Distances to Globular Clusters using RR Lyrae Variables

Luke Bischoff

Department of Physics, University of Bath, Bath BA2 7AY, United Kingdom

E-mail: `ltb24@bath.ac.uk`

Abstract. RR Lyrae variable stars resident in two globular clusters in the Milky Way galaxy, Palomar 5 and Palomar 13, are analysed and the distances to the globular clusters determined. Point spread function photometry was performed on a series of observations taken from the Infrared Array Camera on board the *Spitzer Space Telescope* in the mid-infrared bands, 3.6 μm (channel 1) and 4.5 μm (channel 2), to measure the apparent magnitudes of RR Lyrae stars detected with in the globular clusters. Apparent magnitudes were averaged using a Gaussian local estimation algorithm, which were then corrected for dust extinction effects. Period-luminosity relations were fit to the data for both clusters, from which distances in both the 3.6 μm and 4.5 μm bands were obtained. The average distance to the globular clusters were determined to be $d_{\mathrm{Pal5}} = 20.37 \pm 0.51$ (random) ± 1.60 (systematic) kpc, $d_{\mathrm{Pal13}} = 24.25 \pm 0.62$ (random) ± 1.90 (systematic) kpc which were found to be consistent with previous determinations.

1. Introduction

1.1. The SMHASH program and Spitzer

The measurement of distances to objects throughout the Universe is notoriously difficult, due to the absence of direct three-dimensional information available, and remains an active area for research in astrophysics. Observations made using the *Spitzer Space Telescope* as part of the *Spitzer* Merger History and Shape of the Galactic Halo (SMHASH), alongside the Carnegie RR Lyrae Program (CRRP) can be used to measure distances in the local Universe to a higher accuracy than before [1]. The SMHASH program primarily aims to construct a three-dimensional map of the Milky Way (MW) galaxy [2] by using mid-infrared observations of RR Lyrae (RRL) variable stars in globular clusters and dwarf spheroidal galaxies (dSphs) of the MW. The CRRP, which is complementary to the SMHASH program, studies RRLs in the MW [3] and so, together, these programs are contingent on accurate determinations of distances.

In this paper, distances to the Palomar 5 (Pal 5) and Palomar 13 (Pal 13) globular clusters are calculated using the Infrared Array Camera (IRAC) on board the *Spitzer Space Telescope* as part of the Warm Mission during 2013, using its 3.6 μm and 4.5 μm channels; after *Spitzer's* cryogenic coolant run out, thus ending the use of its two longer wavebands.

1.2. RR Lyrae variables

RR Lyrae stars are Population II variable stars, meaning they are relatively old and metal-poor stars that vary with regular periods of around a few hours to a day [4]. RRLs are frequently found in large quantities within globular clusters which reside in the galactic halo of the MW galaxy. Starting life on the main sequence with masses typically less than that of the Sun, RRLs have evolved past their red-giant stage to lie on the instability strip of the Hertzsprung-Russell diagram [5]. There are three main types of RRL: (i) RRab types, which pulsate in the fundamental frequency, (ii) RRc types,

which pulsate in the first overtone; typically with smaller amplitudes but in a more sinusoidal fashion than RRab types, and, (iii) RRd types, which pulsate simultaneously in the fundamental frequency and the first overtone [1]. Due to their regular variation in brightness, RRL stars are used as standard candles, which are markers used to calibrate distances to cosmological objects [2].

Previous work has typically focused on Cepheid variable stars and the Cepheid period-luminosity (PL) relation, known as the Leavitt law, after being discovered by Henrietta Swan Leavitt in 1908. A PL relation is the relationship between a stars brightness and its period of pulsation. Cepheids, like RRLs, are used as standard candles but are brighter and pulsate with a longer, more stable period of around weeks, when compared to RRLs [4]. The PL relations of Cepheids are well defined in optical wavebands, however this leaves them more susceptible to dust extinction effects [6]. Therefore, RRL PL relations in the mid-infrared bands have several advantages over Cepheids, such as (i) weaker effects due to dust extinction, (ii) metallicity effects are damped in these wavebands, (iii) more sinusoidal and symmetrical light curves, with smaller amplitudes improve precision of average apparent magnitudes, and (iv) a narrower dispersion (or scatter) in the PL relations [2, 4].

The cosmic distance ladder details a process of using cosmological objects, such as Cepheids and RRLs, to calibrate distances from the galactic neighbourhood of around 10 kpc up to 100 Mpc, with an ultimate aim to accurately measure the Hubble constant, H_0. Initially Cepheids and the Leavitt law have been used to climb the distance ladder, however results for the Hubble constant remain inconsistent depending on the approach [4]. Determinations of the Hubble constant using Cepheids have yielded values of $H_0 = 74\pm3$ kms^{-1}Mpc^{-1} [7], where as determinations using the cosmic microwave background have obtained values of $H_0 = 67.8 \pm 0.9$ kms^{-1}Mpc^{-1} [8]. This inconsistency could be resolved by using RRLs as a separate route up the cosmic distance ladder which could improve the precision of the Hubble constant to less than 3% [4]. An agreement for a precise value of the Hubble constant would enable an appreciation of the age, size and expansion of the Universe [9] highlighting the importance of obtaining accurate measurements for these first rungs of the cosmic distance ladder.

2. Data selection and photometry

2.1. Data selection

The IRAC on board *Spitzer* observed a region of around 6×15 arcminutes around the centres of the Pal 5 and Pal 13 globular clusters. Pal 5 [10] and Pal 13 [11] both exhibit long tidal streams that extend beyond their centres due to interaction with, and accretion to, the MW galaxy, and so RRLs residing within these tidal streams are not observed. Each cluster was observed in both the 3.6 μm (IRAC channel 1) and 4.5 μm (IRAC channel 2) wavebands and over a duration of 12 epochs, with one frame in each channel per epoch, calibrated such that at least one full cycle for each RRL in the frame was observed. Prior to use, the raw data was pre-processed into basic calibrated data (BCD) frames by *Spitzer* Science Centre using an established pipeline to reduce the images into a useable state. Image reduction involves, for example, flat fielding of the images, which corrects for variations in pixel sensitivity across the detector, and the removal of biases [12]. Mosaic frames, which is a stitching together of individual BCDs, for each epoch in both IRAC channels were created by the supervisor for this project, Dr Victoria Scowcroft, resulting in the final data in the form of Flexible Image Transport System (`FITS`) files. A pixel phase effect can be exhibited on individual BCDs due to differences in flux, depending on the location within a pixel the light from a point sources falls because of quantum efficiency variations [13], however because the combination of the BCDs into mosaics smooths out this pixel phase effect, correction for it was neglected.

The positions in right ascension (RA) and declination (Dec), periods in days and the type of previously known RRLs were obtained from the *Catalogue of Variable Stars in Galactic Globular Clusters* (CVS), maintained by Christine Clement [14]. Further information was gathered from *Gaia* Data Release 2 (*Gaia* DR2) [15] by considering a region of 10' in radius around the centre of the globular clusters and searching for stars marked as `variable`. The variable stars identified were checked against the RRLs in the CVS to identify candidate RRLs missing from the CVS. Confirmation

of whether the candidate variable stars from the region search in *Gaia* was obtained by matching the *Gaia* `source_id` of the candidate stars to the `gaiadr2.vari_rrlyrae` list containing information on 140784 RRLs; including their periods and types. If the period of the RRL was available from *Gaia*, then this was used over the periods provided in the CVS because some periods in the CVS were obtained in 1962 [16] and RRL periods are not necessarily stable over human lifetimes.

Master mosaic images were also created for each IRAC channel to create `FITS` files of a median image using data from across all 12 epochs. By creating a median image, the master mosaic images have a higher signal-to-noise ratio, making them particularly useful as a comparison to the individual epoch data. The data for the SMHASH program and CRRP were optimized for IRAC channel 1 in the 3.6 μm band and so the channel 2 data, taken in the 4.5 μm band, has a lower signal-to-noise ratio which presented some issues, particularly regarding the individual epoch data, which is explained more in section 2.4 [2].

The individual `FITS` files, as well as the image data as flux values in MJysr^{-1}, contained header information detailing the exposure time, observation date and time and a flux conversion factor for both the epoch files and the master mosaic files. This information was used to convert the flux values from MJysr^{-1} into data number counts, DN, using Equation 1,

$$\text{counts } = \text{flux } \cdot \frac{\text{exposure time}}{\text{flux conversion factor}} \ . \tag{1}$$

where flux is in MJysr^{-1}, exposure time is in seconds and the flux conversion factor in MJysr$^{-1} \cdot$ DN$^{-1} \cdot$ s.

2.2. Photometry

Photometry involves measuring the amount of light, usually in flux or data number counts, from the image produced by a detector or camera, such as IRAC [17]. Once the epoch and master mosaic data was converted into DN counts, photometry was performed on the data to obtain the apparent magnitudes of point sources detected in the frames. There are two main methods for performing photometry that were considered during this project, aperture photometry and point spread function (PSF) photometry. Aperture photometry involves considering the flux of light measured through an aperture centred on the star of radius just larger than the star. A second aperture of larger radius than the aforementioned aperture is also drawn and centred to form an annulus around the star. The average background flux measured in the annulus region is then subtracted from the inner aperture to obtain the flux of the star. There is a problem of optimization here, particularly regarding the radii of the inner and outer apertures, as such it is necessary to ensure that an aperture is chosen that is not too large to include contamination by nearby sources of light but also not too small such that light from the star is being omitted and counted as background flux.

The second method, PSF photometry, involves building a model based on the images of point sources that are detected and typical of a particular detector and fitting the model to point sources in the image frames in order to measure the flux of each point source [17]. The process is usually iterative in that the stars that have been fit and flux values measured, are subtracted from the image to create a residual image. The process of fitting the model to the residual image is then completed which may reveal stars that were not detected in the previous iteration. This iterative process applied by PSF photometry is particularly useful for crowded fields, and because globular clusters are crowded in nature, PSF photometry was chosen over aperture photometry.

2.2.1. PSF model.
A point spread function can be thought of as the image profile produced by a particular instrument, such as IRAC, when a point source is incident on it. The idea of an *effective* PSF (ePSF) model is introduced in Anderson and King [18], which is a net PSF describing the amount of light from a point source falling on an individual pixel. Separate ePSF models were built from the master mosaic frames for each IRAC channel because the model is dependent on the sample of stars present in each frame. This ensures that an accurate ePSF model is produced to reduce the number of false positive and false negative detections of point sources, when the model is eventually fit to the

series of epoch frames. Individual epoch frames were at first used to develop ePSF models to be used in the photometry of that corresponding epoch, however, some epoch frames had particularly low signal-to-noise ratios; especially in channel 2. This meant that inaccurate or failed ePSF models were built, resulting in defects in the PSF photometry process causing inaccurate apparent magnitudes. Therefore, the master frames were used to develop ePSF models, which, as a result of their higher signal-to-noise ratio, improved the quality of the model.

Stars used to construct the ePSF models were detected using a point source detection algorithm, `DAOStarFinder`, which implements the `DAOFIND` routine of the greater `DAOPHOT` procedure, developed by Peter Stetson [19]. A version of the `DAOPHOT` procedure is usefully included in the Python package, `photutils` [20], which was used extensively throughout this project. Bright stars, above a threshold of 50σ from the background level, were chosen for the ePSF model because these stars are more well-defined, with a higher signal-to-noise ratio, to ensure an accurate ePSF model was derived for each IRAC channel. Stars residing near the edge of the master mosaic frames were discounted from the detections and each candidate star to be used in the ePSF model was presented in a list of images and manually inspected for defects. Candidate stars exhibiting defects such as a blending effect, which could be two or more unresolved stars or a galaxy; crowding from neighbouring stars or stars containing artefacts such as bad pixels or cosmic rays. Detection parameters were utilized, which were implemented as part of the source detection algorithm (see section 2.2.2), which greatly reduced detections of candidate stars containing these artefacts, however, the candidate stars were still screened for these defects as some artefacts were still detected. Once these poor quality candidate stars were removed from the list, the ePSF model was rebuilt and saved ready for use in the photometry routine.

2.2.2. PSF photometry. A master source list containing all detected point sources was generated from the master mosaic frame for each IRAC channel in both globular clusters. The star detection algorithm was run on the frames as part of the PSF routine based on each channel's respective ePSF model to populate the master source list, to which it can be referred back to when performing photometry on the series of epoch frames.

Several parameters for the source detection algorithm were used to reduce the number of false positive star detections, these are the full width at half maximum (FWHM) of the stars, a threshold detection value and two parameters relating to the shape and sharpness of the point sources. The optimisation problem must be reconsidered when deciding on the specific values to use in order to select values which are sufficient enough to ensure actual stars are detected but not to such a degree that, for example, random noise is being detected as a star or that the process becomes computationally time expensive. The FWHM parameter was determined based upon analysis of the radial profiles of many uncrowded stars in the frames. The radial profiles of these stars were manually examined using `ImExam`, an affiliated package of the Python package, `astropy`, as in table 1 of The Astropy Collaboration *et al* [21]. After gaining an appreciation for the range of FWHM values, a value of 5.0 pixels was deemed acceptable.

The second parameter considered is the threshold value, measured in standard deviations, σ, above the median background level. Only point sources with peak count values above the determined threshold value will be detected as a star. A plot of the number of sources detected against threshold value was generated with a range of threshold values from 1σ to 10σ, producing an "elbowed" curve, where the elbow of the curve was situated around 5σ to 6σ. A value of 50σ was used in building the ePSF models, but this was a special case as explained in section 2.2.1. Instead, for the source detection algorithm on the master and epoch frames, a value of 6σ was decided for the threshold value after considering the location of the elbow on the x-axis of aforementioned plot.

Another parameter that was analysed was the `round` parameter, which considers the profiles of the point sources and aims to discard sources with directional biases, which could indicate that the source is a galaxy, for example, instead of a star. After visual examination of detected sources, values of ± 0.5 were used. The `sharp` parameter is the final parameter examined here, which considers the sharpness of the profile of the point sources and so setting an upper limit on the sharpness aims to

avoid the source detection algorithm detecting bad pixels or cosmic rays as stars. This value, also determined by visual examination, was set to an upper limit of 0.8 for the construction of the ePSF models, because the most accurate model was desired; the value was relaxed to 0.9 for detections in the master and epoch frames.

Photometry was then performed using an iteratively subtracted PSF routine, which follows the process outlined in `DAOPHOT` [19], whereby the source detection algorithm is run in accordance with the parameters described previously, to measure the flux, in DN counts, of the sources in the epoch frames. The ePSF model constructed for each channel is fit to the epoch data using the Levenberg-Marquardt algorithm and least squares statistic built into `astropy` [21]. A fixed centroids process, which reduces the uncertainty in the determinations of flux [2], was undertaken by using the initial positions of sources in the master source list, to improve the precision of the detected sources within the PSF photometry routine. A residual image was obtained after the fitted sources were subtracted from the image, ready for the next iteration of the process. Pal 5 and Pal 13 are relatively uncrowded compared to other globular clusters and so only two iterations was needed. Once the iterative process was completed, values for the magnitudes of detected stars in DN were obtained for that epoch.

The process was slightly modified for IRAC channel 2 of Pal 13, such that the PSF photometry only considered regions of $50{\times}50$ pixels around known RRLs using the CVS and *Gaia* data due to an extremely bright source in the frame which resulted in no values for magnitudes of sources in most epochs being obtained. The rest of the analysis was the same as in the previous paragraph, only the initial sources stage was changed.

The coordinates of stars detected in the epoch frames were converted from their x and y values into RA and Dec using the given World Coordinate System (WCS) [21] in the header of the `FITS` files and transformed using the International Celestial Reference System (ICRS). This was necessary because the x and y coordinates of the stars in each epoch frame are unique to that frame only and so coordinates common to all epoch frames (including the master frame) were required to enable identification of the same sources in each frame. The PSF photometry routine was completed on all 12 epochs in series, for both channels, each time matching the stars detected in the epoch to the master frame using the WCS transformations. A final list of sources containing their positions in RA and Dec, their magnitudes in DN and uncertainties in DN for each epoch was obtained.

2.3. Determination of apparent magnitudes

The apparent magnitude, m, of each star was calculated using Equation 2,

$$m = ZP - 2.5\log_{10}\left(\text{counts} \cdot \frac{\text{flux conv} \cdot \text{aperture correction} \cdot \text{loc correction}}{\text{exposure time}}\right), \qquad (2)$$

where ZP is the zero-point magnitude for each IRAC channel, which is 18.80 for channel 1 and 18.32 for channel 2 as per the IRAC Handbook [13], *counts* is the flux in DN obtained from the photometry, the flux conversion factor and exposure time used to convert *counts* back into flux values. An aperture correction has also been applied to account for the difference between the apertures used to calibrate IRAC and the aperture of 6 pixels used to perform the photometry. The *loc correction* in Equation 2 is an array-location dependent correction which is a `FITS` file that corrects for scattering and spatial distortion caused by IRAC [13]. The uncertainty in magnitudes was also calculated during the PSF photometry routine and was propagated through into magnitudes. The uncertainties in apparent magnitude primarily arise due to the fitting process of the ePSF model, for example, it's finite accuracy and errors in the fit caused by blending or contamination from nearby stars [19].

2.4. Light curves

Light curves are plotted to identify and examine how the brightness of the stars vary as a function of time or, as in Figure 1, phase. An assortment of light curves are shown in Figure 1 from both Pal 5 and Pal 13 to give some examples of light curves from the data analysed and show some variations between them.
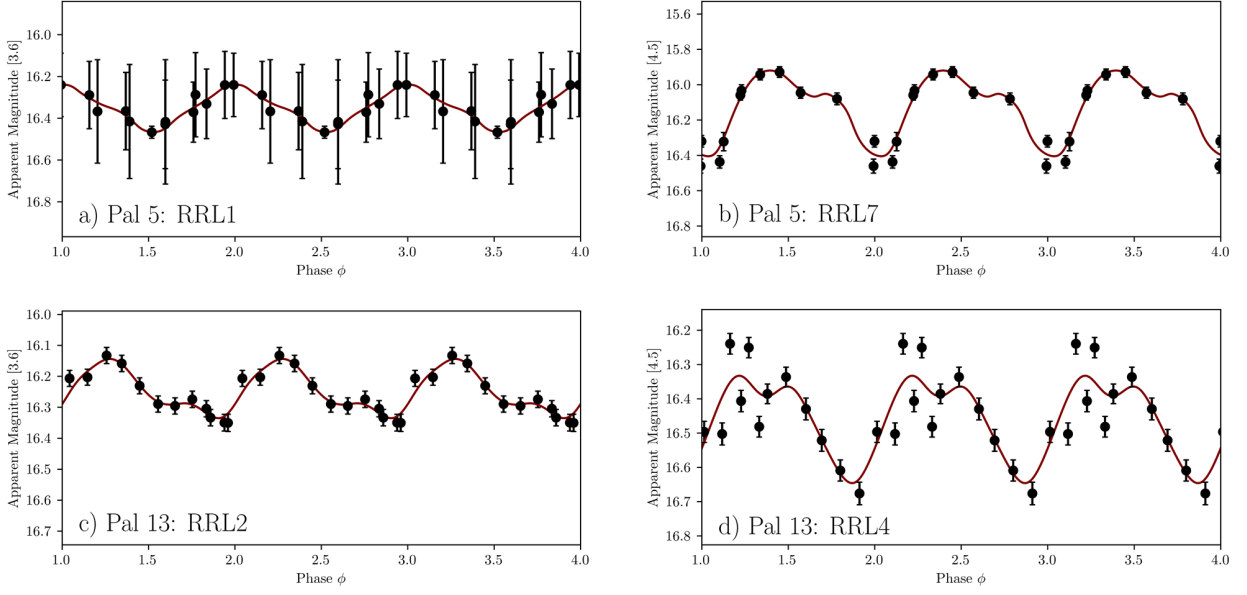
Figure 1. A summary of light curves from Pal 5 (top) and Pal 13 (bottom), displaying the range of curves obtained for the two globular clusters. a) the channel 1 light curve for RRL1, an RRc star in Pal 5 exhibiting large uncertainties in apparent magnitude. b) the channel 2 light curve of RRL7, an RRab star in Pal 5. c) the channel 1 light curve for RR2, an RRab star of Pal 13 displaying a characteristic curve of RRab type stars with minimal defects. d) the channel 2 light curve for RRL4, an RRab star in Pal 13 exhibiting scatter about the GLOESS fit.

The phase of the light curves was obtained by first extracting the date and time each epoch frame was taken from the header of the `FITS` file, and converting this into a Modified Julian Date (MJD). From there, the phase was calculated using Equation 3,

$$\phi = \frac{[MJD]}{P} - \left\lfloor \frac{[MJD]}{P} \right\rfloor , \tag{3}$$

where $[MJD]$ is the time in Modified Julian Date and P is the period of the RRL, similar to equation 10 in Monson *et al* [1]. The phase has been calculated so that multiple periods of the RRLs can be displayed in the light curve because only one period worth of raw data was collected, and so phase is a useful way to check whether the periods smoothly join together. A Gaussian local estimation algorithm (GLOESS) [22] has been used to fit a curve to the light curves and to calculate the time-averaged apparent magnitudes and corresponding uncertainties of the RRLs by parsing the algorithm the phase, apparent magnitude and magnitude uncertainties of each RRL. GLOESS uses second-order polynomials to fit the data locally via interpolation between subsequent data points on the x axis [23] and moves across the data set until the entire light curve has been processed. This time-averaged (mean) apparent magnitude for each RRL was extracted and placed into an array containing its corresponding uncertainty and period. The uncertainty in the mean apparent magnitude was calculated using Equation 4,

$$\sigma_{\langle m \rangle} = \sqrt{ \frac{\Sigma \sigma_i^2}{N^2} + \sigma_{\text{fit}}^2 + A^2 } , \tag{4}$$

where σ_i^2 is the uncertainty in each individual epoch's apparent magnitude for that RRL, N is the number of epochs present in the light curve, σ_{fit}^2 is the standard deviation of the points around the fit and A is the value of the dust extinction.

The average apparent magnitudes were corrected for extinction effects caused by dust in the MW galaxy between the observer and the globular cluster. This results in a reddening of the apparent magnitudes which would influence the measured distances to the globular clusters. The corrections

for dust extinction were obtained from the IRSA Dust Extinction Service Queries built into the `astroquery` module in Python [24], which uses a MW dust map developed in 2011 [25]. Extinction magnitudes, A, were subtracted from the mean apparent magnitudes for all of the RRLs detected. The extinction magnitudes were found to be less than 0.1% of the general sample's mean apparent magnitudes and with the absence of uncertainty values provided for the extinction corrections, the extinction value itself was taken as its uncertainty in Equation 4, which is consistent with, for example, Scowcroft *et al* [6].

3. Period-Luminosity relations and distance measurements

3.1. Fitting PL relations for Pal 5 and Pal 13

In order to calculate the distance to the globular clusters, a comparison between the apparent magnitude of the RRLs and their absolute magnitudes is required. This can be done by comparing a calibrated PL relation from the literature to a PL relation fit, by using the same mathematical form as the calibrated PL relation, to the sample data. The PL relations that have been fit to the data in Figures 2 and 3, for Pal 5 and Pal 13 respectively, use the PL relation data provided in table 3 of Neeley *et al.* [26] (henceforth N19), which is the most recent set of PL relations covering both IRAC channels 1 and 2.

The calibrated PL relations, for both apparent magnitude and absolute magnitude take the form of Equation 5,

$$M = a + b\left(\log_{10}(P) + 0.3\right), \tag{5}$$

where M is the apparent or absolute magnitude, a is the y-intercept of the PL fit, b is the gradient of the PL fit and P is the period of the RRL.

Prior to fitting the PL relations, RRLs identified as type RRc stars, were fundamentalized to enable comparisons between the RRc and RRab stars [27], using Equation 6,

$$\log_{10}(P) = \log_{10}(P_{\text{FO}}) + 0.127, \tag{6}$$

where P is the fundamentalized period and P_{FO} is the first overtone period. The exact form of the calibrated PL relations in absolute magnitudes are shown as Equation 7,

$$M_{[3.6]} = -0.40(\pm 0.03) - 2.78(\pm 0.38)(\log_{10}(P + 0.3)),$$
$$M_{[4.5]} = -0.41(\pm 0.03) - 2.83(\pm 0.39)(\log_{10}(P + 0.3)), \tag{7}$$

where the coefficients of a and b from Equation 5 are given in table 3 of N19. The PL relations in Figures 2 and 3 have been fit by fixing the gradient b constant, while allowing a to vary. Therefore, the PL relations shown in Figures 2 and 3, are of the form,

$$m_{[3.6]} = a_0 - 2.78(\pm 0.38)(\log_{10}(P + 0.3)),$$
$$m_{[4.5]} = a_0 - 2.83(\pm 0.39)(\log_{10}(P + 0.3)), \tag{8}$$

where m is the average apparent magnitude and a_0 is the y-intercept, yet to be found.

The intercepts were found using a least-squares fit to obtain intercept values in both channels for Pal 5 as $a_{0[3.6]} = 16.10 \pm 0.03$ mag and $a_{0[4.5]} = 16.16 \pm 0.03$ mag; and in both channels for Pal 13 as $a_{0[3.6]} = 16.47 \pm 0.02$ mag and $a_{0[4.5]} = 16.55 \pm 0.03$ mag. The dispersion lines in Figures 2 and 3 represent one standard deviation ($\pm 1\sigma$) of the residuals of the RRL data points above and below the fitted PL relations; with a value quoted for each PL relation.
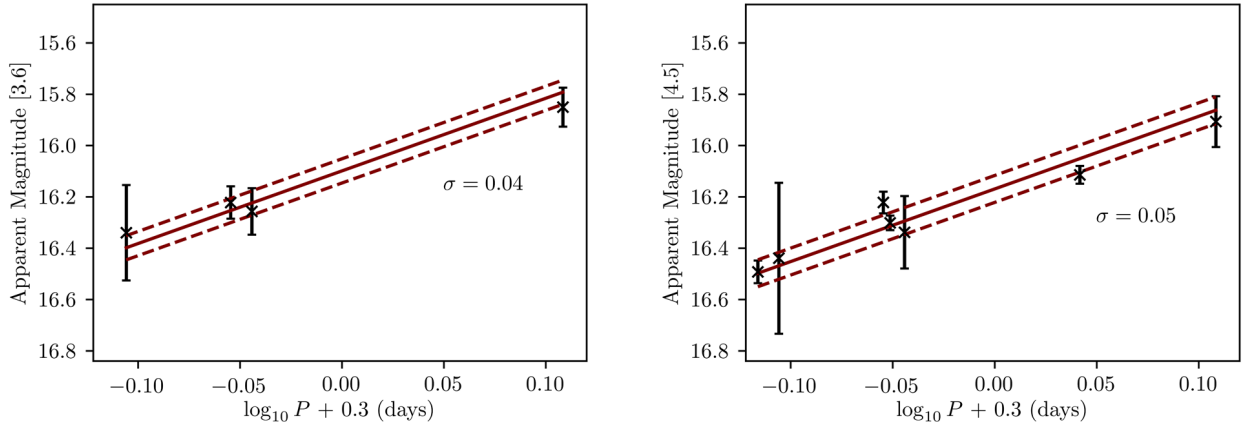
3.2. Calculating distances to Pal 5 and Pal 13

The distance to a globular cluster can be calculated from the distance modulus, which is the difference between the apparent magnitude and the absolute magnitude,

$$\mu = m - M, \tag{9}$$

Table 1. Summary of the results and uncertainties for the apparent magnitude m (as the intercept of the PL relation), distance modulus μ and distance d obtained for Pal 5 and Pal 13.

IRAC Channel	Pal 5			Pal 13		
	m (mag)	μ (mag)	d (kpc)	m (mag)	μ (mag)	d (kpc)
[3.6] $\pm\sigma_{\rm rand}(\pm\sigma_{\rm sys})$	16.10 ±0.03	16.51 $\pm0.04(\pm0.12)$	20.04 $\pm0.37(\pm1.11)$	16.47 ±0.01	16.88 $\pm0.03(\pm0.12)$	23.80 $\pm0.38(\pm1.32)$
[4.5] $\pm\sigma_{\rm rand}(\pm\sigma_{\rm sys})$	16.16 ±0.02	16.57 $\pm0.04(\pm0.12)$	20.69 $\pm0.35(\pm1.15)$	16.55 ±0.01	16.96 $\pm0.03(\pm0.12)$	24.69 $\pm0.49(\pm1.37)$

Figure 2. PL relations in channel 1 (left) and channel 2 (right) for Pal 5. The dashed lines indicate dispersion about the fitted PL, with the region between them being $\pm1\sigma$ of the residuals of the data points, as quoted for each PL relation. The periods of RRc and RRd stars have been fundamentalized in these PL relations. The y-axis scale has be set to enable comparisons between IRAC channels and with the PL relations in Figure 3.
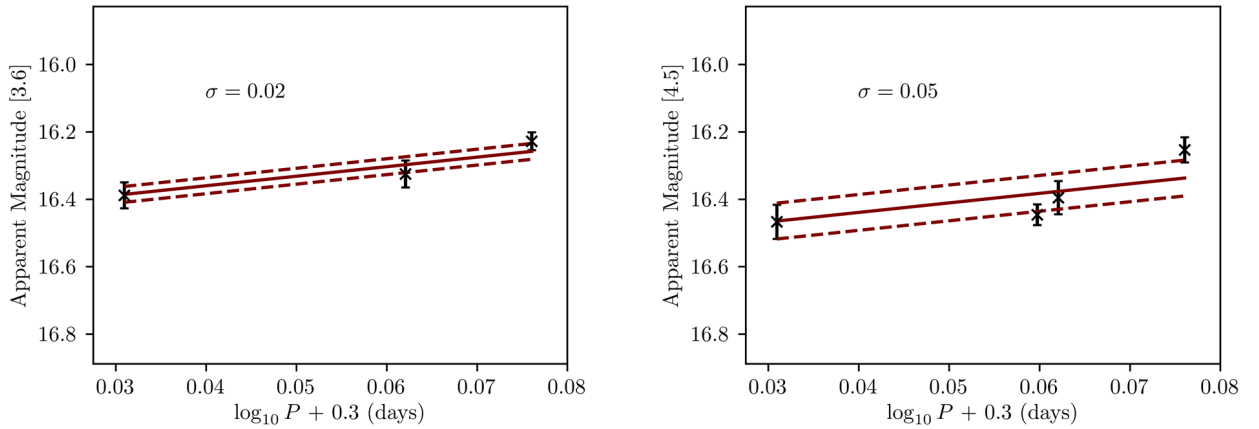
Figure 3. PL relations in channel 1 (left) and channel 2 (right) for Pal 13. The dashed lines indicate dispersion about the fitted PL, with the region between them being $\pm1\sigma$ of the residuals of the data points, as quoted for each PL relation. There were no RRc or RRd stars in this sample of RRLs and so all RRLs in these PL relations are of type RRab. The y-axis scale has been set to more easily compare dispersion to those in Figure 2.

where μ is the distance modulus, m is the apparent magnitude and M is the absolute magnitude. Consider the PL relations in Equations 7 and 8, the distance modulus for each IRAC channel can be found by subtracting Equation 7 from Equation 8, giving a general distance modulus,

$$\mu = m - M = a_0 - a, \tag{10}$$

where a_0 is the calculated intercept from fitting the PL relations and a is the intercept given in N19. The distance modulus for each channel can be obtained by substituting the relevant data for each IRAC channel into Equation 10. Therefore, the distance moduli values in both IRAC channels for Pal 5 are $\mu_{[3.6]} = 16.51 \pm 0.04 \, (\pm 0.12)$ mag and $\mu_{[4.5]} = 16.57 \pm 0.04 \, (\pm 0.12)$ mag and for Pal 13 are $\mu_{[3.6]} = 16.88 \pm 0.04 \, (\pm 0.12)$ mag and $\mu_{[4.5]} = 16.96 \pm 0.04 \, (\pm 0.12)$ mag. The first uncertainty quoted in the distance modulus results is the random uncertainty given by,

$$\sigma_{\text{rand}} = \sqrt{\sigma_a^2 + \sigma_{a_0}^2}, \tag{11}$$

where σ_a is the uncertainty in the calibrated intercept from N19 and in Equation 7 and σ_{a_0} is the uncertainty in the calculated intercept as a result of the fitting process. The uncertainty quoted in brackets is the systematic uncertainty, which is the uncertainty in the PL relation itself and given by,

$$\sigma_{\text{sys}} = \sqrt{\sigma_a^2 + \left(0.3 \cdot \sigma_b^2\right)}, \tag{12}$$

where σ_b is the uncertainty in the gradient of the PL relation as given in N19 and Equation 7.

The distance, d, in parsecs (pc) can be calculated from the distance modulus by,

$$\mu = 5 \cdot \log_{10}(d) - 5, \tag{13}$$

which can be rearranged for d, as shown in Equation 14,

$$d = 10^{\frac{\mu}{5}+1}. \tag{14}$$

Therefore, the determined distances in both IRAC channels to Pal 5 (in kpc) are $d_{[3.6]} = 20.04 \pm 0.37 \, (\pm 1.11)$ kpc and $d_{[4.5]} = 20.69 \pm 0.35 \, (\pm 1.15)$ kpc and for Pal 13 are $d_{[3.6]} = 23.80 \pm 0.38 \, (\pm 1.32)$ kpc and $d_{[4.5]} = 24.69 \pm 0.49 \, (\pm 1.37)$ kpc, where the uncertainties have been propagated through to kpc. A summary of these results are presented in Table 1.

4. Discussion

4.1. Comparing results to previous literature values

The distances to Pal 5 in both IRAC channels in Table 1 are consistent with each other when the random and systematic uncertainties are considered and so a mean distance to Pal 5 is determined to be $d_{\text{Pal5}} = 20.37 \pm 0.51 \, (\pm 1.60)$ kpc, where the uncertainty is calculated as the quadrature sum of the individual uncertainties in each channel. This value can be compared to previous efforts to obtain a distance to Pal 5, such as Price-Whelan *et al* in 2019 where an average distance to Pal 5 of $20.6 \pm 0.36 (\pm 1.13)$ kpc was determined [28]. Therefore, the distance to Pal 5 obtained in this paper is consistent with this result. Another distance to Pal 5 of 23.6 ± 0.7 kpc was obtained by Küpper *et al* [29]. This determination is inconsistent with the distance obtained in this paper. The discrepancy between the values obtained could be due to considerations involving metallicity and the use of a Period-Luminosity-Metallicity (PLZ) relation, which was not considered in the method outlined in section 2, because there seemed to be an absence of data on the metallicities for the RRLs in this paper. Another cause for the discrepancy could be due to the correction for dust extinction for both the results in Table 1 and those in [28], which, as mentioned in Price-Whelan *et al*, was not considered in Küpper *et al*. This would account for the greater distance found in Küpper *et al*, as the extinction is subtracted from the uncorrected apparent magnitudes (section 2.4) which would shift the PL relation down by a constant factor, resulting in a greater intercept and therefore distance.

The distances to Pal 13 in both IRAC channels in Table 1 are also consistent with each other and so a mean distance to Pal 13 is determined to be $d_{\text{Pal13}} = 24.23 \pm 0.62 \, (\pm 1.90)$ kpc. Comparing to the literature, a value of 24.3 ± 1.1 kpc was obtained by Côté *et al* [30], which is consistent with

the distance determined in this paper. Furthermore, a distance of 23.6 ± 0.2 kpc was determined by Shipp *et al* [31], which is also consistent with the distance to Pal 13 obtained in Table 1. Therefore, $d_{\mathrm{Pal13}} = 24.23 \pm 0.62 \ (\pm 1.90)$ kpc is consistent with recent efforts to determine the distance to Pal 13.

4.2. Uncertainties and improvements

There are a number of sources of uncertainty associated with the measurements obtained for Pal 5 and Pal 13. Random sources of error arise from the least-squares fitting model used to determine the y-intercept of the PL fit, alongside the uncertainty in the calibrated intercept given in N19. The PL fit is ultimately based on the apparent magnitudes obtained during the photometry routine, and so photometric uncertainties due to background noise and the signal-to-noise ratio of the frame, blending or crowding of RRLs, nearby bright sources contaminating the background of the star and also errors in the fitting process of the ePSF models.

Some examples of RRLs these sources of uncertainties can be found by referring back to Figure 1. For example, in Figure 1 (a), the light curve, exhibits large error bars despite being in channel 1, which has a higher signal-to-noise ratio. Analysis of the image of this particular star showed evidence of blending or smearing and was found to be in close proximity to two other stars, which may explain the large uncertainties shown for this RRL, which can also be seen in Figure 2. Where as in Figure 1 (b) the channel 2 light curve shows a light curve with epoch data missing for two epochs, which was found to be common with channel 2 data, highlighting the potential effects of signal-to-noise ratio. Figure 1 (d) shows a channel 2 light curve for RRL4, an RRab star in Pal 13 and shows large scatter about its GLOESS fit. After viewing the image of this RRL, it was identified to be very close to two bright stars, which may have caused this scattering of data points.

The uncertainties for both Pal 5 and Pal 13 are around 7% of the determined distance value for both IRAC channels, which is considerably higher than the anticipated 1-2% for the SMHASH program. However, systematic uncertainty accounts for around 75% of the uncertainty associated with these distance measurements, which is caused primarily by the uncertainty in the gradient of the PL relations from N19, with uncertainties of $\sigma_b = 0.38$ and 0.39 for IRAC channels 1 and 2 respectively. If only the random uncertainty is considered then the uncertainty in the distance measurements reduces to around 1.8% for both Pal 5 and Pal 13. As noted in N19, this large systematic uncertainty is caused by the large uncertainties associated with parallax measurements in *Gaia*, which are used to calibrate the PL relations. Therefore, improvements in the calibrated PL relations are required to obtain an ideal accuracy of less than 2%; especially to resolve the discrepancies in the determinations of the Hubble constant discussed in section 1. These improved PL relations could be possible with improvements in parallax measurements in future *Gaia* data releases [32].

Further improvements to the accuracy of the distance measurements could be made by considering the metallicities of the RRL populations in the globular clusters and instead using a PLZ relation. This could reduce uncertainties by up to 1% [26] because of the predicted reduced metallicity effects for RRLs [4]. The small sample sizes in Pal 5 and Pal 13, containing a maximum of seven and four RRLs respectively, could also be a contributing factor. The sample sizes could be improved by including the RRLs residing in the tidal streams of both clusters, potentially improving the overall distance measurements.

4.3. Searching for new candidate RR Lyrae variables

Identification of new candidate RRLs was attempted by analysing a selection of statistical methods to measure variability and applying them to the apparent magnitudes for all detected stars in Pal 5. Three variability indices were chosen from Sokolovsky *et al* [33] (henceforth S17) to test for signs of variability associated with any stars detected that were not indicated in the CVS or *Gaia* DR2.

The first test was to calculate the standard deviation of the measured apparent magnitudes in each epoch for all detected stars using equation 3 in S17. The second test was to measure the inter-quartile range (IQR) which considers the difference between the upper and lower quartiles of

the apparent magnitudes measured for a particular star, splitting the central 50% of the data in half to form two sets of data, and then subtracting the median of each half from each other. The final test was the Welch-Stetson index, which was calculated using equation 16 in S17.

Analysis of the variability indices did not return the anticipated results and so no new candidate variable stars were found in Pal 5. RRLs previously known in the CVS and *Gaia* did not necessarily return high values, indicating variability, thus there was no confidence that stars returning high values for each of these tests were variable, which was confirmed when analysing the light curves of these stars. While unfortunate, it is likely this is due to the small number of epochs and the abundance of RRc type stars in Pal 5, perhaps because of the smaller amplitudes associated with RRc stars.

5. Future work

The next step would be to apply the work carried out onto more globular clusters in the MW galaxy, particularly those that are more crowded, such as the IC4499 globular cluster. IC4499 contains around 97 known RRLs from an initial investigation of the CVS, so this may require adjustments to the PSF photometry code to further optimize it for crowded fields.

A further development would be to consider the metallicities of the RRLs in the clusters, which could improve the accuracy of the distance measurements and potentially reducing the uncertainties associated with the measurements.

6. Conclusion

Distances to the Pal 5 and Pal 13 globular clusters have been calculated using data collected from the *Spitzer* IRAC in the mid-infrared spectrum. An average distance of $d_{\mathrm{Pal5}} = 20.37 \pm 0.51(\pm 1.60)$ kpc to the Pal 5 globular cluster was obtained and found to be consistent with the most recent measurement using a similar method, when the random and systematic uncertainties are considered. An average distance of $d_{\mathrm{Pal13}} = 24.23 \pm 0.62(\pm 1.90)$ kpc to the Pal 13 globular cluster was calculated and consistent with previous measurements undertaken.

Improvements in the systematic uncertainties as a result of the calibrated PL relations would significantly help achieve the aims of the SMHASH program. This could be achieved by improvements in the parallax measurements in future *Gaia* data releases, and significantly reduce the uncertainties in distance measurements [26, 32]. These developments are particularly exciting as this could help work towards refining the Hubble constant and its uncertainty as each rung of the cosmic distance ladder is climbed; having implications on the current understanding of the scale and evolution of the Universe [7].

References

[1] Monson A J, Beaton R L, Scowcroft V, Freedman W L, Madore B F, Rich J A, Seibert M, Kollmeier J A and Clementini G 2017 *The Astronomical Journal* **153** 96
[2] Garofalo A *et al.* 2018 *Monthly Notices of the Royal Astronomical Society* **481** 578–595

[3] Muraveva T, Garofalo A, Scowcroft V, Clementini G, Freedman W L, Madore B F and Monson A J 2018 *Monthly Notices of the Royal Astronomical Society* **480** 4138–4153

[4] Beaton R L *et al.* 2016 *The Astrophysical Journal* **832** 210

[5] Muraveva T, Clementini G, Garofalo A and Cusano F 2020 *Monthly Notices of the Royal Astronomical Society* **499** 4040–4053

[6] Scowcroft V, Freedman W, Madore B F, Monson A J, Persson S E, Seibert M, Rigby J R and Sturch L 2011 *The Astrophysical Journal* **743** 76

[7] Freedman W L, Madore B F, Scowcroft V, Burns C, Monson A, Persson S E, Seibert M and Rigby J 2012 *The Astrophysical Journal* **758** 24

[8] Planck C *et al.* 2016 *Astronomy & Astrophysics* **594** A13

[9] Freedman W L *et al.* 2001 *The Astrophysical Journal* **553** 47–72

[10] Erkal D, Koposov S E and Belokurov V 2017 *Monthly Notices of the Royal Astronomical Society* **470** 60–84

[11] Bradford J D, Geha M, Muñoz R R, Santana F A, Simon J D, Côté P, Stetson P B, Kirby E and Djorgovski S G 2011 *The Astrophysical Journal* **743** 167

[12] Reach W T *et al.* 2005 *Publications of the Astronomical Society of the Pacific* **117** 978–990

[13] Reach W T *et al.* 2006 *Infrared Array Camera Data Handbook, Version 3.0* (Spitzer Science Centre, Pasadena)

[14] Clement C M, Muzzin A, Dufton Q, Ponnampalam T, Wang J, Burford J, Richardson A, Rosebery T, Rowe J and Hogg H S 2001 *The Astronomical Journal* **122** 2587–2599

[15] Gaia C *et al.* 2018 *Astronomy & Astrophysics* **616** A1

[16] Kinman T and Rosino L 1962 *Publications of the Astronomical Society of the Pacific* **74** 499–506

[17] Howell S B 2006 *Handbook of CCD Astronomy* (Cambridge University Press) chap 5, pp 102–134

[18] Anderson J and King I R 2000 *Publications of the Astronomical Society of the Pacific* **112** 1360–1382

[19] Stetson P B 1987 *Publications of the Astronomical Society of the Pacific* **99** 191–222

[20] Bradley L *et al.* 2020 astropy/photutils: 1.0.1 URL `https://doi.org/10.5281/zenodo.4049061`

[21] Price-Whelan A M *et al.* 2018 *The Astronomical Journal* **156** 123

[22] Persson S E, Madore B F, Krzeminski W, Freedman W L, Roth M, and Murphy D C 2004 *The Astronomical Journal* **128** 2239–2264

[23] Monson A J, Freedman W L, Madore B F, Persson S E, Scowcroft V, Seibert M and Rigby J R 2012 *The Astrophysical Journal* **759** 146

[24] Ginsburg A *et al.* 2019 *The Astronomical Journal* **157** 98

[25] Schlafly E F and Finkbeiner D P 2011 *The Astrophysical Journal* **737** 103

[26] Neeley J R *et al.* 2019 *Monthly Notices of the Royal Astronomical Society* **490** 4254–4270

[27] Iben Jr I 1974 *Annual Review of Astronomy and Astrophysics* **12** 215–256

[28] Price-Whelan A M, Mateu C, Iorio G, Pearson S, Bonaca A and Belokurov V 2019 *The Astronomical Journal* **158** 223

[29] Küpper A H W, Balbinot E, Bonaca A, Johnston K V, Hogg D W, Kroupa P and Santiago B X 2015 *The Astrophysical Journal* **803** 80

[30] Côté P, Djorgovski S G, Meylan G, Castro S and McCarthy J K 2002 *The Astrophysical Journal* **574** 783–804

[31] Shipp N, Price-Whelan A M, Tavangar K, Mateu C and Drlica-Wagner A 2020 *The Astronomical Journal* **160** 244

[32] Neeley J R *et al.* 2015 *The Astrophysical Journal* **808** 11

[33] Sokolovsky K *et al.* 2017 *Monthly Notices of the Royal Astronomical Society* **464** 274–292