, and 0.025 (2.2σ) for model C.

Table 2 provides goodness of fit quantified by the maximum likelihood and the Bayesian information criteria (BIC). Since all three extensions to the baseline model are constrained solely by the data in the SN calibration sample, we compute both metrics using ln L_{SN,cal} from the corresponding data block. We find strong evidence favouring model B (ΔBIC ≪ −6) over the baseline model, model A yields a substantially better fit than the baseline model, but its free intrinsic scatter σ_{int,cal} turns out to be a redundant parameter (no detection: the maximum likelihood found at σ_{int,cal} = 0 and fit yielding merely an upper bound, see Table 2 and Fig. 2), and for this reason, the model is less preferred than model B. The predictive power of model C is substantially reduced by the fact that a large part of the anomaly is effectively absorbed by intrinsic scatter, which is fixed at the value inferred from the cosmological sample. Consequently, model C is less favoured than the baseline model in terms of BIC, although it is fully consistent with the data.

The strongest discrepancy between β_{cal} and β is found for model B, which is the most favoured by the data among the three extensions to the baseline model considered here. The vanishing intrinsic scatter in the calibration sample assumed in this model indicates that independent colour correction results in maximally improved precision of the SN standardization in the calibration sample. This can also be realized by fitting a model with a universal colour correction (β = β_{cal}) and free σ_{int,cal}. The fit yields σ_{int,cal} = 0.135 ± 0.034, demonstrating directly that the assumption of universal SN colour correction substantially degrades the precision of the SN standardization in the calibration sample. It is also worth emphasizing that model B has greater predictive power than the baseline model in a broader context. As we discuss in the following section, the new SN colour correction in the calibration sample can be straightforwardly linked to the standard dust extinction known from the MW.
Table 2. Best-fitting primary parameters of the SN data block measured from observations of Cepheids and Type Ia SNe assuming the baseline model and its three extensions: model A (free $\beta_{\text{cal}}$ and $\sigma_{\text{int,cal}}$), model B (free $\beta_{\text{cal}}$ and $\sigma_{\text{int,cal}}$ ≡ 0), and model C (free $\beta_{\text{cal}}$ and $\sigma_{\text{int,cal}}$ ≡ $\sigma_{\text{int}}$). Best-fitting results are provided in the form of the posterior mean values and errors containing 68% per cent of marginalized probabilities (with 68% per cent upper limit for $\sigma_{\text{int,cal}}$ in model A). The Hubble constant for models A, B, and C is derived by assuming reference colour $c_{\text{ref}} = 0$ in equation (8). In general, the Hubble constant determination depends on the choice of $c_{\text{ref}}$ in these cases, and the Planck value is obtained for $c_{\text{ref}} \lesssim \mathrm{0.13}$, as discussed in Section 3.3. Since extra parameters in the new models are constrained solely by the data from the SN calibration block, the goodness of fit is quantified in terms of the corresponding maximum likelihood $\Delta \ln L_{\text{SN,cal,max}} = \ln L_{\text{SN,cal,max}}(\text{new model}) - \ln L_{\text{SN,cal,max}}(\text{baseline model})$ and the Bayesian Information Criterion $\Delta \text{BIC} = \Delta \sigma \ln(19) - 2 \Delta \ln L_{\text{SN,cal,max}}$, where $\Delta \sigma = 2, 1, 1$ for models A, B, and C, respectively, and $N = 19$ is the number of SNe in the SN calibration sample. Best-fitting parameters of the Cepheid data block are listed in Table A1.

<table>
<thead>
<tr>
<th>SN parameters</th>
<th>Baseline</th>
<th>Model A</th>
<th>Model B (most favoured)</th>
<th>Model C</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>3.109$^{+0.112}_{-0.112}$</td>
<td>3.060$^{+0.114}_{-0.115}$</td>
<td>3.057$^{+0.114}_{-0.113}$</td>
<td>3.062$^{+0.114}_{-0.113}$</td>
</tr>
<tr>
<td>$\sigma_{\text{int}}$</td>
<td>0.115$^{+0.008}_{-0.008}$</td>
<td>0.115$^{+0.008}_{-0.008}$</td>
<td>0.115$^{+0.008}_{-0.008}$</td>
<td>0.115$^{+0.008}_{-0.008}$</td>
</tr>
<tr>
<td>$\beta_{\text{cal}}$</td>
<td>$\equiv \beta$</td>
<td>$\equiv \beta$</td>
<td>$\equiv \beta$</td>
<td>$\equiv \beta$</td>
</tr>
<tr>
<td>$\sigma_{\text{int,cal}}$</td>
<td>$\equiv \sigma_{\text{int}}$</td>
<td>$\equiv \sigma_{\text{int}}$</td>
<td>$\equiv \sigma_{\text{int}}$</td>
<td>$\equiv \sigma_{\text{int}}$</td>
</tr>
<tr>
<td>$M_B$</td>
<td>$-19.245^{+0.038}_{-0.037}$</td>
<td>$-19.223^{+0.037}_{-0.037}$</td>
<td>$-19.216^{+0.032}_{-0.032}$</td>
<td>$-19.228^{+0.040}_{-0.040}$</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>$0.129^{+0.008}_{-0.009}$</td>
<td>$0.129^{+0.008}_{-0.009}$</td>
<td>$0.129^{+0.008}_{-0.009}$</td>
<td>$0.129^{+0.008}_{-0.008}$</td>
</tr>
<tr>
<td>$H_0$ [km s$^{-1}$ Mpc$^{-1}$]</td>
<td>73.142$^{+1.315}_{-1.306}$</td>
<td>73.864$^{+1.285}_{-1.283}$</td>
<td>74.101$^{+1.124}_{-1.137}$</td>
<td>73.708$^{+1.419}_{-1.413}$</td>
</tr>
</tbody>
</table>

Goodness of fit

| $\Delta \ln L_{\text{SN,cal,max}}$ | 0 | 6.72 | 6.72 | 1.09 |
| $\Delta \text{BIC}_{\text{SN,cal}}$ | 0 | $-7.55$ | $-10.49$ | 0.77 |

3.3 Hubble constant

When defining the SN absolute magnitude $M_B$ and the corresponding standardized magnitudes, we have the freedom to choose an arbitrary reference colour parameter $c_{\text{ref}}$, i.e.

$$m_B = \mu + M_B + \alpha x_1 + \beta (c - c_{\text{ref}})$$

for SN block,

$$m_B = \mu + M_B + \alpha x_1 + \beta_{\text{cal}} (c - c_{\text{ref}})$$

for SN cal block. (8)

Changing $c_{\text{ref}}$ in the baseline model ($\beta = \beta_{\text{cal}}$) automatically modifies $M_B$ inferred from the data, but it does not affect the remaining parameters, including $H_0$. From this point of view, the commonly adopted $c_{\text{ref}} = 0$ is an arbitrary choice which does not have any implications for the Hubble constant determination and can be merely motivated by the minimization of a correlation between $M_B$ and $\beta$. However, this situation changes when the colour correction is not universal and differs between the calibration and cosmological SN samples, i.e. $\beta \neq \beta_{\text{cal}}$. In this case, adopting $c_{\text{ref}} \neq 0$ changes the SN absolute magnitude (with respect to the $c_{\text{ref}} = 0$ case) differently in the calibration and cosmological samples, by $\Delta M_B = \beta_{\text{cal}} c_{\text{ref}}$ and $\Delta M_B = \beta c_{\text{ref}}$, respectively. The only way to reconcile SN absolute magnitudes in both SN samples is to adjust the Hubble constant so that the difference in $\Delta M_B$ is fully compensated by the corresponding shift in distance modulus $\Delta \mu$,

$$\Delta \mu = (\beta - \beta_{\text{cal}}) c_{\text{ref}}.$$  

(9)

This means that every choice of $c_{\text{ref}}$ has its own unique best-fitting value of the Hubble constant inferred from the same observational data.

The values of the Hubble constant measured for models A, B, and C and listed in Table 2 are obtained for an SN absolute magnitude defined for the same reference colour as in the baseline model ($c_{\text{ref}} = 0$). Since the reference colour is very close to the median colour in both SN samples, the resulting constraints on the Hubble constant agree fairly well with the result from the baseline model. For small reference colours, i.e. $|c_{\text{ref}}| < 1$, we derive from equation (9) that the best-fitting Hubble constant shifts by $\Delta H_0$ with respect to its value obtained for $c_{\text{ref}} = 0$, as given by the following equation:

$$\Delta H_0/H_0 \approx (\ln 10/5)(\beta_{\text{cal}} - \beta) c_{\text{ref}}.$$  

(10)

Using the above approximation, we can show that models A, B, or C can recover the Planck value of the Hubble constant for $c_{\text{ref}} \approx -0.16, -0.13, -0.18$, respectively. The reference colours required to obtain the Planck value of the Hubble constant coincide with the blue end of the observed colour distribution in the cosmological and calibration samples. Interestingly, the reference colour for the preferred model B is within the $1\sigma$ range of the intrinsic colour distribution derived from the Pantheon+ SN sample, assuming an exponential distribution of dust reddening (Popovic et al. 2021).

Following the above estimates of $c_{\text{ref}}$ required for a reduction of the Hubble constant to the Planck value, we repeat the full analysis with models A, B, and C for a range of $c_{\text{ref}}$ values between $-0.175$ and $0.025$. The only modification of the corresponding likelihoods listed in Table 1 occurs in the two SN blocks where the reference colour modifies SN distance moduli according to equation (8). Fig. 4 shows the resulting constraints on the Hubble constant as a function of the reference colour. We find that all three models reduce the Hubble constant tension to sub-2$\sigma$ confidence levels when $c_{\text{ref}} \approx -0.06$ and recover the Planck measurement when $c_{\text{ref}} \approx -0.15$. The strongest modification of the best-fitting $H_0$ as a function of $c_{\text{ref}}$ occurs for model B ($\sigma_{\text{int,cal}} = 0$), which yields the largest difference between the colour correction slopes in the calibration sample and cosmological sample. The apparent reduction of the Hubble constant tension appears primarily due to gradually decreasing best-fitting $H_0$ and to a much lesser extent due to larger errors. The latter results from a degeneracy between the colour correction slope and $H_0$. This degeneracy is minimized at reference colour $c_{\text{ref}} \approx -0.02$, at which a correlation between $\beta_{\text{cal}}$ and $M_B$ happens to vanish for this particular SN calibration sample, but it gradually increases for $c_{\text{ref}} < -0.02$ or $c_{\text{ref}} > -0.02$. The effect of vanishing correlation between $\beta_{\text{cal}}$ and $M_B$ is visible in Fig. 4 as the smallest errors and differences between best-
fitting $H_0$ values at $c_{\text{ref}} \approx -0.02$. We find that the errors in $H_0$ measurements for $c_{\text{ref}} = -0.125$ are $\sim 60$ per cent larger than for $c_{\text{ref}} = 0$.

## 4 DISCUSSION

Our analysis shows that the assumption of a universal colour correction expressed in equation (4) does not hold for the SN sample used in the local determination of the Hubble constant. Colour correction in the calibration sample requires a steeper slope than for the remaining SNe in the Hubble flow. A direct implication of this finding is that the Hubble constant determination becomes dependent on the choice of a reference colour, with the Planck value recovered for $c_{\text{ref}} \approx -0.13$ (model B). This new property is entirely inferred from the observational data.

As shown in Fig. 4, the Hubble constant tension decreases to just $1\sigma$ when the adopted reference colour in equation (8) is the...
mean intrinsic colour of Type Ia SNe, as measured recently from the Pantheon+ sample (Popovic et al. 2021). It is also important to notice that the relative colour \( c - c_{\text{ref}} \) in the calibration sample becomes strictly positive (see Fig. 1). With this mind, it is natural to suspect that the colour correction in this case is predominantly driven by dust extinction. This picture is consistent with the probabilistic model of SN colours used by Popovic et al. (2021) for which the red tail of the observed colour distribution arises entirely from reddening due to dust. Following this interpretation, we conclude that a low, nearly Planckian value of the Hubble constant (or a longer distance scale) obtained in our analysis for \( c_{\text{ref}} \approx -0.13 \) results most likely from a stronger dust extinction and thus higher intrinsic brightness (\( \Delta m \approx 0.17 \) for \( c_{\text{ref}} = -0.13 \) and model B) of SNe in the calibration sample than in the baseline model.

Our analysis recovers the *Planck* value of the Hubble constant for reference colour \( c_{\text{ref}} \approx -0.13 \) (the most favoured model B), which happens to coincide with the bluest colour of SNe in both the calibration and cosmological samples. This brings us to a proposal in which full agreement with the *Planck* cosmology can be restored when we assume that intrinsic SN colour is close to \(-0.13\) (with negligible scatter), while redder SN colours result simply from reddening by intervening dust. The required intrinsic colour can be obtained from SN observations following a similar approach to that presented in Popovic et al. (2021) and Brout & Scollnic (2021), but with a modified model describing the distribution of colour excess, \( E(B - V) \). Popovic et al. (2021) employs an exponential distribution of \( E(B - V) \) proposed by Mandel, Narayan & Kirshner (2011). This model can be justified as a maximum entropy solution given the mean value of the colour excess. However, this is only a motivation based on information theory, which does not necessarily reflect any physical constraints such as the 3D distribution of dust and SNe in galaxies. In fact, the distribution of apparent colours seen in Fig. 1 can be reproduced by a wider range of combinations of intrinsic colour and reddening distributions. In particular, the simplest solution motivated by minimizing the Hubble constant tension is a single-valued intrinsic colour \( c_{\text{ref}} \approx -0.13 \) (or a very narrow distribution with the mean value equal to \( c_{\text{ref}} \)) and a colour excess distribution given by \( c - c_{\text{ref}} \). We think that this scenario is worth further exploration as a possible revision of the exponential distribution of dust reddening assumed in the present Bayesian models (see e.g. Mandel et al. 2017; Thorp et al. 2021). One should also keep in mind that the intrinsic colour of an SN may depend on other intrinsic properties (e.g. Foley, Sanders & Kirshner 2011).

Assuming that the colour excess with respect to reference colour \( c_{\text{ref}} \approx -0.13 \) arises predominantly as dust reddening, i.e. \( c - c_{\text{ref}} = E(B - V) \), we can interpret \( \beta \) (or \( \beta_{\text{cal}} \)) in our analysis with models A, B, or C as the extinction coefficient \( R_E \). In this picture, galaxies in the calibration sample appear to be analogues of the MW in terms of their dust properties. The extinction coefficient in the calibration sample determined in our analysis (\( R_{E} \approx 4.6 \pm 0.4 \) for model B in Table 2) is fully consistent with the mean extinction coefficient measured in the MW, i.e. \( R_{E} \approx 4.3 \) (Cardelli, Clayton & Mathis 1989; Fitzpatrick 1999). Furthermore, MW-like extinction in the calibration sample readily improves the precision of the SN standardization decreasing the intrinsic scatter in distance moduli from \( \sigma_{\text{int-cal}} = 0.135 \pm 0.034 \) to \( \sigma_{\text{int-cal}} < 0.097 \). Finally, the measured extinction coefficient in the calibration sample ensures full consistency with the extinction correction applied to Cepheid observations, which is based on \( R_{E} = 4.3 \) (Riess et al. 2019, 2021a).

The above dust-based interpretation of SN colours is less straightforward in the cosmological sample. SN host galaxies in the Hubble flow are characterized by a substantially lower coefficient of the
colour correction with $\beta = 3.06 \pm 0.11$. This result is consistent with many other studies of various SN cosmological sample (see e.g. Scolnic et al. 2018; Jones et al. 2019) and it has been a long-lasting problem whether it can be interpreted as extinction using physical models of dust. Although Bayesian analysis of Type Ia SNe suggests that a broad distribution of $R_B$ spanning between $R_B \approx 2$ and $R_B \approx 5$ is favoured by SN data (as a means to reduce an appreciable fraction of scatter in the Hubble diagram; see Brout & Scolnic 2021; Popovic et al. 2021), a complete physical picture connecting dust properties and extinction inferred from observations is still missing. In particular, the extinction coefficient measured in the MW falls into a rather wide range $3.5 \leq R_B \leq 6.5$ (Cardelli et al. 1989), while values of $2$–$3$ are not seen. The problem becomes particularly pressing in the context of our study in which the apparent differences between the extinction properties in galaxies of the calibration and cosmological samples can be used to reconcile the local and CMB-based measurements of the Hubble constant.

5 SUMMARY AND CONCLUSIONS

We have reanalysed Cepheid and Type Ia SN observations used in the most precise local determination of the Hubble constant to date to test the universality of the commonly used phenomenological colour correction for SN standardization given by equation (5) proposed by Tripp (1998). Our main results are as follows:

(i) SN data in the calibration sample (galaxies with observed Cepheids) and cosmological sample (galaxies in the Hubble flow) are inconsistent with the assumption of a universal colour correction. Standardization of SNe in the calibration sample requires a steeper slope found at $2.2 \sigma$ confidence level assuming universal intrinsic scatter, at $2.4 \sigma$ confidence level when fitting intrinsic scatter independently in the calibration sample, and at $3.8 \sigma$ confidence level assuming vanishing intrinsic scatter in the calibration sample. The latter model maximizing the difference between the SN colour corrections in the two SN samples is the most favoured by the data. Accounting for the SN colour correction inferred directly from the SN calibration data eliminates the necessity for including intrinsic scatter in the corrected SN peak magnitudes.

(ii) The difference between SN colour corrections in the calibration and cosmological samples inevitably makes the Hubble constant measurement dependent on the choice of reference colour setting the absolute magnitude in both SN samples. While the Hubble constant value of Riess et al. (2021b) is recovered for $c_{\text{ref}} = 0$, we find that gradually lower values are measured when using gradually bluer reference colour, i.e. $c_{\text{ref}} < 0$. We recover the Planck value for $c_{\text{ref}} \approx -0.13$, which happens to coincide with the blue end of the apparent colour distribution.

(iii) The slope of the SN colour correction in the calibration sample coincides numerically with the mean extinction coefficient found in the MW. This suggests that galaxies in the calibration sample – unlike SN host galaxies in the cosmological sample – are analogues to the MW in terms of their dust extinction properties.

(iv) The minimum physical scenario required to obtain the Planck value of the Hubble constant assumes that the SN intrinsic colour is $-0.13$ (with relatively small scatter) and the observed colour distribution results predominantly from dust reddening with MW-like extinction in the calibration sample.

Our study opens up a new avenue for understanding the physical origin of the Hubble constant tension. In this proposal, the tension arises from the insufficiently accurate standardization of Type Ia SNe resulting from a poor understanding of dust extinction and SN intrinsic colours. Further investigation and perhaps more observations will be needed to test this scenario and address arising questions. Analysis of alternative calibration samples based on TRGB or SBF distance calibrations will allow us to verify whether the new colour correction shown in our study is a special property of host galaxies with well-observed Cepheids or a generic property of SN host galaxies in the nearby universe. Additional tests can be enabled by spectroscopic observations of the local SN environments, which provide independent constraints on dust content along the lines of sight to SNe. In general, the empirically determined SN colour correction can result from two physical effects: reddening due to dust and/or a possible modulation related to SN intrinsic colour. Disentangling these effects will be essential for understanding the colour correction from first principles. This will perhaps require a better synergy between advanced data analysis methods and physically motivated models of dust extinction and light emission by Type Ia SNe, following the recent studies by Mandel et al. (2017), Thorp et al. (2021), and Popovic et al. (2021). Another avenue worth pursuing is observations of SNe in near-infrared, which minimize the effect of dust extinction. Recent studies reported the Hubble constant estimates consistent with the SH0ES value (Burns et al. 2018; Dhawan, Jha & Leibundgut 2018; Jones et al. 2022). However, more effort is needed to reduce the measurement errors, which are currently too large to confirm conclusively the Hubble constant tension.

ACKNOWLEDGEMENTS

This work was supported by a VILLUM FONDEN Investigator grant (project number 16599). RW thanks Adriano Agnello, Charlotte Angus, Christa Gall, Darach Watson, Luca Izzo, and Nandita Khetan for discussions, and David Jones for constructive comments. The authors thank the anonymous referee for constructive comments that helped improve this work.

DATA AVAILABILITY

No new data were generated or analysed in support of this research.

REFERENCES

Brouss D. et al., 2022, preprint (arXiv:2202.04077)
Di Valentino E. et al., 2021, Class. Quantum Gravity, 38, 153001

MNRAS 515, 2790–2799 (2022)
Appendix A: Best-Fitting Parameters of the Cepheid Data

Table A1. Best-fitting parameters of the Cepheid data block measured from observations of Cepheids and Type Ia SNe assuming the baseline model and its three extensions: model A (free $\beta_{\text{cal}}$ and $\sigma_{\text{cal}}$), model B (free $\beta_{\text{cal}}$ and $\sigma_{\text{cal}} \equiv 0$) and model C (free $\beta_{\text{cal}}$ and $\sigma_{\text{cal}} \equiv \sigma_{\text{cal}}$). Best-fitting results are provided in the form of the posterior mean values and errors containing 68 per cent of marginalized probabilities.

<table>
<thead>
<tr>
<th>Cepheid parameters</th>
<th>Baseline</th>
<th>Model A</th>
<th>Model B (most favoured)</th>
<th>Model C</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_{V,160}$</td>
<td>$-5.929^{+0.014}_{-0.014}$</td>
<td>$-5.929^{+0.014}_{-0.014}$</td>
<td>$-5.929^{+0.014}_{-0.014}$</td>
<td>$-5.929^{+0.014}_{-0.014}$</td>
</tr>
<tr>
<td>$b_W$</td>
<td>$-3.292^{+0.012}_{-0.012}$</td>
<td>$-3.292^{+0.012}_{-0.012}$</td>
<td>$-3.292^{+0.012}_{-0.012}$</td>
<td>$-3.292^{+0.012}_{-0.012}$</td>
</tr>
<tr>
<td>$z_W$</td>
<td>$-0.213^{+0.049}_{-0.048}$</td>
<td>$-0.209^{+0.047}_{-0.047}$</td>
<td>$-0.208^{+0.046}_{-0.046}$</td>
<td>$-0.211^{+0.048}_{-0.048}$</td>
</tr>
<tr>
<td>$z_p$ [μas]</td>
<td>$-17^{+4}_{-4}$</td>
<td>$-17^{+4}_{-4}$</td>
<td>$-17^{+4}_{-4}$</td>
<td>$-17^{+4}_{-4}$</td>
</tr>
</tbody>
</table>

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