

The Gaia and HST Cepheid Scales and the Tension in Hubble's Constant

Christian Kirkham

Level 4 Project, MPhys Physics and Astronomy

Supervisors: Professor Tom Shanks and Professor Nigel Metcalfe

Department of Physics, Durham University

Submitted: April 20, 2022

Abstract

This project aims to investigate the effect on Hubble's Constant of the systematic errors in Gaia EDR3 parallaxes, for which the Gaia team have given a suggested correction. Investigating EDR3 quasar parallaxes demonstrates that the ecliptic latitude dependence of the parallax bias is removed by the Gaia correction but a small residual median offset of $\pi_{\text{QSO,median}} = -1.1 \pm 0.6 \mu\text{as}$ is found. Using the difference between the parallax of open cluster stars and the cluster barycentre, the dependence of the bias on G magnitude was found to be corrected well for stars with magnitude $G > 9$ mag but a residual median parallax of $\pi_{\text{clus,median}} = -3 \pm 4 \mu\text{as}$ was found for bright stars with $G < 9$ mag. These medians were combined into a parallax zeropoint of $zp = +4 \pm 4 \mu\text{as}$, which is of the order of values found in the literature. Applying this zeropoint to the SH0ES method of calibrating and calculating Hubble's constant yielded a change in H_0 of the order of $\sim 0.01\%$ when combined with multiple distance anchors and so is unable to account for the Hubble tension. A best-fit value of $H_0 = 75.2 \pm 1.2 \text{ km sec}^{-1} \text{ Mpc}^{-1}$ was determined using the combination of anchors of EDR3 Milky Way Cepheid parallaxes, LMC Detached Eclipsing Binaries and NGC 4258 Water Masers. This value is consistent with the SH0ES value but more Cepheid photometry in Cepheid-SNIA hosts is required for a more accurate determination. Using main-sequence fitting to open cluster colour-magnitude diagrams to investigate the Wesenheit magnitude reddening law showed that the goodness-of-fit of the MS fit increased with distance. Using a modified reddening law which included a distance dependence showed that nearby clusters followed the standard reddening law of $R \sim 3.1$ while distant clusters diverged from the expected law, suggesting that a modified law must be incorporated into future calculations of Hubble's constant to properly calibrate lower rungs of the distance ladder.



CONTENTS

1. Introduction	3
1.1. Motivation	3
1.2. Gaia Early Data Release 3	4
1.3. SH0ES	5
1.4. Galactic Reddening	7
2. Data Sources and Methods	8
2.1. Gaia EDR3	8
2.1.1. Quasars	8
2.1.2. Open Clusters	9
2.2. Calculating Hubble's Constant	9
2.2.1. Cepheid P-L Relation Calibration	9
2.2.2. SNIA Calibration	10
2.2.3. Simultaneous Fit	10
2.2.4. Redshift-Distance Calibration	12
2.3. Main Sequence Fitting	12
3. Results and Discussion	14
3.1. Gaia EDR3 Parallax Bias	14
3.1.1. Ecliptic Latitude Dependence	14
3.1.2. G Magnitude Dependence	15
3.2. Comparing EDR3 with HST Parallaxes	16
3.3. The Hubble Constant	17
3.3.1. SH0ES Method	18
3.3.2. Cepheid-Redshift Diagram	23
3.4. Galactic Reddening	24
4. Conclusions	26
4.1. Gaia EDR3	26
4.2. Galactic Reddening	27
4.3. Hubble Tension	28
Acknowledgments	29
References	29
A. Anchor Combinations	31
B. Cosmic Distance Ladder	33
C. Main Sequence Fits	34
D. MCMC Method	35

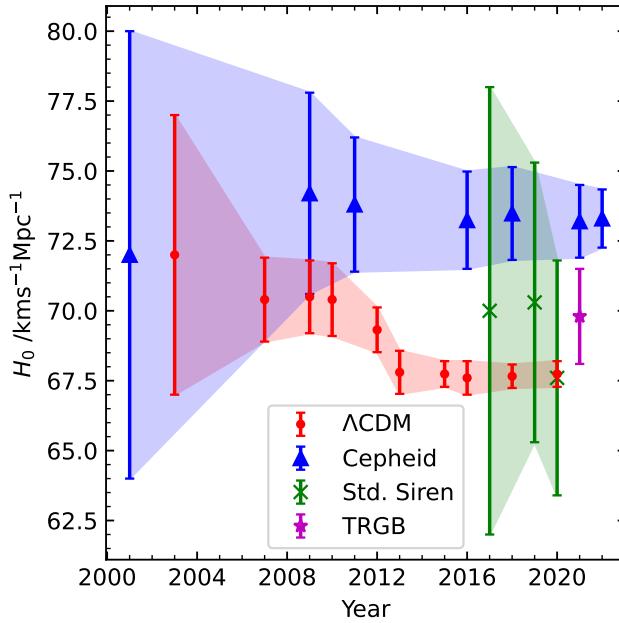


FIG. 1: Historical values of the Hubble constant over time, denoted by the method by which it was determined. It is clear that the Planck Λ CDM values and the SH0ES Cepheid values are in tension. Other values from sources such as gravitational wave Standard Sirens and the Tip of the Red Giant Branch are generally between the two main values.

1. INTRODUCTION

1.1. Motivation

Determining the Hubble constant, H_0 , precisely is one of the most important goals of modern cosmology as quantities such as the age and size of the universe depend on its value (Perivolaropoulos 2022). The Hubble constant is the constant of proportionality in the Hubble-Lemaitre law,

$$v = H_0 d, \quad (1)$$

which relates the measured recession velocity, v , of an object to its distance from the observer, d , and is a measure of the current expansion rate of the universe. This law was first derived from Einstein's General Theory of Relativity by Lemaître (1927) and verified observationally by Hubble (1929), with values of H_0 calculated to be $625 \text{ km sec}^{-1} \text{ Mpc}^{-1}$ and $500 \text{ km sec}^{-1} \text{ Mpc}^{-1}$ respectively. From hereafter the units of H_0 will be omitted for brevity.

Almost a century later, astronomers have determined values of H_0 to within a couple of percent, but the measured value differs depending on the method used, as can be seen in figure 1. The two leading values are the Planck Collaboration's Λ CDM (cosmological model with a cosmological constant, Λ , and cold dark matter, CDM) value and the SH0ES team's Cepheid value, both of which are in increasing tension with each other as measurements improve over time.

Determined by measuring the power spectrum of temperature fluctuations in the Cosmic Microwave Background, the Planck team found the value of H_0 to be 67.74 ± 0.46 (Planck

(Collaboration et al. 2020) and CMB measurements to be in very good agreement with the predictions of Λ CDM. The value of the Planck H_0 is dependent on the choice of cosmological model, so relieving the tension from this side requires a change in the physics of cosmology. Some attempts have been made to look for alternative models in the Planck data, such as Early Dark Energy at the time of recombination (Smith et al. 2022), but as yet none are as compelling as Λ CDM.

The Supernova Cosmology for H_0 and a Dark Energy Equation of State (SH0ES) team's value of H_0 depends on Hubble Space Telescope (HST) observations of Cepheid variable stars and Pantheon+ observations of Type IA supernovae (SNIA), determining a value of 73.04 ± 1.04 (Riess et al. 2022, hereafter R22). This latest value places the tension between the SH0ES and Planck values of H_0 at the key value of 5σ , signifying a new discovery. Possible ways to relieve the Hubble tension using the SH0ES method will be discussed in section 1.3.

It is this which motivates this report, in particular the potential to relieve the Hubble tension by analysing the SH0ES data and methods. This work will first consider systematic errors in parallaxes from the Gaia Early Data Release 3 and investigate a method to correct for these biases. This bias correction will then be used to compare Gaia parallaxes to HST trigonometric parallaxes, as used in previous SH0ES investigations. Next, the corrected Gaia parallaxes will be used in conjunction with SH0ES Cepheid photometry and Pantheon SNIA observations to determine the effects of the biases on the value of the Hubble constant. Finally, Gaia data will be used to investigate the need for a change to the dust reddening law which could be used to relieve the Hubble tension.

1.2. Gaia Early Data Release 3

Located at Lagrange Point 2, the Gaia satellite uses two telescopes to measure parallaxes and proper motions for millions of stars across the whole sky (Lindegren et al. 2021a). As a result of the huge number of trigonometric distances to many galactic and LMC Cepheid variable stars, this data is a very good tool for calibrating the Cepheid Period-Luminosity relation, serving as a robust first rung in the cosmological distance ladder (see section 1.3).

Despite this, by Gaia Early Data Release 3, a parallax bias has been found in the data due to a thermal difference between the two telescopes (Lindegren et al. 2021b, hereafter L21b). This systematic bias is known to have a dependence on the ecliptic latitude, G magnitude and colour (measured using either the effective wavelength or the pseudocolour, depending on the astrometric solution) of each individual star. There have been a few independent attempts to correct for this parallax bias (Flynn et al. 2022, Maíz Apellániz 2022, Wang et al. 2022), all of which attempt to improve on the correction provided by the Gaia team in L21b. This correction is given using the tabulated form,

$$Z_{\text{EDR3}}(G, \nu_{\text{eff}}, \beta) = \sum_j \sum_k q_{jk}(G) c_j(\nu_{\text{eff}}) b_k(\beta), \quad (2)$$

which corrects the raw Gaia parallaxes, π_{raw} , as

$$\pi_{\text{corr}} = \pi_{\text{raw}} - Z_{\text{EDR3}}. \quad (3)$$

Here, G is the G magnitude, ν_{eff} is the effective wavelength, and β is the ecliptic latitude of the star. The coefficient $c_j(\nu_{\text{eff}})$ describes the bias due to the colour of the stars with indices, $j = 0 \dots 4$, which correspond to polynomials of the form

$$j = \begin{cases} 0, & \text{Zeroth order in } \nu_{\text{eff}} \\ 1, & \text{Linear first order in } \nu_{\text{eff}} \\ 2, & \text{Cubic in } \nu_{\text{eff}} \\ 3, & \text{First order in } \nu_{\text{eff}} \text{ for } \nu_{\text{eff}} \leq 1.24 \text{ (red stars)} \\ 4, & \text{First order in } \nu_{\text{eff}} \text{ for } \nu_{\text{eff}} > 1.72 \text{ (blue stars)} \end{cases} \quad (4)$$

All $c_j(\nu_{\text{eff}})$ values are clamped outside the range $1.24 < \nu_{\text{eff}} \leq 1.72$ unless specified.

The coefficient $b_k(\beta)$ describes the bias due to the ecliptic latitude, β , of the stars with indices, $k = 0 \dots 2$, which correspond to polynomials in $\sin \beta$ of the form

$$k = \begin{cases} 0, & \text{Zeroth order in } \sin \beta \\ 1, & \text{First order in } \sin \beta \\ 2, & \text{Quadratic in } \sin \beta \end{cases} \quad (5)$$

To calculate the offset, the pre-calculated values of the coefficients given in the tables in L21b can be interpolated. Alternatively the Gaia team have provided a Python package, `zeropoint`, which can be used to calculate Z_{EDR3} ¹.

1.3. SH0ES

The SH0ES method for determining the Hubble constant uses HST photometry and depends on three rungs of calibration of the cosmological distance ladder: geometric distances to calibrate the Cepheid period-luminosity (PL) relation, Cepheid photometric distances to calibrate the SNIA light curves and SNIA photometric distances to calibrate the redshift-distance relation (R22). The Hubble constant is calculated from this final step, alongside the deceleration parameter, q_0 , as a fit parameter to the Hubble diagram.

The geometric calibration of the Cepheids is performed using four distance measurements, known as distance anchors. These four anchors are as follows:

- **Milky Way Cepheids** whose distances can be measured directly from EDR3 parallaxes, as discussed earlier. There are currently 74 Cepheids with both SH0ES HST photometry and EDR3 parallaxes.
- **NGC 4258 Water Maser**, a bright source which gives a geometric distance of $D_{\text{N4258}} = 7.576 \pm 0.082 \text{ (stat)} \pm 0.076 \text{ (sys)} \text{ Mpc}$ (Reid et al. 2019).
- **20 LMC Detached Eclipsing Binaries (DEB)** which give a geometric distance of $D_{\text{LMC}} = 49.59 \pm 0.09 \text{ (stat)} \pm 0.54 \text{ (sys)} \text{ kpc}$ (Pietrzyński et al. 2019).
- **1 M31 Detached Eclipsing Binary** which gives a geometric distance of $D_{\text{M31}} = 744 \pm 33 \text{ kpc}$ (Vilardell et al. 2010).

¹ <https://pypi.org/project/gaiadr3-zeropoint/>

These four anchors, combined with HST photometry of Cepheids, are used to calibrate the Cepheid PL relation. First discovered by [Leavitt & Pickering \(1912\)](#), this relation links the pulsation periods of Cepheid-type variable stars with their intrinsic luminosity, providing an invaluable method of measuring distances known as a standard candle. The form of the PL relation used by the SH0ES team is given as,

$$M_H^W = M_{H,1}^W + b_W(\log P - 1) + Z_W \Delta[\text{O}/\text{H}], \quad (6)$$

where M_H^W is the reddening-free absolute Wesenheit magnitude of the Cepheid, P is the fundamental period of the Cepheid pulsation, $\Delta[\text{O}/\text{H}]$ is the metallicity of the Cepheid, and $M_{H,1}^W$, b_W and Z_W are the parameters of the P-L relation ([Riess et al. 2016](#), hereafter R16). The Wesenheit magnitude, or Wesenheit function, in the V band as defined by [Madore & Freedman \(1991\)](#) is

$$M_V^W = M_V - R_V(B - V), \quad (7)$$

where M_V is the V-band magnitude, $(B - V)$ is the apparent colour, and $R_V = A_V/E(B - V)$ is the ratio of the total absorption to the selective absorption and is a constant $R_V \sim 3.1$. This is a useful quantity for comparing standard candles with the reddening effect of dust removed but should not be considered to be a conventional magnitude.

The form of the H-band Wesenheit magnitude used by the SH0ES team and hence this investigation is given in ([Riess et al. 2021](#), hereafter R21) as

$$m_H^W = m_{\text{F160W}} - 0.386(m_{\text{F555W}} - m_{\text{F814W}}), \quad (8)$$

where the HST passbands can be approximated to the standard photometric colours as F160W = H, F555W = V and F814W = I.

Alongside the necessity to properly determine the Gaia EDR3 parallax bias to accurately calibrate the PL relation, [Efstathiou \(2020\)](#) has identified a number of other possible issues with the SH0ES anchors. One such issue is that of crowding in NGC 4258 due to the majority of observed stars lying close to the centre of the galaxy. As a result, there is a high star density which could cause issues when performing photometry on N4258 Cepheids. This issue has been alleviated somewhat by [R22](#) who increased their Cepheid set by observing the less-crowded outer regions of the galaxy.

To use these Cepheid calibrations to determine the value of the Hubble constant, it is necessary to use Type IA Supernovae to measure distances out to where the Hubble flow dominates over any peculiar velocities, allowing the use of redshifts to measure the expansion of the universe. As such, the SH0ES team's observations have focussed on two main areas: firstly, HST photometry of Cepheid stars in SNIA hosts ([R16](#)) and second, Pantheon+ photometry and redshifts of Type IA supernova out to high distances ([Scolnic et al. 2022](#)).

The focus of this work will be on systematic errors in the SH0ES data and its methods in order to investigate ways to reduce the Hubble tension. Other explanations, beyond the scope of this work, are that we are living in a local underdensity known as a “Local Hole” ([Wong et al. 2022](#)) which could reduce the tension with the Planck H_0 from 5σ to $\sim 3\sigma$.

1.4. Galactic Reddening

As described in equation 7, there is a global reddening law assumed in all SH0ES calculations, whereby the value of R is assumed to be constant for all stars and all galaxies. This assumption was challenged by [Mortsell et al. \(2021a\)](#) who allowed R_W to vary between galaxies in the Wesenheit function

$$m_H^W = m_H - R_W(V - I), \quad (9)$$

which is a generalised form of the [R21](#) equation 8. Assuming a fixed value of R_W yielded a value of $H_0 = 70.4 \pm 1.5$, while allowing it to vary between hosts gave a lower value of $H_0 = 57.3 \pm 2.7$, with the values being in 1.9σ and 3.7σ tension with the Planck value respectively. This approach was criticised in [R22](#), as a varying value of R_W would imply a change to the black-body physics of Cepheid variable stars. They argue that the ratio of the luminosity, m_H , to the intrinsic colour, $(V - I)$, is a constant quantity determined by the Stefan-Boltzmann law, and so cannot be varied between galaxies.

A second method for varying the reddening law outlined in [Mortsell et al. \(2021a\)](#) and refined in [Mortsell et al. \(2021b\)](#) is to vary the effect of dust on observations, using a variant of the Wesenheit function given by

$$m_H^W = m_H - R_E E(V - I), \quad (10)$$

where $E(V - I)$ is the colour excess or extinction. This is an approach also adopted by [Follin & Knox \(2018\)](#) and involves varying R_E between galaxies as the behaviour of dust may vary across the sky and universe. By keeping a global R_E , [Mortsell et al. \(2021a\)](#) find $H_0 = 69.6 \pm 1.5$, in 1.4σ tension with Planck, while by allowing R_E to vary between galaxies they find $H_0 = 66.9 \pm 1.3$, which is in agreement with the Planck value. [Mortsell et al. \(2021b\)](#) improves on this approach by adopting a colour cut for Cepheids with $E(V - I) < 0.15$, obtaining a value of $H_0 = 68.6 \pm 2.6$, which agrees with the Planck value.

It is possible to investigate reddening law variations within the Milky Way using independent means. Several studies have mapped the variation of R within the galaxy ([Fitzpatrick & Massa 2007](#), [Schlafly et al. 2016](#)) finding that, depending on the line of sight, the value of R_V can vary in the range $R_V \sim 3 - 4$. Given that the SH0ES methodology assumes a constant $R_V = 3.1$, ignoring this variation could cause issues when dereddening Milky Way Cepheid magnitudes.

Effects of this changing reddening law could be evident in the data presented in [Hoyle et al. \(2003\)](#). Here, zero-age main sequence (ZAMS) fitting was used to infer the extinction of and distance to 11 young open clusters, each of which contained Cepheid variable stars. Of particular interest are the ZAMS fits to the $(U - B)$ against $(B - V)$ plots (UBV plot), as these are independent of the distance and depend only on the extinction, $E(B - V)$. It can be seen by visual inspection that the quality of the ZAMS fit worsens with increasing distance, possibly suggesting that their assumption of $R = 3.3$ is incorrect for more distant clusters. If confirmed, this could have implications for the calibration of Hubble's constant using photometry of Milky Way Cepheids.

The second aim of this investigation is to test the suggestion that R has a distance dependence by using main-sequence fitting and Gaia EDR3 to find how the goodness-of-fit varies with distance when assuming a constant R . Next, the reddening law will be modified to include a distance term and a fitted free “X-factor” parameter to inspect how X varies with distance.

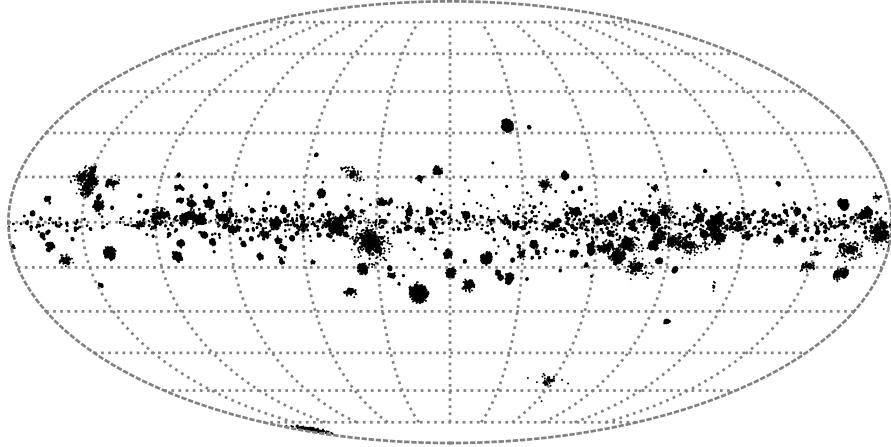


FIG. 2: Positions of the 237,535 open cluster stars of the [CG18](#) catalogue plotted in galactic coordinates.

2. DATA SOURCES AND METHODS

2.1. Gaia EDR3

2.1.1. Quasars

In order to investigate the ecliptic latitude dependence of the Gaia EDR3 parallax bias, it is necessary to find objects whose parallaxes are known independently. As they are some of the most distant known objects in the universe, quasars can be assumed to have a parallax of zero making them a useful tool for investigating the bias. This work will be making use of the Gaia catalogue of quasar candidates, `gaiaedr3.agn_cross_id`², which crossmatches EDR3 data with 17 external AGN catalogues as described in [L21b](#).

Before the data is used, a number of filters must be applied to remove objects with poor quality astrometry. As per [Groenewegen \(2021\)](#), objects with astrometric goodness-of-fit, $-4 < \text{astrometric_gof_al} < 5$, and renormalised unit weight error, $\text{ruwe} < 1.4$, are selected from the dataset and all other objects removed. These filters are applied to all Gaia EDR3 data used in this work.

Once these filters have been applied to the quasar data, the parallax is plotted against ecliptic latitude and binned median parallaxes are calculated in 0.5 degree bins. The same plot can then be produced for the quasar parallaxes after the [L21b](#) correction has been applied to judge the effectiveness of the Gaia correction. Taking a median of the full set of corrected parallaxes gives a residual median parallax, $\pi_{\text{QSO},\text{median}}$, which can be treated as a global offset that must be subtracted from all EDR3 parallaxes.

² <https://doi.org/10.17876/gaia/edr.3/10>

2.1.2. Open Clusters

To investigate the G magnitude dependence of the parallax bias, it is necessary to use a different dataset since quasars are generally dim with $G \gtrsim 16$ mag, and cannot be used to determine the bias for the brighter magnitudes of Milky Way Cepheids. Instead, it is possible to use open clusters whose stars cover a wide range of magnitudes to determine parallax biases by subtracting the parallax of the cluster member star, π_* , from the parallax of the cluster barycentre, π_c . The median value of $\pi_* - \pi_c$ should be zero, with any biases being found by taking medians of magnitude bins.

In the investigation, open cluster stars were taken from the [Cantat-Gaudin et al. \(2018\)](#) (hereafter CG18) catalogue. This catalogue of 237,535 stars, seen mapped in figure 2, uses Gaia DR2 data so the data must be updated with EDR3 parallaxes, making it necessary to recalculate the cluster parameters. To calculate the parallax of the cluster barycentre, the median parallax, $\tilde{\pi}_*$, of all stars with a cluster membership probability of 0.5 or greater was calculated. From this, the median absolute deviation (MAD),

$$\text{MAD} = \text{median}(|\pi_{*,i} - \tilde{\pi}_*|), \quad (11)$$

was found. Stars with an MAD of 3 or greater were removed from the data, and the median cluster parallax recalculated to give the value of π_c as per the method of [CG18](#).

Once the cluster parallaxes have been calculated, a plot of the parallax difference against G magnitude can be made with medians binned in 1 mag bins. The same plot can then be produced for the relative cluster parallaxes after the [L21b](#) correction has been applied to judge the effectiveness of the Gaia correction. If there are any residual biases for some magnitudes, a median, $\pi_{\text{clus,median}}$, can be found which can be combined with the quasar median to give an overall zeropoint offset,

$$zp = -(\pi_{\text{clus,median}} + \pi_{\text{QSO,median}}). \quad (12)$$

This zeropoint offset can then be used to further correct EDR3 parallaxes in the appropriate G magnitude range, modifying equation 3 to be

$$\pi_{\text{corr}} = \pi_{\text{raw}} - Z_{\text{EDR3}} + zp. \quad (13)$$

Note the sign convention used here to align with other works ([Freedman 2021](#), [Riess et al. 2021](#)) for comparison.

2.2. Calculating Hubble's Constant

To judge how the choice of geometric distance anchors affects the value of H_0 , the process will be repeated for various combinations of anchors, including raw EDR3 parallaxes, EDR3 parallaxes with [L21b](#) correction, EDR3 parallaxes with [L21b](#) and zp correction, NGC 4258, M31, and the LMC.

2.2.1. Cepheid P-L Relation Calibration

The aim of this rung is to fit the parameters of the PL relation given in equation 6, namely M_H^W , b_W and Z_W , to HST photometry of Cepheid variable stars. The Cepheid data used in

this work comes from a number of sources. H-band photometry for 372 M31 and 139 N4258 Cepheids are provided in R16. Photometry for 74 Milky Way Cepheids comes from R21 and LMC photometry of 70 Cepheids is given in Riess et al. (2019). R22 also makes use of a large ground-based sample of LMC Cepheids for anchoring but due to differences between the photometric systems used, this data will not be considered here.

Note that some care must be taken to calculate the metallicity correctly as

$$\Delta[\text{O}/\text{H}] = [\text{O}/\text{H}]_{\text{cep}} - [\text{O}/\text{H}]_{\odot}, \quad (14)$$

where $[\text{O}/\text{H}]_{\odot} = 8.69$ is the solar metallicity (Asplund et al. 2009). In R21 the relation $[\text{Fe}/\text{H}] = [\text{O}/\text{H}]$ is assumed.

To incorporate the geometric distance measurements given in section 1.3, the geometric distance modulus, μ_{geom} , is given from the distance, D (in parsecs), or the parallax, π (in millarcseconds), with

$$\mu_{\text{geom}} = 5 \log_{10} D - 5 = 10 - 5 \log_{10} \pi. \quad (15)$$

Once the PL parameter fit has been made, its accuracy can be evaluated by reversing the PL relation and comparing the photometric Cepheid distance moduli with the geometric distance moduli. This photometric distance modulus can be calculated as

$$\mu_{\text{cep}} = m_H^W - (M_{H,1}^W + b_W(\log_{10} P - 1) + Z_W \Delta[\text{O}/\text{H}]), \quad (16)$$

which when plotted against μ_{geom} should follow a 1:1 relation if the fit has been performed correctly.

2.2.2. SNIA Calibration

To calibrate the supernovae measurements, it is necessary to use the PL relation to calculate the distances to each SNIA host. The distance moduli are determined by fitting a single Cepheid distance, μ_{cep} , to all Cepheids in the respective SNIA host. The catalogue of Cepheid H-band photometry in R16 includes data for 19 SNIA hosts.

The absolute B-band peak magnitude of SNIA is fitted to HST B-band SNIA photometry from R16. Since all SNIA have the same peak magnitude, it is possible to fit a single magnitude, M_B , to all SNIA hosts simultaneously. Supernovae can then be used as measures of distance, finding the distance modulus as

$$\mu_{\text{SNIA}} = m_B - M_B. \quad (17)$$

2.2.3. Simultaneous Fit

Equations 6 and 17 are fitted simultaneously to all anchor Cepheids, SNIA host Cepheids and SNIA using the method of total least squares. This is done analytically using the following

system of equations (Riess et al. 2009). The vector of measurements is given by

$$\mathbf{y} = \begin{pmatrix} m_{i,j}^W - \mu_i \\ \vdots \\ m_{k,l}^W \\ \vdots \\ m_k^B \\ \vdots \end{pmatrix}, \quad (18)$$

where $m_{i,j}^W$ is the apparent Wesenheit magnitude of the j th Cepheid in the i th anchor, μ_i is the geometric distance to the i th anchor, $m_{k,l}^W$ is the apparent Wesenheit magnitude of the l th Cepheid in the k th SNIA host, and m_k^B is the apparent B magnitude of the supernova in the k th SNIA host. Individual EDR3 parallaxes are considered to be separate anchors. The vector of fit parameters is given by

$$\mathbf{q} = \begin{pmatrix} M_{H,1}^W \\ b_W \\ Z_W \\ M_B \\ \mu_k \\ \vdots \end{pmatrix}, \quad (19)$$

where $M_{H,1}^W$, b_W and Z_W are the Cepheid PL relation parameters, M_B is the absolute peak SNIA magnitude and μ_k is the distance modulus of the k th SNIA host. The equation matrix is given by

$$\mathbf{L} = \begin{pmatrix} 1 & \log_{10} P_{i,j} & \Delta[\text{O/H}]_{i,j} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \log_{10} P_{1,l} & \Delta[\text{O/H}]_{1,l} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \log_{10} P_{19,l} & \Delta[\text{O/H}]_{19,l} & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 & 0 & \cdots & 1 \end{pmatrix}, \quad (20)$$

where $P_{i,j}$ and $\Delta[\text{O/H}]_{i,j}$ are the period and metallicity of the j th Cepheid in the i th anchor and $P_{k,l}$ and $\Delta[\text{O/H}]_{k,l}$ are the period and metallicity l th Cepheid in the k th SNIA host. The matrix of measurement errors is given by

$$\mathbf{C} = \text{diag}(\sigma_{M,i,j}^2, \dots, \sigma_{m,k,l}^2, \dots, \sigma_{m,k}^2, \dots), \quad (21)$$

where $\sigma_{M,i,j}$ is the standard error of the absolute magnitude of the j th Cepheid in the i th anchor, $\sigma_{m,k,l}$ is the standard error of the apparent magnitude of the l th Cepheid in the k th SNIA host, and $\sigma_{m,k}$ is the standard error of the apparent magnitude of the supernova in the k th SNIA host.

The set of equations is solved analytically using

$$\mathbf{q}_{\text{best}} = (\mathbf{L}^T \mathbf{C} \mathbf{L})^{-1} \mathbf{L}^T \mathbf{C}^{-1} \mathbf{y}. \quad (22)$$

The standard error in \mathbf{q}_{best} is taken as the square root of the diagonal of the covariance matrix, $(\mathbf{L}^T \mathbf{C} \mathbf{L})^{-1}$, and the reduced chi-squared value is given by

$$\chi_{\nu}^2 = \frac{1}{\nu} (\mathbf{y} - \mathbf{L}\mathbf{q})^T \mathbf{C}^{-1} (\mathbf{y} - \mathbf{L}\mathbf{q}), \quad (23)$$

where $\nu = N - n$ is the degrees of freedom with N , the number of data points, and n , the number of free parameters.

2.2.4. Redshift-Distance Calibration

This investigation uses SNIA redshifts and apparent B-band peak magnitudes from [Scolnic et al. \(2018\)](#). This catalogue consists of 1,048 supernovae with redshifts in the range $0.01 < z < 2.3$, known as the Pantheon sample. Here, redshifts in the CMB frame will be used to remove the effects of peculiar velocities on the data.

To calculate H_0 from this data requires using the redshift-distance relation in the form

$$\log_{10} \left(cz \left[1 + \frac{1}{2}(1 - q_0)z - \frac{1}{6}(1 - q_0 - 3q_0^2 + j_0)z^2 \right] \right) = 0.2m_B + a_B. \quad (24)$$

Which can be used to plot the LHS of the equation against the apparent B-band magnitude, $0.2m_B$, and fitting for the deceleration parameter, q_0 , and the y-intercept of this plot, a_B . The method of chi-squared minimisation is then used to fit for a_B and q_0 . This fit is independent of the choice of anchors. These fit parameters can be combined with the previous absolute SNIA peak magnitude, M_B , fit to calculate the value of Hubble's constant,

$$H_0 = 10^{0.2M_B + a_B + 5}. \quad (25)$$

2.3. Main Sequence Fitting

To test the reddening law variation in the Milky Way, this investigation will use the [CG18](#) catalogue of open clusters. Following [CG18](#) values of the extinction, $E(B - V)$, will be taken from [Kharchenko et al. \(2013\)](#), hereafter K13. This provides a set of 1,073 clusters whose distances range between $\sim 0.1 - 10$ kpc, spanning a significant portion of the Milky Way disc. The distribution of clusters in the galactic coordinates can be seen in figure 2, with the majority of clusters lying in the galactic plane.

All analysis will be performed in the Gaia passbands, G , G_{BP} and G_{RP} with the star colour being taken as $(G_{BP} - G_{RP})$. To correct for dust extinction, the following formulae will be used:

$$G_0 = G - R_V E(B - V), \quad (26)$$

and

$$(G_{BP} - G_{RP})_0 = \frac{1}{2} R_V E(B - V), \quad (27)$$

where a subscript 0 denotes an intrinsic value. These derive from the relations that $A_G \sim 3.1E(B - V)$ ([Malhan et al. 2018](#)) and $A_G \sim 2E(BP - RP)$ ([Andrae et al. 2018](#)). Note that

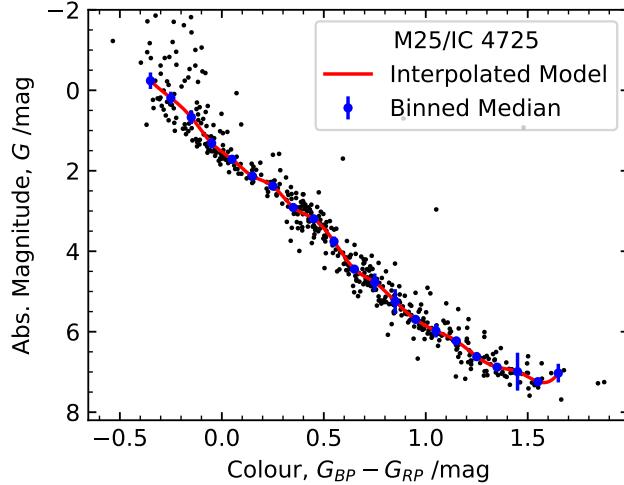


FIG. 3: Plot of a main-sequence model produced by interpolating the binned medians of a CMD of the open cluster M25.

the similar formulae quoted in Tisserand et al. (2022) are incorrect and cannot be used to correct for dust extinction.

Before performing main-sequence (MS) fitting, a MS model must first be prepared. As it is close to Earth and provided a good UBV fit in Hoyle et al. (2003), the model used in the investigation will be based on the MS of the open cluster M25 (IC 4725 in the CG18 catalogue). After colour correction, a series of binned medians in 0.1 mag bins will be made to the CMD plot and a continuous MS model produced by cubic interpolating the medians. The data, medians and model can be seen in figure 3.

The goodness-of-fit to the M25 MS model is then evaluated as a reduced chi-squared statistic calculated using the formula,

$$\chi_{\nu}^2 = \frac{1}{N-1} \sum_i \left(\frac{G_i - f[(BP - RP)_i]}{\sigma} \right)^2, \quad (28)$$

where σ is the standard deviation of the fit residuals, $\{(BP - RP)_i, G_i\}$ is the set of N stars which overlap with the range of the M25 MS model, $-0.4 < (BP - RP) < 1.8$ mag, and $f[(BP - RP)]$ is the interpolated MS model. Normalising the statistic by the degrees of freedom of the fit, $\nu = N - 1$, allows for comparison between clusters.

To filter out clusters which are misidentified, combinations of multiple clusters, or have a sufficiently different age to M25 two arbitrary offsets, ΔG and $\Delta(BP - RP)$, are used to shift the model and are fitted to the cluster data. If the best fit offsets result in a chi-squared statistic of $\chi_{\nu}^2 > 2$, the cluster is considered to be bad and is removed from the data. The chi-squared statistic is then recalculated for the clusters without the arbitrary offsets and the variation of goodness-of-fit with distance modulus, $(m - M)$, is plotted.

To investigate the possibility of a distance dependent reddening law, the law is modified like so,

$$A_G \sim A_V = R'E(B - V) = 3.1 \cdot X \left(\frac{d_{\text{cluster}}}{d_{\text{M25}}} \right) \cdot E(B - V), \quad (29)$$

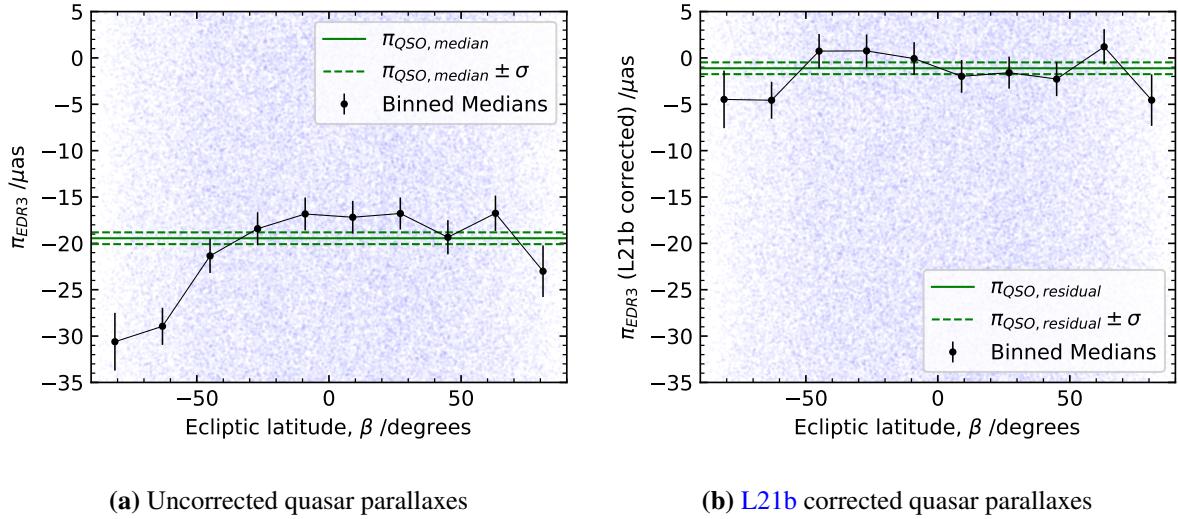


FIG. 4: Plot of 1,614,173 quasar EDR3 parallaxes before (left) and after applying the [L21b](#) correction (right). The quasar parallaxes are plotted in blue. There is a residual median parallax of $-1.1 \pm 0.6 \mu\text{as}$ after correction.

where d_{cluster} is the distance to the open cluster, d_{M25} is the distance to M25 and X is a fit parameter which will be referred to as the “X-factor”. This modified law is normalised such that $X = 1$ when fitted to M25.

3. RESULTS AND DISCUSSION

3.1. Gaia EDR3 Parallax Bias

3.1.1. Ecliptic Latitude Dependence

The results of analysing the EDR3 parallax bias dependence on ecliptic latitude can be seen in figure 4. Figure 4a shows the raw parallaxes before applying the [L21b](#) correction while figure 4b shows the corrected parallaxes. Binned medians are shown to indicate the variation of the parallax bias with ecliptic latitude. On both plots, the median of the whole sample, $\pi_{QSO, median}$, and its standard error has been plotted.

In the uncorrected plot, it is clear that there is a definite dependence on ecliptic latitude, with quasars at low ecliptic latitudes tending to be below the median of the whole sample. This is indicative of the EDR3 parallax bias as the requirement for Gaia’s heatshield to point at the Sun keeps the satellite aligned with the ecliptic plane. As well as the ecliptic latitude variation, there is an overall median parallax of $-19.4 \pm 0.4 \mu\text{as}$. This is smaller than the median quasar parallax of $-29 \mu\text{as}$ found for DR2 ([Luri et al. 2018](#)), showing an improvement in raw parallax measurements. It is hoped that the future release of Gaia DR4 will be able to improve these raw measurements further.

After [L21b](#) correction it can be seen that much of the ecliptic latitude variation has been removed, with binned medians which are consistent with the global median. As such, there is no need to further correct for ecliptic latitude dependence. Despite this, there is a residual

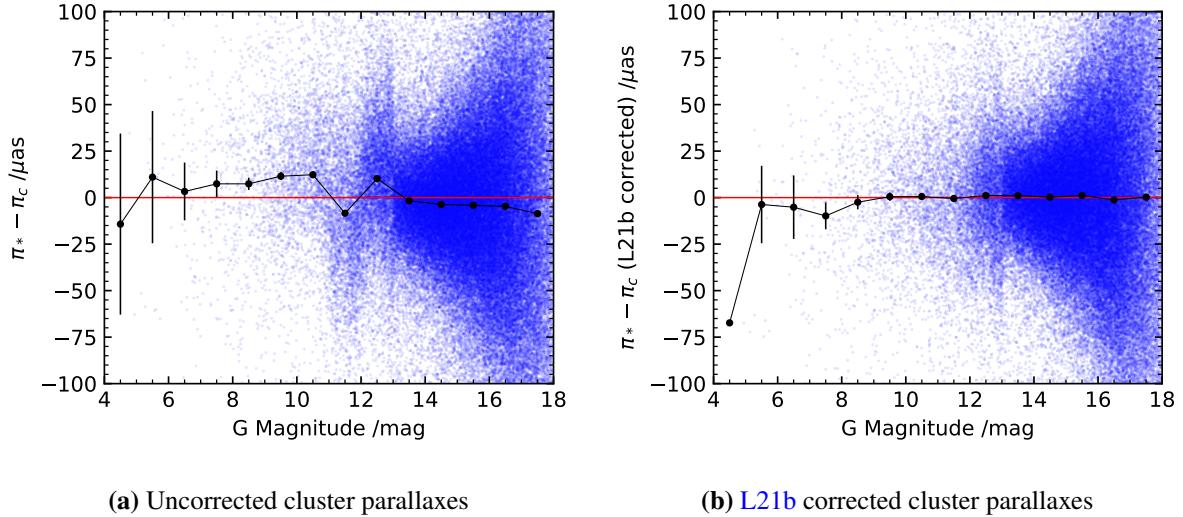


FIG. 5: Plot of 237,535 CG18 cluster EDR3 parallaxes before (left) and after applying the L21b correction (right). The star parallaxes are plotted in blue and the binned medians in black. A red line is plotted at zero to show the ideal unbiased parallax difference.

median parallax, $\pi_{\text{QSO,residual}} = -1.1 \pm 0.6 \mu\text{as}$, that will be considered when calculating zp as in equation 12. This is a small offset, on the order of 0.1–1% for Milky Way Cepheid parallaxes, but may have a non-negligible effect on the value of Hubble’s constant.

3.1.2. *G* Magnitude Dependence

The results of finding the relative cluster star parallaxes from the 237,535 CG18 stars are shown in figure 5. The uncorrected raw EDR3 parallaxes are shown in figure 5a while the L21b corrected parallaxes are shown in figure 5b. On both plots the binned medians in 1 mag bins have been plotted to indicate the G magnitude dependence of the parallax bias and a line at $\pi_* - \pi_c = 0$ has been plotted to show the ideal median value.

For the uncorrected parallaxes, it is evident that there is a parallax bias at all magnitudes, particularly for brighter and dimmer stars. In particular the deviation at 12th magnitude is due to the overlap between gates, mechanisms which are used to mitigate saturation of the Gaia detector ([L21b](#)).

For most part the L21b correction corrects for the magnitude dependence, removing the deviation at 12th magnitude and moving the median relative parallax to zero for stars dimmer than 9th magnitude. Despite this, there is still an offset for bright stars with $G < 9$ mag. This represents an overcorrection by the Gaia formulation and, given that the R21 set of Cepheids have magnitudes in the range $5 < G < 10$ mag, it will be necessary to take this overcorrection into account before anchoring the PL relation.

This zeropoint offset for bright stars has been identified from a number of different sources (Flynn et al. 2022, Freedman 2021, Ripepi et al. 2021, Zinn 2021) and can be seen in more detail in figure 6. Here a median residual parallax and its standard error for stars with $G < 9$ mag has been plotted. This median parallax, $\pi_{\text{clus,median}} = -3 \pm 4 \mu\text{as}$, is the value of the bright

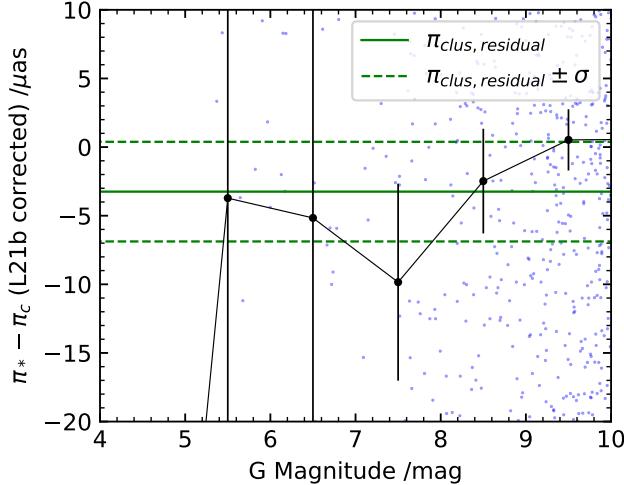


FIG. 6: Residual bright star median, $\pi_{\text{clus,median}} = -3 \pm 4 \mu\text{as}$, after L21b correction for CG18 cluster relative parallaxes. Cluster star parallaxes are plotted in blue.

star residual median that will be used in this investigation. Since this is a median of relative parallaxes only, to find a parallax offset it is necessary to combine it with the global quasar parallax offset as in equation 12. This gives an overall parallax zeropoint of $zp = +4 \pm 4 \mu\text{as}$.

As with the quasar parallax offset, this is a small quantity, corresponding to a 0.4–4% offset for Milky Way Cepheids, although for the furthest stars this may have a non-negligible effect and so this offset will be applied to all stars with $G < 9$ mag. This can be compared with the zeropoint offset found by Maíz Apellániz (2022) who used a similar method with stellar clusters to calculate an independent alternative to the L21b correction. Their zeropoint offset of $zp_{\text{MA22}} = +6.9 \pm 2.2 \mu\text{as}$ is of the same order and sign as this work’s zeropoint, validating this approach.

A major drawback of the stellar cluster approach to determining the bright star zeropoint is the lack of stars at bright magnitudes. In the CG18 catalogue there is just a single star with a magnitude $G < 5$ mag, so any medians in this range are highly unreliable. As a result, it is not possible to properly correct the parallaxes of the brightest Cepheids, with their parallax bias being inferred from stars in range $5 \lesssim G < 9$ mag.

This work has followed the methods of other papers in assuming that the bright star zeropoint is a constant and independent of magnitude, however it is possible that this is a poor assumption, as figure 6 may suggest. Rather than taking a single median and applying this to all bright stars, it may be necessary in future to fit a linear offset dependence with magnitude or even a higher-order polynomial fit.

3.2. Comparing EDR3 with HST Parallaxes

Alongside Gaia parallaxes for Milky Way Cepheids, a set of trigonometric parallaxes measured using the Hubble Space Telescope are used to anchor SH0ES measurements of the Hubble constant. These parallaxes are a combination of 7 Cepheids observed with the HST’s Wide Field Camera 3 (WFC3) (Benedict et al. 2007) and 10 Cepheids observed using spatial scanning with

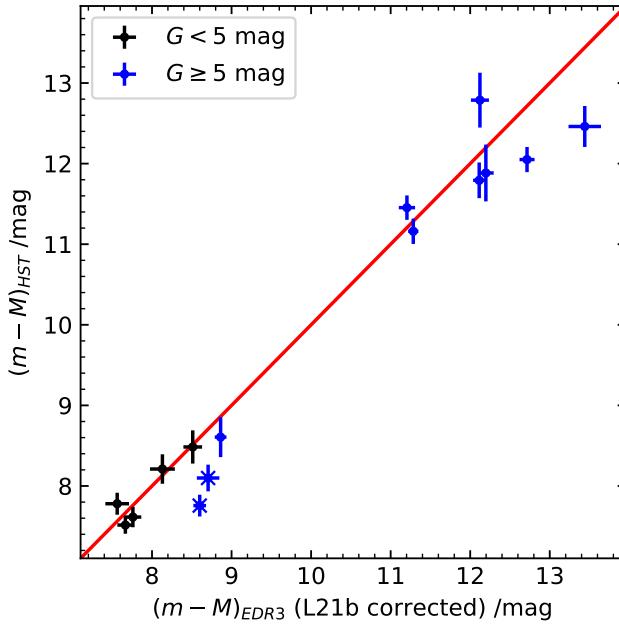


FIG. 7: Comparison of EDR3 distance moduli with HST trigonometric distance moduli for a sample of 15 Milky Way Cepheids. The stars RT Aql and FF Aur have been marked with crosses due to the difference between the two distance samples as per [Owens et al. \(2022\)](#). Plotted in red is the line of 1:1 correspondence.

the HST’s Fine Guidance Sensor (FGS) ([Riess et al. 2018a](#)) to form a set of 17 Cepheids in total. For comparison to earlier papers, this investigation will follow the method of [Riess et al. \(2018b\)](#) and exclude δ Cep and Y Sgr, leaving a set of 15 Cepheids.

These HST parallaxes can be compared with Gaia EDR3 parallaxes, similar to methods performed in [Shanks et al. \(2019\)](#) and [Owens et al. \(2022\)](#), the results of which are shown in figure 7. While at distances below 10 mag the stars lie close to the 1:1 line, there are some outliers, namely RT Aql and FF Aur which have been highlighted with crosses and were also identified as outliers in [Owens et al. \(2022\)](#). At higher distances, there is a larger difference of order ~ 1 mag between EDR3 and HST parallaxes, with few of the stars lying within error of the 1:1 line.

The points in figure 7 have been coloured according to their G magnitude to investigate whether saturation effects are causing the distance mismatch. It can be seen that this is not the case as the dimmer stars appear to be further than the brighter stars from the line. As it has been shown that the L21b correction adequately corrects parallaxes to within a few microarcseconds, this mismatch must be primarily due to errors in the HST trigonometric parallaxes. Given these HST parallaxes are still used to anchor the SH0ES measurements ([R22](#)), it may have a noticeable effect on the value of the Hubble constant.

3.3. The Hubble Constant

The results of using the SH0ES method to determine Hubble’s constant will be presented in this section. The galaxy distance moduli and absolute SNIA peak magnitude given here will be

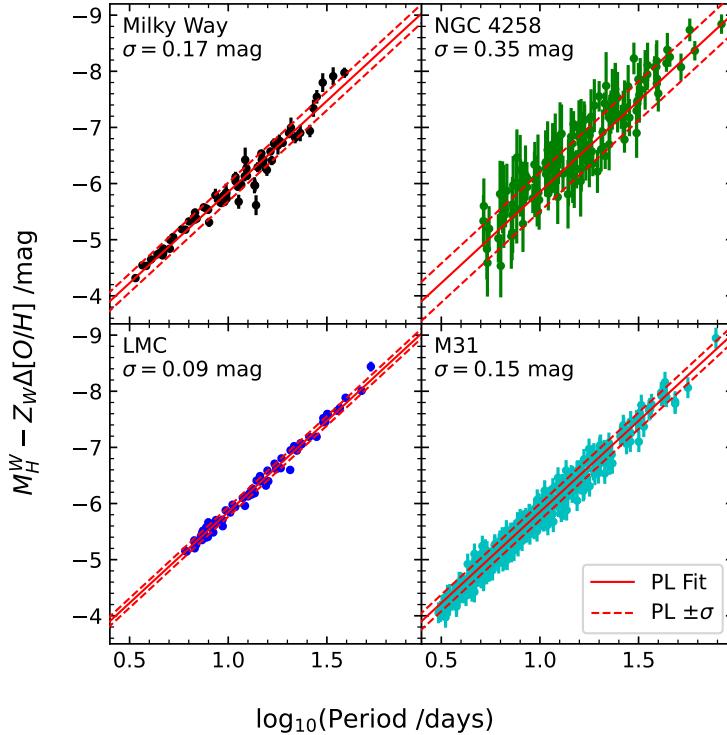


FIG. 8: Plot of the PL relation fit for all four geometric distance anchors, Milky Way EDR3 parallaxes, LMC and M31 Detached Eclipsing Binaries and NGC 4258 Water Masers. Plotted in red is the PL relation fit (solid line) and the corresponding PL dispersion, σ , (dashed line). All stars have been shifted by their metallicity for comparison. In this plot, M31 Cepheids were not fitted for and are shown for reference.

calculated using the best combination of all anchors - Milky Way Cepheids with L21b and zp correction, LMC and N4258 - to demonstrate the individual results of the rungs of the distance ladder. Later, the choice of anchors will be varied and their effect on the fit parameters and the distance ladder as a whole will be demonstrated.

3.3.1. SH0ES Method

The results of fitting the period-luminosity relation in equation 6 are shown in figure 8. M31 was not fitted for in this plot, and is shown for reference. In the figure are Cepheids in the four distance anchors and the simultaneously fitted PL relation and its dispersion, σ , measured as the standard deviation of the residuals. For the Milky Way and the LMC, the fit is good and the dispersion is small. However for NGC 4258 there is a larger dispersion and the points tend to lie above the PL fit. This may be a result of either a poor geometric distance measurement to NGC 4258 or poor metallicity measurements, both of which could contribute to a vertical shift of the points in figure 8.

Since it is based on measurements of only two DEBs, the error in the geometric distance measurement to M31 is large at 4% and may be insufficient to calibrate the PL relation accurately. As a result, calibrations of Hubble's constant which use M31 as a distance anchor in this

work will not be considered when determining the best combination of anchors. This is similar to the method adopted by R16 who include the distance modulus of M31 as a fit parameter rather than using the geometric distance measure.

The SH0ES metallicity measurements are discussed in Mortsell et al. (2021b), who suggest that R21’s assumption of $[Fe/H] = [O/H]$ should be modified to give $[O/H] = [Fe/H](1 + k)$ with a constant value $k \sim -0.5 \pm 0.25$. They find that this affects the calibration of the Milky Way Cepheids, changing the SH0ES measurement of H_0 from $73.2 \pm 1.3 \rightarrow 72.1 \pm 1.4$.

There is no indication from the data of a discontinuity at $\log_{10} P = 1$ as suggested by Ngeow et al. (2005), although a more detailed investigation would be needed to determine the existence of the discontinuity. Fits by R16 and R22 have found $< 1\sigma$ evidence of a break in the relation at $P = 10$ days and use a single slope fit, as in this work. Future investigations could make use of the Gaia DR2 catalogue of Cepheid variable stars (Holl et al. 2018) which provides a large set of 8,890 fundamental classical Cepheids in the Milky Way and LMC, providing a rich dataset to explore the P-L relation. This catalogue is expected to be superseded in June 2022 with the release of the full Gaia Data Release 3.

The results of fitting the period-luminosity relation for 19 Cepheid-SNIA hosts to determine their distance moduli can be seen in figure 9. Plotted in red is the PL relation (solid line) and the dispersion (dashed line), σ , which is calculated by taking the standard deviation of the fit residuals. For each galaxy the fitted distance modulus has been given with its respective standard error.

For many of the galaxies, the distance modulus has been determined to a precision of the order of a tenth of a percent. This demonstrates the competency of this method for determining distances to distant objects with Cepheid photometry. This is demonstrated most strikingly with the galaxy NGC 4424, where the distance modulus has been determined with an error of 0.3% despite the PL relation being fitted to just three stars. It is clear, however, that more precise determinations of galaxy distances will require further observations of Cepheids in distant SNIA hosts. It is hoped that over the next few years these Cepheid sets will continue to grow at a fast pace, with examples of new improvements including R22 which increases the number of observed Cepheids in N4424 from three to nine, and increases the total number of Cepheid-SNIA hosts from 19 to 37.

Table I shows a plot of the Cepheid distances from R16 compared with this work’s Cepheid distances to the 19 SNIA-Cepheid hosts. It can be seen that while the two distances are consistent, this work’s distances are generally shorter than those found by R16. As this corresponds to a greater SNIA absolute peak magnitude, this difference will result in a larger value of H_0 , worsening the tension with Planck’s H_0 value.

The results of fitting a constant absolute peak B-band magnitude to the Cepheid host SNIAe are shown in figure 10. The solid red line represents a 1:1 correspondence between the supernova and Cepheid distances while the dashed red lines represent the dispersion of the data, $\sigma = 0.12$ mag, calculated as the standard deviation of the residuals. The dispersion of the SNIA distance moduli is consistent with the intrinsic scatter of ~ 0.15 mag for SNIA luminosities (Sanders et al. 2021), the cause of which is unknown.

While these methods rely solely on Cepheid distances to determine the SNIA magnitude, a growing sample of Tip of the Red Giant Branch (TRGB) stars can be used to add an independent distance measure to complement Cepheid distances (De Jaeger et al. 2022, Dhawan

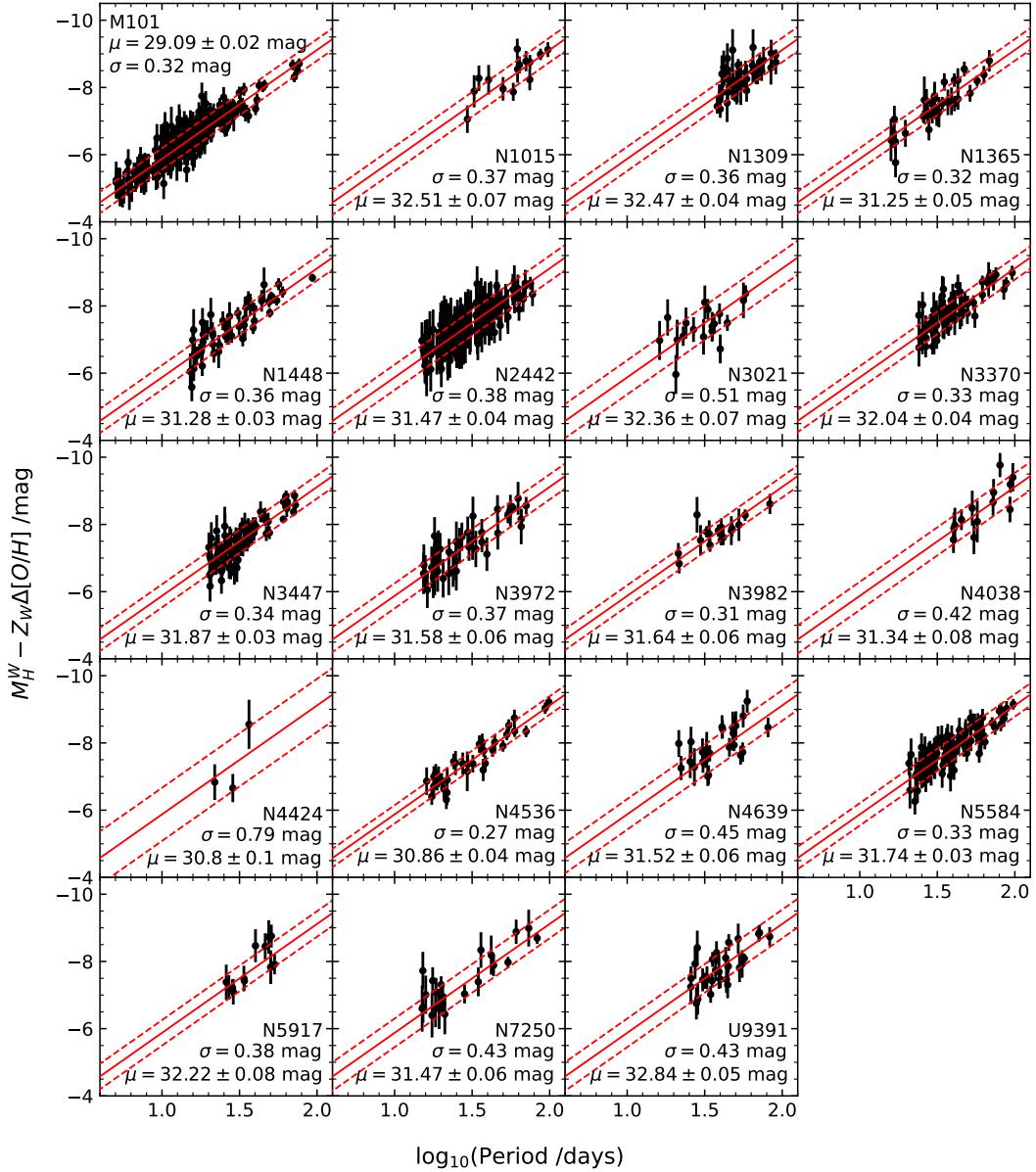


FIG. 9: Plot of the PL relation fits for all 19 SNIa and Cepheid host galaxies. Plotted in red is the PL relation (solid line) and the PL dispersion, σ , (dashed line). All stars have been shifted by their metallicity for comparison.

et al. 2022). This sharp discontinuity in the evolutionary track of a red giant star has an I-band magnitude which is insensitive to the stellar metallicity and so can be used as a reliable standard candle for distance measurements (Sakai 1999). Preliminary results suggest that TRGB distance measurements support the SH0ES value of H_0 over the Planck value, although a larger set of TRGB-SNIa hosts is required before a Cepheid-independent TRGB value of H_0 can be considered.

The results of fitting the final rung of the SH0ES distance ladder are shown in table II and figure 11. As the form of the redshift-distance relation used here is independent of the value of M_B , the fit parameters in figure 11 are independent of the choice of anchors and are found to be $q_0 = -0.61 \pm 0.01$ and $a_B = 0.7179 \pm 0.0008$. The value of the jerk parameter is not fitted

Galaxy	$\mu_{\text{This Work}} / \text{mag}$	Error	$\mu_{\text{R16}} / \text{mag}$	Error
M101	29.09	0.02	29.135	0.045
N1015	32.51	0.07	32.497	0.081
N1309	32.47	0.04	32.525	0.055
N1365	31.25	0.05	31.307	0.057
N1448	31.28	0.03	31.311	0.045
N2442	31.47	0.04	31.511	0.053
N3021	32.36	0.07	32.498	0.090
N3370	32.04	0.04	32.072	0.049
N3447	31.89	0.03	31.908	0.043
N3972	31.58	0.06	31.587	0.0702
N3982	31.64	0.06	31.737	0.069
N4038	31.34	0.08	31.290	0.112
N4424	30.8	0.1	31.080	0.292
N4536	30.86	0.04	30.906	0.053
N4639	31.52	0.06	31.532	0.071
N5584	31.74	0.08	31.786	0.046
N5917	32.22	0.08	32.263	0.102
N7250	31.47	0.06	31.499	0.078
U9391	32.84	0.05	32.919	0.063

TABLE I: Comparison of this work’s Cepheid distances to SNIA-Cepheid hosts with the Cepheid distances from [R16](#). It can be seen that this work’s distances are generally shorter than those found by [R16](#).

for and is taken to be $j_0 = 1$ ([R16](#)).

The effect of the choice of anchors on all fit parameters can be seen in appendix [A](#) and all three rungs of the distance ladder can be seen together in appendix [B](#). The best combination of anchors is found to be Milky Way EDR3 [L21b](#) and zp corrected parallaxes, LMC and NGC 4258, resulting in a value of $H_0 = 75.2 \pm 1.2$, similar to the results of the SH0ES team who found in [R16](#) that the best combination was also MW, LMC and N4258. The best-fit PL parameters are $M_{H,1}^W = -5.875 \pm 0.006$ mag, $b_W = -3.24 \pm 0.02$ mag and $Z_W = -0.13 \pm 0.03$ mag dex $^{-1}$, with a best-fit SNIA peak magnitude of $M_B = -19.21 \pm 0.03$. While this anchor combination has a larger chi-squared statistic than other fits, it is not affected by the inaccurate distance to M31 and has been corrected for the bright star zeropoint so was chosen as the best combination. Comparing this to the analysis performed in [R22](#), the SH0ES team now incorporate all four anchors alongside other sources such as TRGB measurements and ground-based Cepheid photometry.

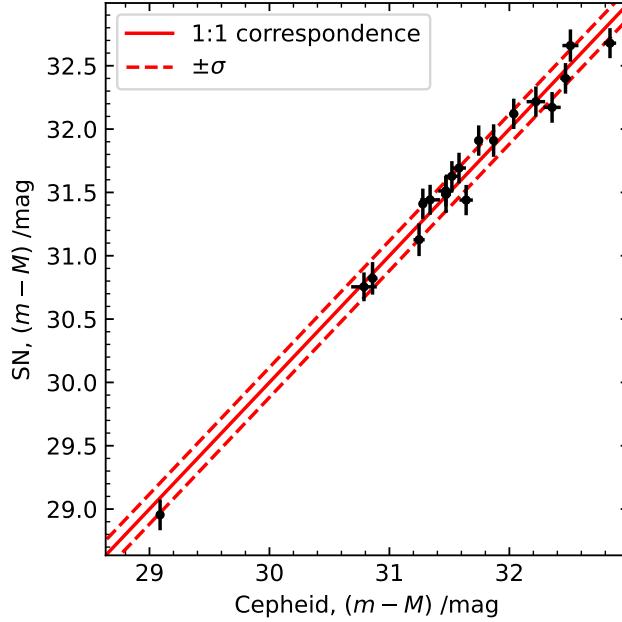


FIG. 10: Result of calibrating the SNIA distance moduli to the Cepheid distance moduli. The solid red line represents a 1:1 correspondence between the supernova and Cepheid distances while the dashed red lines represent the dispersion of the data, $\sigma = 0.12$ mag. This is consistent with the intrinsic SNIA luminosity scatter of ~ 0.15 mag

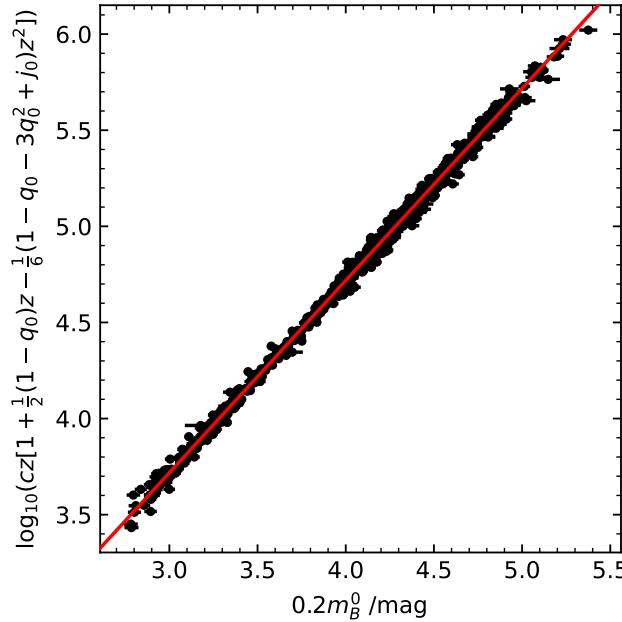


FIG. 11: Result of calibrating the redshift-distance relation on a Hubble diagram for 1,048 supernovae. The fit shown here is for $H_0 = 75.2 \pm 1.2$ and $q_0 = -0.61 \pm 0.01$.

For multiple anchors the slope of the PL relation, b_W , is insensitive to change, staying consistent with the R16 value of -3.26 ± 0.02 mag. However for single anchors, this slope varies greatly with the unexpectedly low slope from NGC 4258 likely being caused by the high scatter

Anchors	H_0	σ	χ^2_ν
<i>Four Anchors</i>			
MW (raw) + M31 + LMC + N4258	75.4	1.2	1.16
MW (L21b) + M31 + LMC + N4258	76.3	1.2	1.13
MW (L21b + zp) + M31 + LMC + N4258	76.4	1.2	1.13
<i>Three Anchors</i>			
MW (L21b + zp) + LMC + N4258	75.2	1.2	1.24
MW (L21b + zp) + M31 + N4258	76.9	1.2	1.09
MW (L21b + zp) + M31 + LMC	77.0	1.2	1.12
<i>One Anchor</i>			
MW (raw)	74.0	1.3	1.17
MW (L21b)	77.8	1.4	1.19
MW (L21b + zp)	78.7	1.4	1.18
M31	79.2	1.4	0.99
LMC	74.5	1.6	1.19
N4258	73.0	1.4	1.09

TABLE II: Table demonstrating the effect of the choice of anchors on the calculated value of Hubble constant. Highlighted in bold is the preferred anchor choice.

in the data. The best constraint on the slope with the smallest error comes from the LMC, which gives a slope of $b_W = -3.29 \pm 0.02$ mag.

As can be seen in table II, applying the correction for a bright star zeropoint, zp , has a $\sim 1\%$ effect on the Hubble constant when the Milky Way is the sole anchor. However, when constrained by other anchors, this offset is reduced to a small effect of the order of $\sim 0.01\%$ on the value of the Hubble constant. By making EDR3 parallaxes larger and hence Cepheid distances smaller, it widens the gap between the SH0ES and Planck H_0 values and as such is not able to explain the Hubble tension.

3.3.2. Cepheid-Redshift Diagram

While most approaches calibrate the redshift-distance relation from supernovae distances, it is possible to calibrate the redshifts and hence the Hubble constant directly from Cepheid distances. This leads to the interesting result seen in figure 12 where the best-fit Hubble constant is found to be $H_0 = 66.7 \pm 0.1$, which is in close agreement with the Planck value. Here, the solid red line is the line of 1:1 correspondence while the dashed red lines are the dispersion, $\sigma = 0.65$ mag, or the standard deviation of the residuals.

Since none of the SH0ES SNIA-Cepheid host galaxies are in the R16 Pantheon catalogue, the Carnegie Supernova Project I, Data Release 3 (CSP-I DR3) (Burns et al. 2018) has been

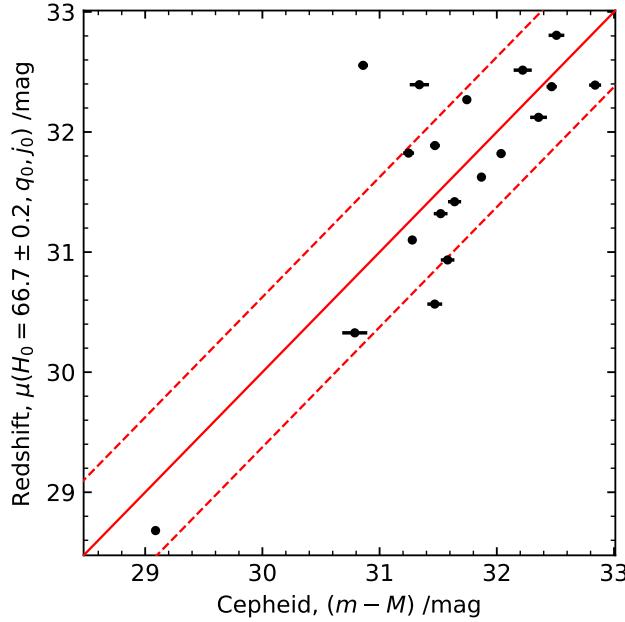


FIG. 12: Plot of redshift distance modulus against Cepheid distance modulus with a fit value of $H_0 = 66.7 \pm 0.2$ where the deceleration parameter is taken to be $q_0 = -0.61$. The solid red line is the line of 1:1 correspondence while the dashed red lines are the dispersion, $\sigma = 0.62$ mag, of the residuals.

used, providing CMB-frame redshifts for all 19 SNIA-Cepheid host galaxies. Note that CSP-I is now incorporated into the Pantheon+ catalogue (Scolnic et al. 2022). Although these galaxies are local and are not dominated by the Hubble flow, the CMB frame correction should remove the effects of peculiar velocities.

As redshift measurements were not a part of CSP-I, the catalogue uses data from a combination of the NASA/IPAC Extragalactic Database and the Sloan Digital Sky Survey (Krisciunas et al. 2017) and are quoted without error. As a result it is not possible to determine how much of the scatter is caused by uncertainties in the redshift measurements, with the Cepheid distance errors serving as the only reference.

3.4. Galactic Reddening

The results of fitting the M25 MS model to clusters in the CG18 catalogue using the standard reddening law with $R_V = 3.1$ are shown in figure 13. It is evident from the figure that there is an increase in goodness-of-fit with distance, supporting the results from visual inspection of Hoyle et al. (2003). As the data has been filtered to remove clusters with main-sequences which have a different morphology to that of M25, it is unlikely that this increase with distance is due to a change in the population type as the remaining clusters will be of a similar age to M25 and will not be data aberrations such as two clusters being counted one.

Examples of cluster color-magnitude diagrams (CMD) with the M25 MS model can be seen in appendix C. These are an assortment of clusters at varying distances to demonstrate how the goodness-of-fit measurement corresponds to the data. In general the closest clusters are excellent fits to the model, with very little scatter and low goodness-of-fit, χ^2_ν . For clusters

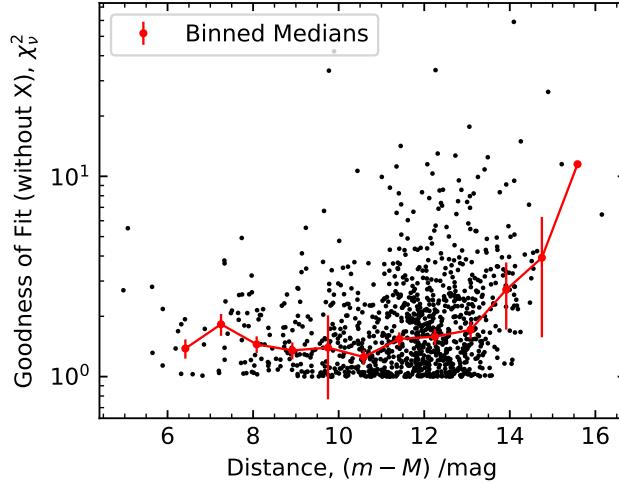


FIG. 13: Plot of goodness-of-fit, χ^2_ν , against distance modulus for the M25 MS fit to open clusters in the CG18 catalogue using the standard reddening law with $R_V = 3.1$. It can be seen that there is an increasing trend between goodness-of-fit and distance, suggesting that a distance-dependent term needs to be incorporated into the dust reddening law.

at a further distances, the scatter of the data is higher but the data still fits the model well, although for some clusters there is evidence of other non-main sequence features in the data which increase χ^2_ν . For the furthest clusters the data is very poor with a large scatter and few stars, lying away from the model. For some distant clusters, the shape of the M25 main sequence does not match the shape of its CMD and so these clusters will be filtered out when applying the arbitrary shifts to the M25 model.

One possible explanation for the worsening of the M25 fit as distance increases could be inaccurate values of $E(B - V)$ from K13. An investigation by Cantat-Gaudin et al. (2020), which determined the absorption, A_V , using machine learning found smaller values of $E(B - V)$ at higher extinctions than those found by K13. Given that $E(B - V)$ is expected to increase with distance (Obermeier et al. 2016), it is possible that use of this catalogue could contribute to the increase in goodness-of-fit.

The results of fitting the “X-factor” from equation 29 to the CG18 clusters are shown in figure 14. In red are the values of X and R'_V expected from the standard reddening law. It can be seen that for most clusters, the X-factor follows a $X \propto 1/d_{\text{cluster}}$ relation, which corresponds to the standard $R'_V = 3.1$. However, for higher distances there is more scatter in R'_V and more of a trend towards higher reddening. This is expected behaviour as in Schlafly et al. (2017) it was found that reddening generally increases with distance.

The increased scatter in R'_V at higher distances may be due to the inclusion of all lines-of-sight in the same dataset. It was seen in Schlafly et al. (2016) that $R(V)$ depends on the position as well as the distance of the object so it may be necessary to consider a more complex form of the reddening law which includes position terms. In the Milky Way these position terms would depend on structures in the galaxy, and could be used to determine the size and shape of the spiral arms.

The effect of using the modified reddening law can be seen in figure 15. The red line is the

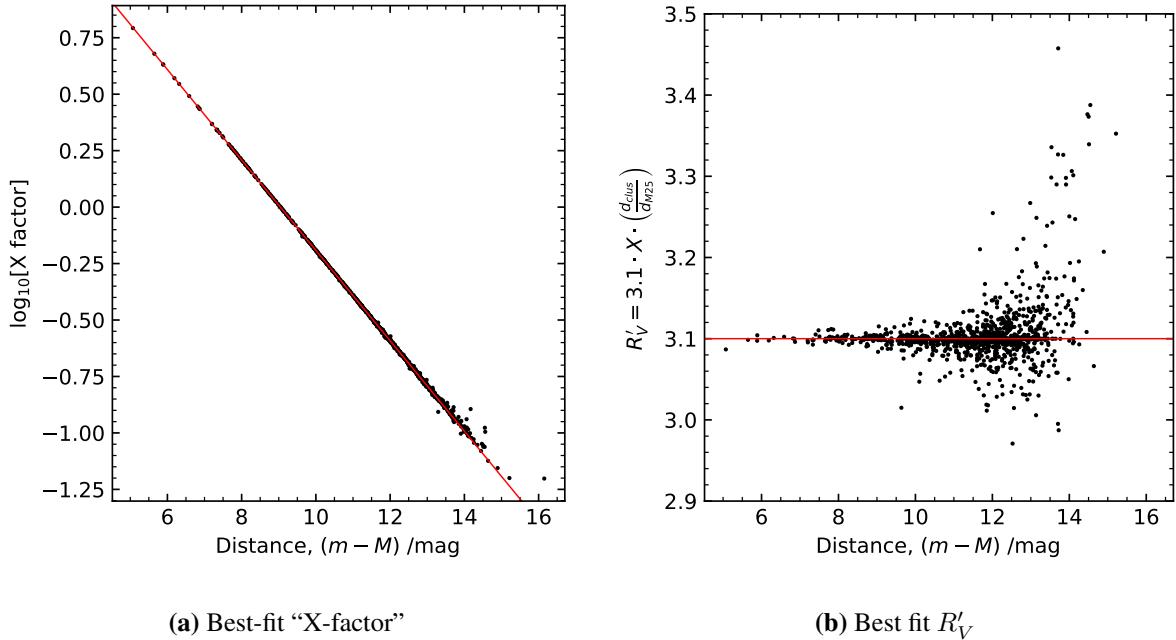


FIG. 14: Plot of the best-fit correction to the reddening law against distance modulus. The results expected from the standard reddening law have been plotted as a red line. In general the X-factor follows a $1/d_{\text{cluster}}$ relation, giving the standard $R'_V = 3.1$, however at higher distances there is some deviation from the expected law.

line of 1:1 correspondence indicating no change in goodness-of-fit. It can be seen from the plot that there is a negligible change in the goodness-of-fit due to the use of the modified reddening law. This may indicate that a more complicated form of the distance dependence of the law is required as distant clusters are still being found to be a poor fit to the model.

As suggested by Mortsell et al. (2021a), it is possible to apply the reddening law results found here to the calibration of Hubble’s constant. Within the galaxy, the modified reddening law can be used to recalculate the Milky Way Cepheid Wesenheit magnitudes. Similarly, R_E can be kept as a free parameter when fitting the SNIA hosts, allowing the properties of the dust in these galaxies to vary.

4. CONCLUSIONS

4.1. Gaia EDR3

Using quasars to determine the dependence of the parallax bias on ecliptic latitude, it was found that the correction suggested by the Gaia team (L21b) removed much of the latitude variation but undercorrected quasar parallaxes, leaving a residual median parallax of $-1.1 \pm 0.6 \mu\text{as}$. Using open clusters to determine the G magnitude dependence demonstrated that while for stars with $G > 9$ mag the L21b correction removed the bias, a residual undercorrection was left for brighter stars with $G < 9$ mag, corresponding to a median parallax of $-3 \pm 4 \mu\text{as}$.

Combining the two medians gave a parallax zeropoint of $zp = +4 \pm 4 \mu\text{as}$, a small quantity which corresponds to a 0.4 – 4% offset for Milky Way Cepheids. This offset is of the order

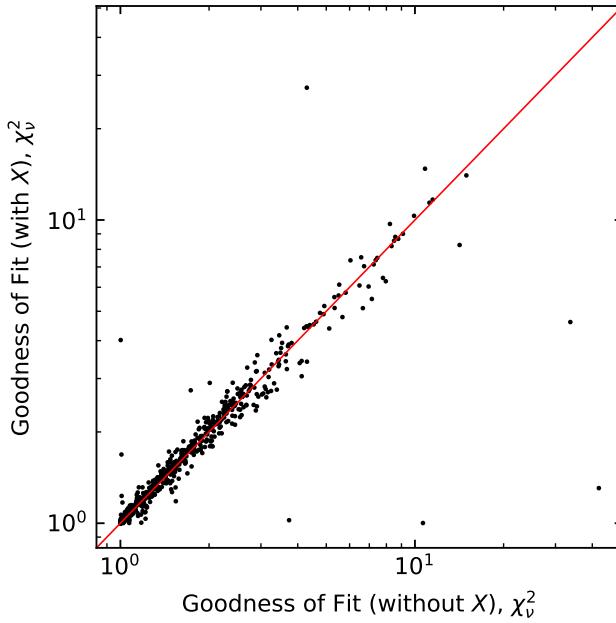


FIG. 15: Comparison of the goodness-of-fit using the standard reddening and using the modified “X-factor” reddening law. The red line represents a one-to-one correspondence. It can be seen that including the distance dependence has a negligible effect on the quality of the fit.

of the zeropoint found by [Maíz Apellániz \(2022\)](#), helping to validate this approach. While for closer Cepheids the offset is negligible, more distant Cepheids will need to be corrected so this correction was applied to the [R21](#) Cepheids to investigate its effect on the value of H_0 . By using the cosmic distance ladder method of the SH0ES team, it was found that including the zeropoint offset had a small effect of $\sim 0.01\%$ on the Hubble constant when combined with other distance anchors, which is not enough to account for the Hubble tension.

The main issue with this method is the lack of bright open cluster stars, resulting in spurious results for $G \lesssim 5$ mag. For better estimates of the parallax zeropoint offset it is necessary to find larger sets of bright stars with independently verifiable parallax measurements, taking into account the colour dependence of the Gaia EDR3 bias. Similarly, the assumption that the bright star offset is a constant for all bright stars may be inappropriate and could be improved using a linear or higher-order polynomial fit to correct for the residual offset.

In future it is hoped that the release of Gaia Data Release 4 will provide even more accurate parallaxes and the biases exhibited in EDR3 will either be reduced or eliminated. In turn this will allow the Cepheid PL relation and hence the value of H_0 to be constrained even further, potentially increasing an already significant tension with the Planck value.

4.2. Galactic Reddening

By fitting a main-sequence model, derived from the main-sequence of M25, to Milky Way open clusters, it was possible to determine that the quality of the fit declined as the distance to the cluster increased. This was an indicator that the standard reddening law — which assumes $A_V = R_V E(B-V)$ where the ratio of the total absorption to the selective absorption, $R_V \sim 3.1$,

is a constant — may need to be modified to include a distance dependence.

A modified version of the reddening law included a linear dependence on the distance to the cluster and a free parameter, X , which was fitted to each cluster. The result of applying this modification found that for lower distances, the modified law followed the standard law, while for higher distances the scatter in R'_V increased with the values skewed towards $R'_V > 3.1$. It is possible that the increased scatter is due to all lines-of-sight being included in the same plot, indicating that further modifications to the law must be made to include position terms.

This modified law will have an impact on the value of Hubble’s constant, as the furthest Cepheids will have an R'_V that differs from 3.1, requiring the Wesenheit magnitudes to be recalculated. Similar modifications can also be applied to SNIA-Cepheid hosts, as done by [Mortsell et al. \(2021a\)](#), implying that the dust properties vary between galaxies. Since dust extinction is still not well-understood, this offers a promising explanation that could help to resolve the Hubble tension.

4.3. Hubble Tension

Following a modified version of the method outlined by the SH0ES team in [R22](#), using a the combination of [L21b](#) and zp corrected EDR3 Milky Way parallaxes, LMC and NGC 4258, it was possible to determine the best-fit value of the Hubble constant of $H_0 = 75.2 \pm 1.2 \text{ km sec}^{-1} \text{ Mpc}^{-1}$ which is consistent with the current best SH0ES value. As discussed, the inclusion of a bright star zeropoint offset due to undercorrection by the Gaia team’s bias correction formulae had a negligible effect and was unable to explain the Hubble tension.

Some issues with method which could contribute to the tension with the Planck H_0 value remain, including the assumption that for the metallicity $[\text{Fe}/\text{H}] = [\text{O}/\text{H}]$ for Milky Way Cepheids. Another possible issue and way to reduce the Hubble tension is by changing the way that dust is treated by allowing the reddening law to vary between galaxies. As the physics behind dust reddening is still poorly understood, this could be a plausible explanation for the tension, as suggested by this work’s investigation of reddening within the Milky Way.

A particularly interesting result is achieved when calibrating redshifts and hence Hubble’s constant directly from Cepheid distances, yielding a best-fit value of $H_0 = 66.7 \pm 0.2 \text{ km sec}^{-1} \text{ Mpc}^{-1}$. There is a large scatter in the data but since the redshifts are quoted without error, it is not possible to determine what is driving it. With more Cepheid-SNIA hosts, the accuracy of this value could be improved, potentially giving a value of H_0 which is consistent with the Planck value.

Comparing EDR3 parallaxes with trigonometric HST parallaxes demonstrates that there is a mismatch between the two distance measurements. Given that the offset is larger than can be explained by the EDR3 parallax biases, this must be due to inaccuracies in the HST parallaxes. Since these HST parallaxes are still used to anchor SH0ES measurements, this inaccuracy may contribute to the Hubble tension.

The future of the Hubble tension lies in the continuing mission to gather more data and observe more objects. Particularly important are the observations of Cepheids in SNIA hosts which provide the crucial step of calibrating the absolute peak SNIA magnitude, the key measurement when determining the Hubble constant. With the growing set of measurements there

is also the opportunity to calibrate H_0 using standard candles other than Cepheids, namely the TRGB. With powerful telescopes such as the JWST and the Vera Rubin Observatory coming online in the next few years, the new data could have the potential to change our view of the universe, confirming Hubble tension solutions such as the Local Hole theory or Early Dark Energy, among others.

ACKNOWLEDGMENTS

I would like to thank my supervisors Tom Shanks and Nigel Metcalfe, both of whom have given me invaluable advice both with project work and with my application for a PhD. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (<https://www.cosmos.esa.int/gaia>), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, <https://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.

REFERENCES

- Andrae R., et al., 2018, *A&A*, 616, A8
 Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, *Annu. Rev. Astron. Astrophys.*, 47, 481
 Benedict G. F., et al., 2007, *AJ*, 133, 1810
 Burns C. R., et al., 2018, *ApJ*, 869, 56
 Cantat-Gaudin T., et al., 2018, *A&A*, 618, A93
 Cantat-Gaudin T., et al., 2020, *A&A*, 640, A1
 De Jaeger T., Galbany L., Riess A. G., Stahl B. E., Shappee B. J., Filippenko A. V., Zheng W., 2022, arXiv:2203.08974 [astro-ph]
 Dhawan S., et al., 2022, arXiv:2203.04241 [astro-ph]
 Efstathiou G., 2020, arXiv:2007.10716 [astro-ph]
 Fitzpatrick E. L., Massa D., 2007, *ApJ*, 663, 320
 Flynn C., Sekhri R., Venville T., Dixon M., Duffy A., Mould J., Taylor E. N., 2022, *MNRAS*, 509, 4276
 Follin B., Knox L., 2018, *MNRAS*, 477, 4534
 Freedman W. L., 2021, *ApJ*, 919, 16
 Groenewegen M. a. T., 2021, *A&A*, 654, A20
 Holl B., et al., 2018, *A&A*, 618, A30
 Hoyle F., Shanks T., Tanvir N. R., 2003, *MNRAS*, 345, 269
 Hubble E., 1929, *PNAS*, 15, 168
 Kharchenko N. V., Piskunov A. E., Schilbach E., Röser S., Scholz R.-D., 2013, *A&A*, 558, A53
 Krisciunas K., et al., 2017, *AJ*, 154, 211
 Leavitt H. S., Pickering E. C., 1912, Harvard College Observatory Circular, 173, 1
 Lemaître G., 1927, *Ann. Soc. Sci. Bruxelles, Ser. B*, 47, 49
 Lindegren L., et al., 2021a, *A&A*, 649, A2
 Lindegren L., et al., 2021b, *A&A*, 649, A4

- Luri X., et al., 2018, [A&A](#), 616, A9
- Madore B. F., Freedman W. L., 1991, [Publ. Astron. Soc. Pac.](#), 103, 933
- Maíz Apellániz J., 2022, [A&A](#), 657, A130
- Malhan K., Ibata R. A., Martin N. F., 2018, [MNRAS](#), 481, 3442
- Mortsell E., Goobar A., Johansson J., Dhawan S., 2021a, arXiv:2105.11461 [astro-ph, physics:gr-qc]
- Mortsell E., Goobar A., Johansson J., Dhawan S., 2021b, arXiv:2106.09400 [astro-ph, physics:gr-qc]
- Ngeow C.-C., Kanbur S. M., Nikolaev S., Buonaccorsi J., Cook K. H., Welch D. L., 2005, [MNRAS](#), 363, 831
- Obermeier C., et al., 2016, [A&A](#), 587, A49
- Owens K. A., Freedman W. L., Madore B. F., Lee A. J., 2022, arXiv:2201.00733 [astro-ph]
- Perivolaropoulos L., 2022, arXiv:2201.08997 [astro-ph, physics:gr-qc, physics:hep-ph, physics:hep-th]
- Pietrzyński G., et al., 2019, [Nature](#), 567, 200
- Planck Collaboration et al., 2020, [A&A](#), 641, A6
- Reid M. J., Pesce D. W., Riess A. G., 2019, [ApJ](#), 886, L27
- Riess A. G., et al., 2009, [ApJ](#), 699, 539
- Riess A. G., et al., 2016, [ApJ](#), 826, 56
- Riess A. G., et al., 2018a, [ApJ](#), 855, 136
- Riess A. G., et al., 2018b, [ApJ](#), 861, 126
- Riess A. G., Casertano S., Yuan W., Macri L. M., Scolnic D., 2019, [ApJ](#), 876, 85
- Riess A. G., Casertano S., Yuan W., Bowers J. B., Macri L., Zinn J. C., Scolnic D., 2021, [ApJ](#), 908, L6
- Riess A. G., et al., 2022, arXiv:2112.04510 [astro-ph]
- Ripepi V., et al., 2021, [MNRAS](#), 508, 4047
- Sakai S., 1999, in Sato K., ed., Vol. 183, Cosmological Parameters and the Evolution of the Universe. p. 48
- Sanders J. L., Belokurov V., Man K. T. F., 2021, [MNRAS](#), 506, 4321
- Schlafly E. F., et al., 2016, [ApJ](#), 821, 78
- Schlafly E. F., Peek J. E. G., Finkbeiner D. P., Green G. M., 2017, [ApJ](#), 838, 36
- Scolnic D. M., et al., 2018, [ApJ](#), 859, 101
- Scolnic D., et al., 2022, arXiv:2112.03863 [astro-ph]
- Shanks T., Hogarth L. M., Metcalfe N., Whitbourn J., 2019, [MNRAS](#), 490, 4715
- Smith T. L., Lucca M., Poulin V., Abellan G. F., Balkenhol L., Benabed K., Galli S., Murgia R., 2022, arXiv:2202.09379 [astro-ph, physics:hep-ph]
- Tisserand P., et al., 2022, arXiv:2112.07693 [astro-ph]
- Vilardell F., Ribas I., Jordi C., Fitzpatrick E. L., Guinan E. F., 2010, [A&A](#), 509, A70
- Wang C., Yuan H., Huang Y., 2022, [ApJ](#), 163, 149
- Wong J. H. W., Shanks T., Metcalfe N., Whitbourn J. R., 2022, [MNRAS](#), 511, 5742
- Zinn J. C., 2021, [ApJ](#), 161, 214

Appendix A: Anchor Combinations

See table on next page.

Anchors	$M_{H,1}^W$ (mag)	σ	b_W (mag)	σ	Z_W (mag dex $^{-1}$)	σ	M_B	σ
<i>Four anchors</i>								
MW (raw) + M31 + LMC + N4258	-5.875	0.005	-3.24	0.02	-0.09	0.02	-19.20	0.03
MW (L21b) + M31 + LMC + N4258	-5.854	0.005	-3.24	0.02	-0.04	0.02	-19.18	0.03
MW (L21b + zp) + M31 + LMC + N4258	-5.849	0.005	-3.24	0.02	-0.03	0.02	-19.17	0.03
<i>Three Anchors</i>								
MW (L21b + zp) + LMC + N4258	-5.875	0.006	-3.24	0.02	-0.13	0.03	-19.21	0.03
MW (L21b + zp) + M31 + N4258	-5.857	0.006	-3.23	0.02	-0.02	0.05	-19.17	0.03
MW (L21b + zp) + M31 + LMC	-5.848	0.005	-3.22	0.02	-0.02	0.02	-19.16	0.03
<i>One Anchor</i>								
MW (raw)	-5.92	0.01	-3.20	0.03	-0.23	0.06	-19.25	0.04
MW (L21b)	-5.85	0.01	-3.14	0.03	-0.19	0.06	-19.14	0.04
MW (L21b + zp)	-5.84	0.01	-3.13	0.03	-0.20	0.06	-19.11	0.04
M31	-5.77	0.02	-3.21	0.03	-0.19	0.08	-19.10	0.04
LMC	-5.90	0.02	-3.22	0.02	-0.19	0.08	-19.23	0.05
N4258	-6.03	0.03	-3.08	0.04	-0.18	0.07	-19.27	0.04

TABLE III: Fit parameters for the first and second rungs of the cosmic distance ladder. The preferred combination of anchors is highlighted in bold.

Appendix B: Cosmic Distance Ladder

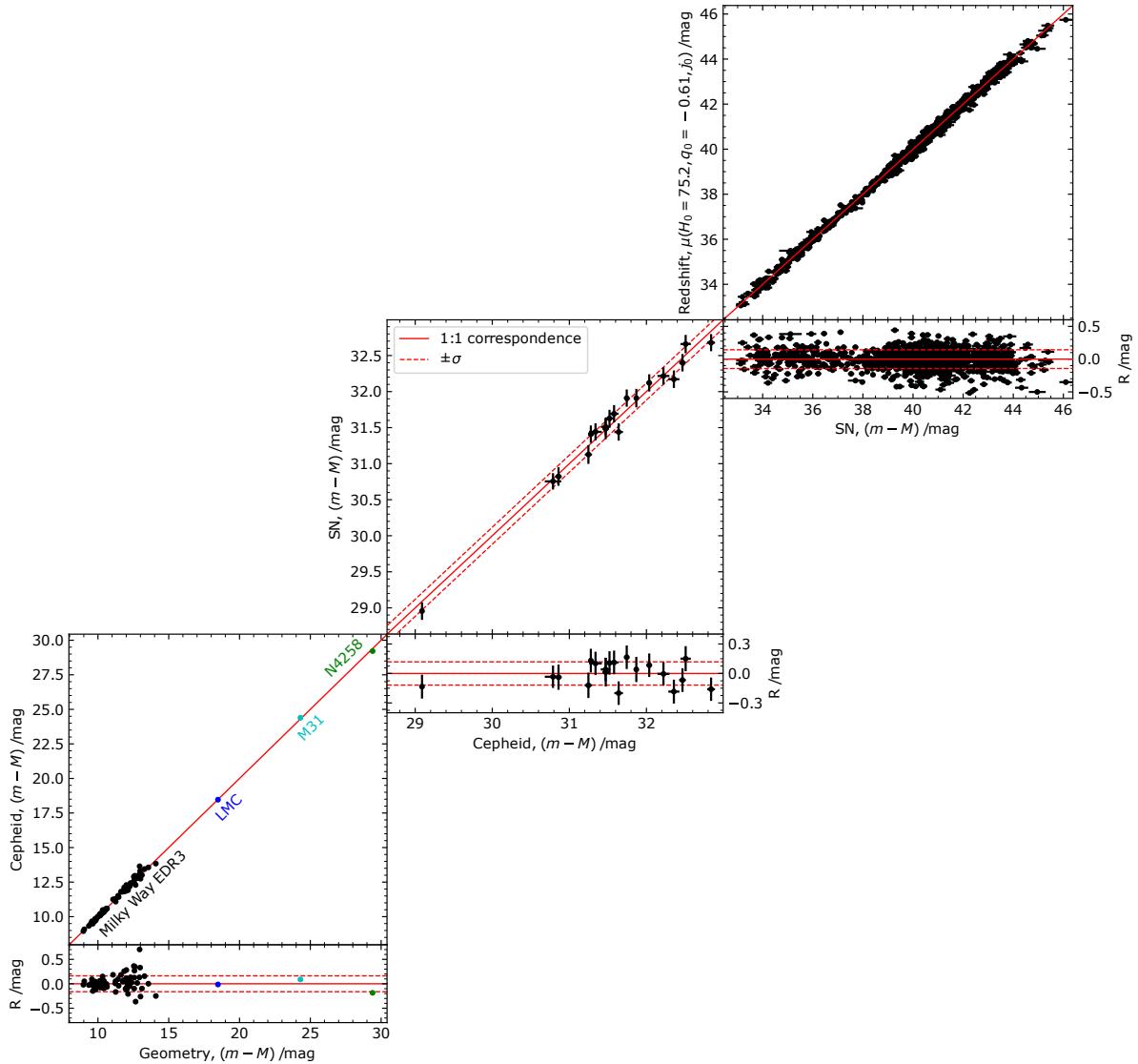


FIG. 16: Plot of all three rungs of the cosmic distance ladder for the anchor combination, Milky Way with L21b and zp correction, LMC and N4258.

Appendix C: Main Sequence Fits

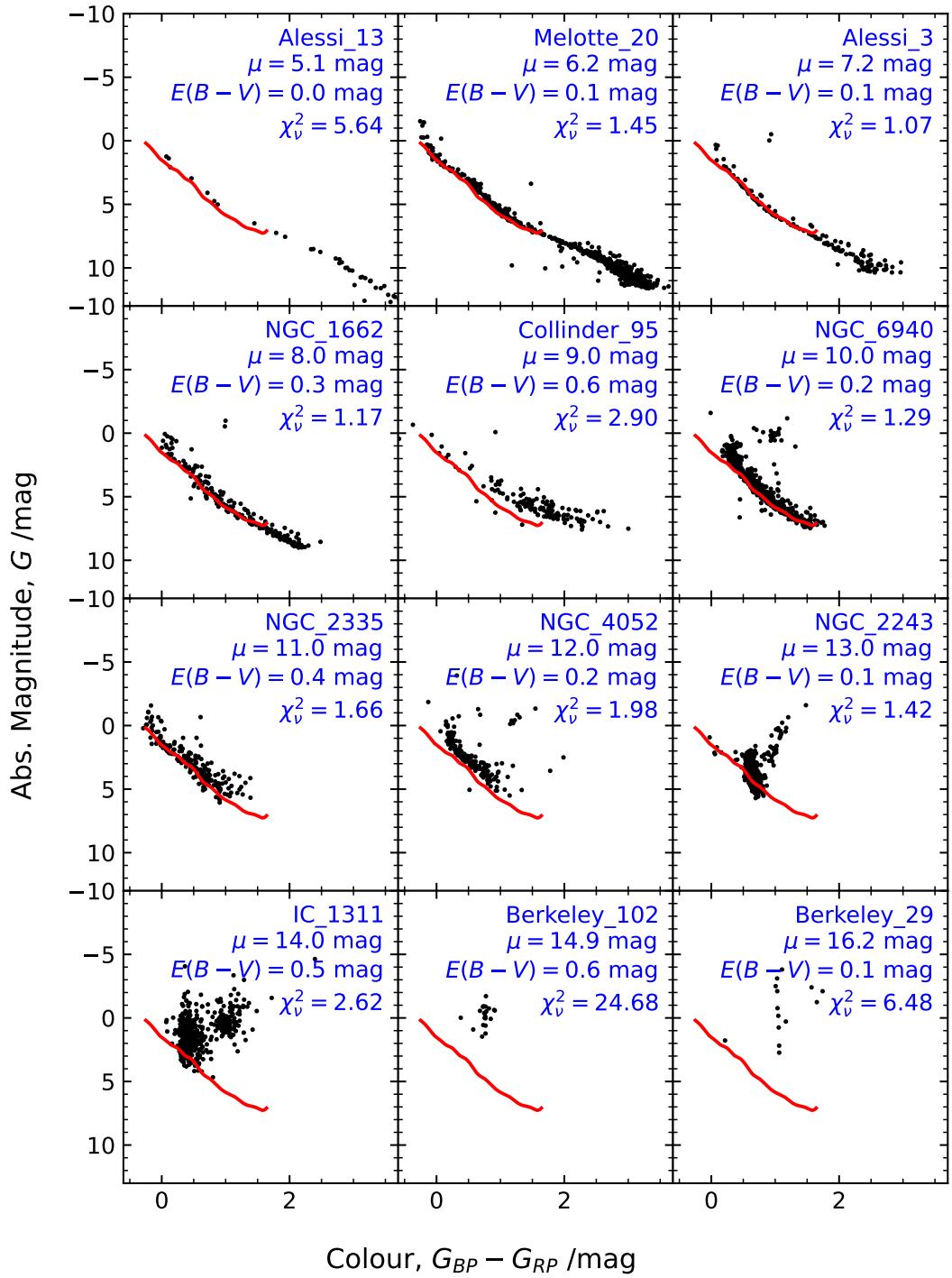


FIG. 17: Plots of the main-sequence fits using the standard reddening law for a selection of clusters at different distances. Note the shape of the CMD of the cluster IC 1311, an example of a cluster which is filtered out by fitting arbitrary shifts to the M25 model.

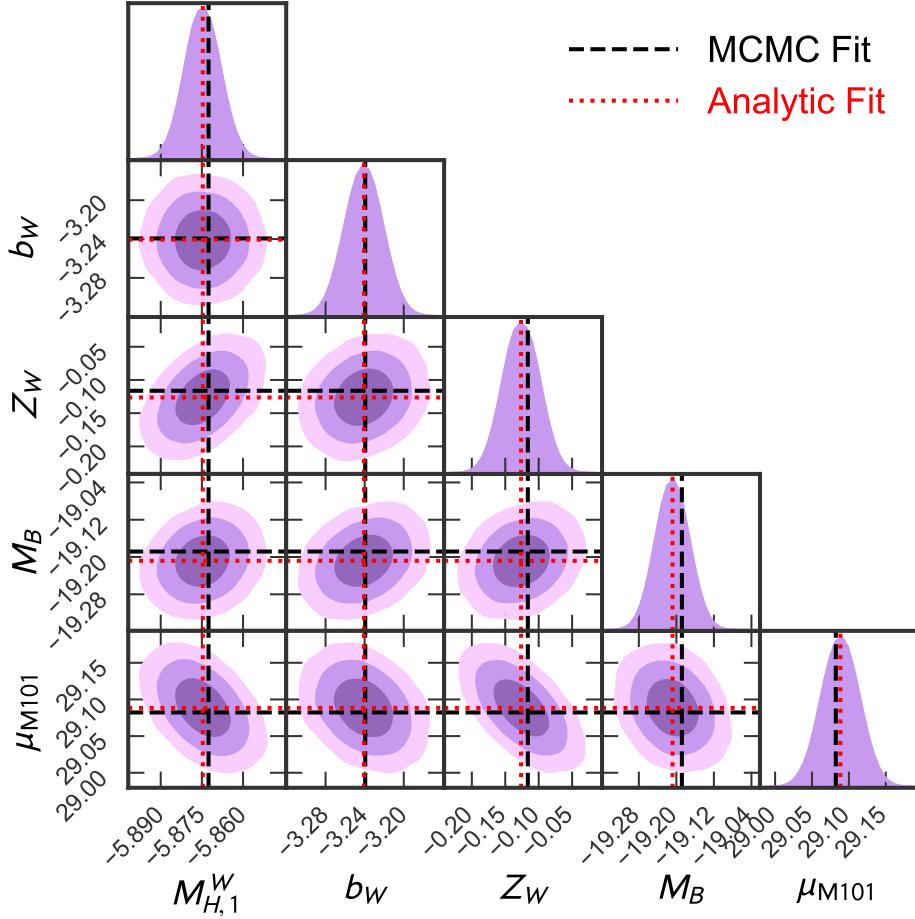


FIG. 18: Corner plot from fitting the distance ladder using MCMC. The best fit values of the parameters with the greatest likelihood are plotted as black lines and the analytic fit parameters are plotted in red. The contours represent the 1σ , 2σ and 3σ percentiles. The distance modulus of M101 is included due to the apparent covariance between μ_{M101} and parameters such as $M_{H,1}^W$ and Z_W .

Appendix D: MCMC Method

A numerical alternative to the analytical method outlined in section 2.2.3 is to use Markov-Chain Monte Carlo sampling. The advantage of this method over the analytical solution is its ability to demonstrate covariances between the fit parameters. Here, the posterior is

$$P(\mathbf{q}|\mathbf{L}, \mathbf{C}, \mathbf{y}) \propto P(\mathbf{q})P(\mathbf{L}, \mathbf{C}, \mathbf{y}|\mathbf{q}), \quad (\text{D1})$$

where \mathbf{q} , \mathbf{y} , \mathbf{L} and \mathbf{C} have the same definitions as in section 2.2.3. A simple χ^2 likelihood is used with

$$P(\mathbf{L}, \mathbf{C}, \mathbf{y}|\mathbf{q}) = \exp \left[-\frac{1}{2} \sum_i \left(\frac{y_i - (\mathbf{L}\mathbf{q})_i}{\sigma_i} \right)^2 \right], \quad (\text{D2})$$

where σ_i is the diagonal of \mathbf{C} such that $\mathbf{C} = \sigma_i \delta_{ii}$. A uniform prior is adopted of the form

$$P(\mathbf{q}) = \prod_i \begin{cases} 1 & \text{if } q'_i - 10\sigma_{q,i} \leq q_i \leq q'_i + 10\sigma_{q,i} \\ 0 & \text{otherwise} \end{cases}, \quad (\text{D3})$$

where $\mathbf{q}' \pm \sigma_{\mathbf{q}}$ are the best-fit parameters from the analytical method. To perform the fit, the python package `emcee` is used with 500 walkers and 1024 iterations.

The histograms of the samples can be seen in figure 18, using the best-fit combination of anchors MW (L21b + zp), LMC and N4258. The best-fit parameters are found as the maximum value of the likelihood function, $P(L, C, y|\mathbf{q})$. The distance modulus of M101 is included alongside the PL and SNIA parameters due to the apparent covariance between μ_{M101} and PL parameters such as $M_{H,1}^W$ and Z_W . As all fit parameters agree with the results of the analytical fit, this method is included in the appendix as a supplemental methodology.

The full corner plot with all 23 parameters is available online³.

³ <https://i.imgur.com/aToPhER.png>