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

Radoslaw Wojtak\textsuperscript{1} * & Jens Hjorth\textsuperscript{1}
\textsuperscript{1}DARK, Niels Bohr Institute, University of Copenhagen, Jagtvej 128, 2200 Copenhagen, Denmark

9 August 2024

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 standardisation in the calibration sample (galaxies with observed Cepheids) require 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 that of the Milky Way ($R_B \approx 4.6 \pm 0.4$) 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 standardisation. 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 which coincides with the blue end of the observed SN colour distribution ($c \approx -0.13$), the Hubble constant from Planck observations of the CMB (assuming a flat ΛCDM 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 require further investigations.

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

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 CMB assuming a flat ΛCDM cosmological model (Planck Collaboration et al. 2020) has recently drawn much attention. With $H_0 = 73.04 \pm 1.04 \, \text{km s}^{-1} \, \text{Mpc}^{-1}$ derived from the most recent advances in Cepheid and SN observations of the SH0ES collaboration (for Supernova H0 for the Equation of State Riess et al. 2021b) and $H_0 = 67.36 \pm 0.54 \, \text{km s}^{-1} \, \text{Mpc}^{-1}$ inferred from the Planck data (Planck Collaboration et al. 2020), it is currently the strongest divergence in cosmological measurements with a confidence level reaching 5σ. 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 ΛCDM cosmological model.

Multiple extensions or modifications of the standard ΛCDM 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 hypothesised 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 oscillations (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; Fondi et al. 2022; Vagnozzi 2021). Furthermore, early dark energy also has been proven to affect the power spectrum derived from the CMB in a way

* E-mail: radek.wojtak@nbi.ku.dk

© 0000 RAS
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 which 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 that fact the Hubble constant measured locally is not a direct observable which can be used as a prior in cosmological analyses with type Ia SN data. Finding an expansion history which 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 analog obtained in the so-called inverse distance ladder method based on distance scales calibrated with the Planck data (Efstathiou 2021; Calamara & Marra 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 which 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 (SBF; 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 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 standardise 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 standardisation 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 standardisation 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 summarise 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) and Riess et al. (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, $m_{F160W}$, and pulsation periods, $P$, of Cepheids observed in the Milky Way ($L_{MW}$; Riess et al. 2021a), the Large Magellanic Cloud ($L_{LMC}$; Riess et al. 2019) and 20 galaxies ($L_{gal}$; Riess et al. 2016) including 19 type Ia SNe host galaxies and the megamaser galaxy NGC 4258. These measurements constrain distance moduli $\mu$ via the following equation:

$$\Delta \mu = \frac{1}{H_0} \left( \frac{D_L}{D_s} - 1 \right),$$
\[ m_{W}^{F160W} = \mu + M_{F160W} + b_{W}(\log_{10} P - 1) + z_{W}\Delta[O/H], \]  
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 \eqref{eq:m_W} incorporates corrections due to metallicity \( \Delta[O/H] \) which is directly measured for all Milky Way Cepheids and local environments of Cepheids in galaxies of the calibration sample. Following \cite{Riess2019} 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 \cite{Riess2016} we derive \( \Delta[O/H] \) assuming the solar metallicity given by the calibration of these metallicity estimates, i.e., 8.93 \cite{AndersGrevesse1989}. Wesenheit magnitudes of Cepheids in the 20 calibration galaxies are calculated from magnitudes \( m_{F160W} \) and colours \( m_{F555W} - m_{F814W} \) provided by \cite{Riess2016} using the relation

\[ m_{W}^{F160W} = m_{F160W} + R(m_{F555W} - m_{F814W}) \]  
with \( R = 0.386 \) corresponding to an extinction coefficient \( R_{0} = 4.3 \) in the reddening law of \cite{Fitzpatrick1999}, as assumed in \cite{Riess2019} and \cite{Riess2021a}.  

Distances to Cepheids in the Milky Way are constrained by \textit{Gaia} measurements of their parallaxes, \( \pi \), \cite{GaiaCollaboration2021}. Following \cite{Riess2021a} we include zero point \( z_{P} \) as a free parameter in order to obtain an unbiased relation between observed parallaxes, \( \pi \), and distance moduli using the following equation

\[ \pi + z_{P} = 10^{-0.2(m_{\mu} - 10)}. \]  

As described explicitly in Table \ref{table:likelihoods} the likelihood for the Milky Way 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 \( L_{\text{cal}} \) are given by the measurement uncertainties in \( m_{F160W} \) provided by \cite{Riess2016} and \cite{Riess2019}. Distance moduli of the LMC and the 20 calibration 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 \cite{Pietrzyński2019} and to NGC 4258 from megamasers \cite{Reid2019} are included in likelihoods \( L_{LMC} \) and \( L_{\text{cal}} \). Together with parallax distances to Cepheids in the Milky Way, these measurements serve as the anchors of the cosmological distance scale entering the Hubble constant determination.  

The last two data blocks in Table \ref{table:likelihoods} include observations of type Ia SNe. We use light curve parameters obtained by \cite{Scollnic2015} using the SALT2 fitting methodology \cite{Guy2005,ScollnicKessler2016}. Distance moduli are derived from peak apparent magnitudes, \( M_{B} \), by applying corrections related to the light curve shape quantified by \( z_{1} \) 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 \cite{Tripp1998},  

\[ m_{B} = \mu + M_{B} - \alpha z_{1} + \beta c, \]  
where B-band absolute magnitude, \( M_{B} \), and correction coefficients \( \alpha \) and \( \beta \) are measured directly from the SN data. Errors in distance moduli adopted in the SN likelihoods are given explicitly in Table \ref{table:SNlikelihoods}. They include all elements of covariance matrices obtained for individual SNe, intrinsic scatter \( \sigma_{\text{int}} \) and extra error due to unconstrained peculiar velocities (included only in \( L_{SN} \)) with \( \sigma_{s} = 200 \) km s\(^{-1}\) \cite{Carrick2015}.  

For the cosmological data block \( L_{SN} \) we selected SNe using the same criteria as those adopted by \cite{Riess2016}. The primary selection condition is given by redshift range \( 0.23 < z < 0.15 \) and cuts in light curve parameters: \( |c| < 0.3 \) and \( |z_{1}| < 3 \). In addition, SNe with low quality fits are rejected. SNe passing the fit quality check (223) are characterised by sufficient goodness of fit (fitprob > 0.001) and well-constrained light-curve parameters with errors < 1.5 for \( z_{1} \) < 2 days 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 characterised by an independent slope, \( \beta_{\text{cal}} \), of the colour correction, i.e.,  

\[ m_{B} = \mu + M_{B} - \alpha z_{1} + \beta_{\text{cal}}, \]  
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 \( (\beta_{\text{cal}} \equiv \beta) \) and intrinsic scatter \( (\sigma_{\text{int} \equiv \sigma_{\text{int,cal}}}) \) 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.,  

\[ L \propto L_{LMC} \times L_{SN} \times L_{\text{cal}} \times L_{4258} \times L_{\text{LMC \, dist}} \times L_{\text{SN \, cal}} \times L_{SN}. \]  

Our model is described by 11 primary parameters \( (H_{0}, M_{F160W}, b_{W}, z_{W}, z_{P}, M_{B}, \alpha, \beta, \sigma_{\text{int}}, \beta_{\text{cal}}, \sigma_{\text{int,cal}}) \) out of which two \( (\beta_{\text{cal}} \text{ and } \sigma_{\text{int,cal}} \text{ – the colour correction slope and intrinsic scatter in the SN calibration sample}) \) are optional. In addition, 21 latent variables are constrained as a 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 fit model. In particular, in our study we scrutinise the intrinsic consistency between the calibration \( (L_{SN,\text{cal}}) \) and cosmological \( (L_{SN}) \) SN data blocks.  

We use a Markov Chain Monte Carlo technique to integrate the posterior probability function and find best fit model parameters. The chains are computed with the \textit{emcee} code \cite{Foreman2013}. Unless explicitly stated, best-fit parameters are provided as the posterior mean values and errors given by 16th and 84th percentiles of the marginalised probability distributions. The 1σ and 2σ confidence contours shown in all figures contain 68 and 95 per cent, respectively, of the corresponding 2-dimensional marginalised 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 et al. 2020). The respective formula for distance modulus reads

$$
\mu(z) = \mu_{\text{Planck}}(z) + 5 \log_{10}(H_{0\text{Planck}}/H_0),
$$

where all quantities with 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 variation 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 which 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 (MF160W,bW,zW,p) and SN data (MB,α,β,σint,H0). 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 fit 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 \((M_V, b_V, z_{\text{min}})\) 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 \(z \lesssim 0.1\) SN samples (e.g. Jones et al. 2019). The best fit 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 summarised in Riess et al. (2021b). Our measurement is at a 4\(\sigma\) discrepancy with the Hubble constant determination from Planck observations assuming a flat ΛCDM cosmological model (Planck Collaboration et al. 2020). This tension increases to a 5\(\sigma\) 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 \((\mu_1,...,\mu_{19}\) in Table I) to check if the distribution of residuals in the SN block confirm 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 panel in Figure 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 underestimation of blue SNe \((c < 0)\) appears 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. Figure 1 also shows that 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 Figure 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
Table 2. Best fit 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 $\sigma_{\text{int,cal}} \equiv \sigma_{\text{int}}$), and model C (free $\beta_{\text{cal}}$ and $\sigma_{\text{int,cal}} \equiv \sigma_{\text{int}}$). Best fit results are provided in the form of the posterior mean values and errors containing 68 per cent of marginalised probabilities (with 68 percent upper limit for derived assuming reference colour models are constrained solely by the data from the SN calibration block, the goodness of fit is quantified in terms of the corresponding Planck value). The constraints on SN parameters ($M_B$, $\alpha$, $\beta$, $\sigma_{\text{int}}$, $\beta_{\text{cal}}$, $\sigma_{\text{int,cal}}$) and the Hubble constant in model A are shown in Figure 2. The red lines indicate combinations of parameters reducing model A to the baseline model, i.e., $- \beta = \beta_{\text{cal}}$ and $\sigma_{\text{int}} = \sigma_{\text{int,cal}}$. The right 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).

### SN parameters

|       | baseline | model A | model B | model C |
|-------|----------|---------|---------|---------|
| $\beta$ | $3.109^{+0.112}_{-0.112}$ | $3.060^{+0.114}_{-0.115}$ | $3.057^{+0.114}_{-0.113}$ | $3.062^{+0.114}_{-0.114}$ |
| $\sigma_{\text{int}}$ | $0.115^{+0.008}_{-0.008}$ | $0.115^{+0.008}_{-0.008}$ | $0.115^{+0.008}_{-0.008}$ | $0.113^{+0.008}_{-0.008}$ |
| $\beta_{\text{cal}}$ | $\equiv \beta$ | $4.285^{+0.504}_{-0.503}$ | $4.565^{+0.384}_{-0.379}$ | $4.110^{+0.542}_{-0.550}$ |
| $\sigma_{\text{int,cal}}$ | $\equiv \sigma_{\text{int}}$ | $< 0.097$ | $\equiv 0$ | $\equiv \sigma_{\text{int}}$ |
| $M_B$ | $-19.249^{+0.038}_{-0.037}$ | $-19.222^{+0.037}_{-0.037}$ | $-19.216^{+0.032}_{-0.032}$ | $-19.228^{+0.040}_{-0.040}$ |
| $\alpha$ | $0.129^{+0.008}_{-0.009}$ | $0.129^{+0.008}_{-0.008}$ | $0.130^{+0.008}_{-0.008}$ | $0.129^{+0.008}_{-0.008}$ |
| $H_0$ [km s$^{-1}$ Mpc$^{-1}$] | $73.142^{+1.315}_{-1.306}$ | $73.864^{+1.285}_{-1.283}$ | $74.101^{+1.124}_{-1.137}$ | $73.708^{+1.419}_{-1.413}$ |

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

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 independent parameter (model A), vanishing ($\sigma_{\text{int,cal}} \equiv 0$, model B), or equal to the analogous scatter in the cosmological sample ($\sigma_{\text{int,cal}} \equiv \sigma_{\text{int}}$, model C). We summarise the best fit 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 ($M_B$, $\alpha$, $\beta$, $\sigma_{\text{int}}$, $\beta_{\text{cal}}$, $\sigma_{\text{int,cal}}$) and the Hubble constant in model A are shown in Figure 2. The red lines indicate combinations of parameters reducing model A to the baseline model, i.e., $- \beta = \beta_{\text{cal}}$ and $\sigma_{\text{int}} = \sigma_{\text{int,cal}}$. The right 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, 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, $\beta_{\text{cal}}$, in the calibration sample is larger than in the cosmological sample $\beta$. We find $\beta_{\text{cal}} - \beta = 1.23 \pm 0.54$ for model A, $\beta_{\text{cal}} - \beta = 1.51 \pm 0.40$ for model B and $\beta_{\text{cal}} - \beta = 1.05 \pm 0.56$ for model C. Figure 3 shows the marginalised posterior distribution for $\beta_{\text{cal}} - \beta$ obtained for the three models. Numerical estimation of the probability that $\beta_{\text{cal}} \leq \beta$ yields $0.017 (2.4\sigma)$ for model A, $\sim 10^{-4} (3.8\sigma)$ for model B and 0.025 (2.2\sigma) for model C.

Table 2 provides goodness of fit quantified by the maximum likelihood and the Bayesian Information Criterion (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 ($\Delta \text{BIC} \ll -6$) over the baseline model. Model A yields a substantially better fit than the baseline model, but its free intrinsic scatter $\sigma_{\text{int,cal}}$ turns out to be a redundant parameter (no detection: the maximum likelihood found at $\sigma_{\text{int,cal}} = 0$ and fit yielding merely an upper bound, see Table 2 and Figure 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...
Figure 2. Constraints on parameters of the SN standardisation and the Hubble constant obtained for model A in which SN colour correction and intrinsic scatter are assumed to be independent in the calibration and cosmological samples. The blue dashed lines show combinations of parameters which reduce model A to the baseline model which assumes a universal colour correction and intrinsic scatter across all SNe, i.e., $\beta = \beta_{\text{cal}}$ and $\sigma_{\text{int}} = \sigma_{\text{int,cal}}$. The results demonstrate that SNe in the calibration sample require stronger colour correction ($\beta_{\text{cal}} > \beta$) than those in the cosmological sample. It is also apparent that the intrinsic scatter in the calibration sample with an independent SN colour correction is consistent with 0, which is not the case of the baseline model. The contours show 1σ and 2σ confidence regions containing 68 and 95 per cent of 2-dimensional marginalised probability distributions. The summary statistics above diagonal panels shows the median values and errors given by 16-th and 84-th quantiles of the marginalised probability distributions.

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 fully consistent with the data.

The strongest discrepancy between $\beta_{\text{cal}}$ and $\beta$ 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
standardisation in the calibration sample. This can be also realised by fixing a model with a universal colour correction \((\beta = \beta_{\text{cal}})\) and free \(\sigma_{\text{int-cal}}\). The fit yields \(\sigma_{\text{int-cal}} = 0.135 \pm 0.034\) demonstrating directly that the assumption of universal SN colour correction substantially degrades the precision of the SN standardisation in the calibration sample. It is also worth emphasizing that model B has a 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 Milky Way.

### 3.3 Hubble constant

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

\[
\begin{align*}
    m_B &= \mu + M_B + \alpha x_1 + \beta(c - c_{\text{ref}}) & \text{for SN block,} \\
    m_B &= \mu + M_B + \alpha x_1 + \beta_{\text{cal}}(c - c_{\text{ref}}) & \text{for SN cal block.}
\end{align*}
\]

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 minimisation 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 moduli \(\Delta \mu\):

\[
\Delta \mu = (\beta - \beta_{\text{cal}})c_{\text{ref}}. \tag{9}
\]

This means that every choice of \(c_{\text{ref}}\) has its unique best fit value of the Hubble constant inferred from the same observational data.

The values of the Hubble constant measured for models A, B, C and listed in Table 2 are obtained for a 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 eq. (9) that the best fit Hubble constant shifts by \(\Delta H_0\) with respect to its value obtained for \(c_{\text{ref}} = 0\) as given by the following equation:

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

Using the above approximation we can show that models A, B or C can recover the Planck value of the Hubble constant respectively for \(c_{\text{ref}} \approx -0.16, -0.13, -0.18\). 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 sample. Interestingly, the reference colour for the preferred model B is within the 1σ 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 eq. (8). Figure 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-2σ 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 fit \(H_0\) 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 fit \(H_0\) 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 minimised 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 be-
tween best fit $H_0$ 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 eq. (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 1σ when the adopted reference colour in eq. (5) 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. 4). 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 $c_{\text{ref}} \lesssim -0.1$ 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 colours 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); Brout & Scolnic (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 et al. (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 minimising the Hubble constant tension is a single-valued intrinsic colour $c_{\text{int}} \approx -0.13$ (or a very narrow distribution with the mean value equal to $c_{\text{int}}$) and a colour excess distribution given by $c - c_{\text{int}}$. 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 a SN may depend on other intrinsic properties (e.g. Foley et al. 2011).

Assuming that the colour excess with respect to reference colour $c_{\text{ref}} \lesssim -0.13$ 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 $R_B$. In this picture, galaxies in the calibration sample appear to be analogs of the Milky Way in terms of their dust properties. The extinction coefficient in the calibration sample determined in our analysis ($R_B \approx 4.6 \pm 0.4$ for model B in Table 2) is fully consistent with the mean extinction coefficient measured in the Milky Way, i.e., $R_B \approx 4.3$ (Fitzpatrick 1999; Cardelli et al. 1989). Furthermore, Milky-Way-like extinction in the calibration sample readily improves the precision of the SN standardisation 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 $R_B = 4.3$ (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 characterised by a sub-
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. Jones et al. 2019; Scolnic et al. 2018) 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 suggest 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 picture picture connecting dust properties and extinction inferred from observations is still missing. In particular, the extinction coefficient measured in the Milky Way falls into a rather wide range $3.5 \lesssim R_B \lesssim 6.5$ (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 standardisation given by eq. (5) proposed by Tripp (1998). Our main results are:

(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. Standardisation of SNe in the calibration sample requires a steeper slope found at 2.2$\sigma$ confidence level assuming universal intrinsic scatter, 2.4$\sigma$ confidence level when fitting intrinsic scatter independent in the calibration sample and at 3.8$\sigma$ confidence level assuming vanishing intrinsic scatter in the calibration sample. The latter model maximising the difference between the 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. (2021) 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 Milky Way. This suggests that galaxies in the calibration sample – unlike SN host galaxies in the cosmological sample – are analogs to the Milky Way 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 Milky-Way-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 insufficiently accurate standardisation 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 if 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); Popovic et al. (2021). Another avenue worth pursuing are 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 (Dhawan et al. 2018; Burns et al. 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.

ACKNOWLEDGMENTS

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 which helped improve this work.

DATA AVAILABILITY

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

REFERENCES

Anders E., Grevesse N., 1989, Geochim. Cosmochim. Acta, 53, 197
Arendse N. et al., 2020, A&A, 639, A57
Birrer S. et al., 2020, A&A, 643, A165
Brout D., Scolnic D., 2021, ApJ, 909, 26
Tension in the SN sector of $H_0$ measurement

Brout D. et al., 2022, arXiv e-prints, arXiv:2202.04077
Burns C. R. et al., 2018, ApJ, 869, 56
Camarena D., Marra V., 2021, MNRAS, 504, 5164
Cardelli J. A., Clayton G. C., Mathis J. S., 1989, ApJ, 345, 245
Carrick J., Turnbull S. J., Lavaux G., Hudson M. J., 2015, MNRAS, 450, 317
Dhawan S., Jha S. W., Leibundgut B., 2018, A&A, 609, A72
Di Valentino E. et al., 2021, Classical and Quantum Gravity, 38, 153001
Efstathiou G., 2014, MNRAS, 440, 1138
Efstathiou G., 2021, MNRAS, 505, 3866
Fitzpatrick E. L., 1999, PASP, 111, 63
Foley R. J., Sanders N. E., Kirshner R. P., 2011, ApJ, 742, 89
Fondi E., Melchiorri A., Pagano L., 2022, arXiv e-prints, arXiv:2203.12930
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Freedman W. L. et al., 2019, ApJ, 882, 34
Gaia Collaboration et al., 2021, A&A, 649, A1
Garnavich P. et al., 2022, arXiv e-prints, arXiv:2204.12060
Guy J., Astier P., Nobili S., Regnault N., Pain R., 2005, A&A, 443, 781
Hill J. C., McDonough E., Toomey M. W., Alexander S., 2020, Phys. Rev. D, 102, 043507
Jones D. O. et al., 2022, arXiv e-prints, arXiv:2201.07801
Jones D. O. et al., 2019, ApJ, 881, 19
Khetan N. et al., 2021, A&A, 647, A72
Knox L., Millea M., 2020, Phys. Rev. D, 101, 043533
Mandel K. S., Scolnic D. M., Kirshner R. P., 2011, ApJ, 731, 120
Mandel K. S., Scolnic D. M., Shariff H., Foley R. J., Kirshner R. P., 2017, ApJ, 842, 93
Mortsell E., Goobar A., Johansson J., Dhawan S., 2021, arXiv e-prints, arXiv:2106.09400
Phillips M. M., 1993, ApJ, 413, L105
Pietrzyński G. et al., 2019, Nature, 567, 200
Planck Collaboration et al., 2020, A&A, 641, A6
Popovic B., Brout D., Kessler R., Scolnic D., 2021, arXiv e-prints, arXiv:2112.04456
Poulin V., Smith T. L., Karwal T., Kamionkowski M., 2019, Phys. Rev. Lett., 122, 221301
Reid M. J., Pesce D. W., Riess A. G., 2019, ApJ, 886, L27
Riess A. G., Casertano S., Yuan W., Bowers J. B., Macri L., Zinn J. C., Scolnic D., 2021a, ApJ, 908, L6
Riess A. G., Casertano S., Yuan W., Macri L. M., Scolnic D., 2019, ApJ, 876, 85
Riess A. G. et al., 2016, ApJ, 826, 56
Riess A. G. et al., 2021b, arXiv e-prints, arXiv:2112.04510
Scolnic D. et al., 2015, ApJ, 815, 117
Scolnic D., Kessler R., 2016, ApJ, 822, L35
Scolnic D. M. et al., 2018, ApJ, 859, 101
Thorp S., Mandel K. S., Jones D. O., Ward S. M., Narayan G., 2021, MNRAS, 508, 4310
Tripp R., 1998, A&A, 331, 815
Vagnozzi S., 2021, Phys. Rev. D, 104, 063524
Visser M., 2004, Classical and Quantum Gravity, 21, 2603
Wong K. C. et al., 2020, MNRAS, 498, 1420
| Cepheid parameters | baseline          | model A          | model B          | model C          |
|--------------------|-------------------|-------------------|-------------------|-------------------|
| \( M_{F160W} \)    | -5.929 ± 0.014    | -5.929 ± 0.014    | -5.929 ± 0.014    | -5.929 ± 0.014    |
| \( b_W \)          | -3.292 ± 0.012    | -3.292 ± 0.012    | -3.292 ± 0.012    | -3.292 ± 0.012    |
| \( z_W \)          | -0.213 ± 0.048    | -0.209 ± 0.047    | -0.208 ± 0.046    | -0.211 ± 0.048    |
| \( zp[^{\mu}\text{as}] \) | -17^{+4}_{-4}    | -17^{+4}_{-4}    | -17^{+4}_{-4}    | -17^{+4}_{-4}    |

Table A1. Best fit 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 \( \sigma_{\text{int-cal}} \)), model B (free \( \beta_{\text{cal}} \) and \( \sigma_{\text{int-cal}} \equiv 0 \)) and model C (free \( \beta_{\text{cal}} \) and \( \sigma_{\text{int-cal}} \equiv \sigma_{\text{int}} \)). Best fit results are provided in the form of the posterior mean values and errors containing 68 per cent of marginalised probabilities.

APPENDIX A: BEST FIT PARAMETERS OF THE CEPHEID DATA