Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **482,** 1275–1303 (2019)
Advance Access publication 2018 September 26

doi:10.1093/mnras/sty2611

The WAGGS project – II. The reliability of the calcium triplet as a metallicity indicator in integrated stellar light

Christopher Usher [1]★ Thomas Beckwith,[1,2] Sabine Bellstedt [3] Adebusola Alabi [4] Leonie Chevalier,[3] Nicola Pastorello,[5] Pierluigi Cerulo,[6] Hannah S. Dalgleish [1] Amelia Fraser-McKelvie [7] Sebastian Kamann [1] Samantha Penny [8] Caroline Foster [9,10] Richard McDermid,[11,12] Ricardo P. Schiavon[1] and Alexa Villaume[4]

[1]*Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK*
[2]*Department of Physics, University of Liverpool, Oliver Lodge, Oxford Street, Liverpool L69 7ZE, UK*
[3]*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, VIC 3122, Australia*
[4]*University of California Observatories, 1156 High Street, Santa Cruz, CA 95064, USA*
[5]*Deakin Software and Technology Innovation Laboratory, Deakin University, Burwood, VIC 3125, Australia*
[6]*Departamento de Astronomia, Universidad de Concepcion, Casilla 160-C, Chile*
[7]*School of Physics and Astronomy, University of Nottingham, Nottingham, NG7 2RD, UK*
[8]*Institute of Cosmology and Gravitation, University of Portsmouth, Dennis Sciama Building, Burnaby Road, Portsmouth PO1 3FX, UK*
[9]*Sydney Institute for Astronomy, School of Physics, A28, The University of Sydney, NSW 2006, Australia*
[10]*ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia*
[11]*Department of Physics and Astronomy, Macquarie University, North Ryde, NSW 2109, Australia*
[12]*Australian Astronomical Observatory, PO Box 915, North Ryde, NSW 1670, Australia*

Accepted 2018 September 19. Received 2018 September 17; in original form 2018 June 11

ABSTRACT

Using data from the WiFeS Atlas of Galactic Globular cluster Spectra, we study the behaviour of the calcium triplet (CaT), a popular metallicity indicator in extragalactic stellar population studies. A major caveat of these studies is that the potential sensitivity to other stellar population parameters such as age, calcium abundance, and the initial mass function has not yet been empirically evaluated. Here, we present measurements of the strength of the CaT feature for 113 globular clusters in the Milky Way and its satellite galaxies. We derive empirical calibrations between the CaT index and both the iron abundance ([Fe/H]) and calcium abundance ([Ca/H]), finding a tighter relationship for [Ca/H] than for [Fe/H]. For stellar populations, 3 Gyr and older the CaT can be used to reliably measure [Ca/H] at the 0.1 dex level but becomes less reliable for ages of ∼2 Gyr and younger. We find that the CaT is relatively insensitive to the horizontal branch morphology. The stellar mass function however affects the CaT strengths significantly only at low metallicities. Using our newly derived empirical calibration, we convert our measured CaT indices into [Ca/H] values for the globular clusters in our sample.

Key words: stars: abundances – globular clusters: general – galaxies: abundances – galaxies: star clusters: general – galaxies: stellar content.

1 INTRODUCTION

Globular clusters (GCs) have two important roles in the study of stellar populations using integrated galaxy light. First, GCs in the Local Group provide important tests for stellar population models and analysis techniques as their ages, metallicities, and abundances are known independently (e.g. Schiavon 2007; Vazdekis et al. 2010; Conroy et al. 2018). Second, GCs play an important role as observational tracers of galaxy stellar light (e.g. Searle & Zinn 1978; Forbes & Bridges 2010; Brodie et al. 2014; Kruijssen et al. 2018). Present in virtually all galaxies with stellar masses $>10^9$ $M_\odot$ (see reviews by Brodie & Strader 2006; Forbes et al. 2018), their high surface brightness allow GCs to be studied at much greater distances than individual stars (GCs have been studied spectroscopically out to a distance of 47 Mpc, Misgeld et al. 2011 and with imaging out to a redshift of $z \sim 0.2$ e.g. Alamo-Martínez et al. 2013). Unlike the luminosity-weighted means provided by analysing

★ E-mail: c.g.usher@ljmu.ac.uk

integrated galaxy light, observations of GCs allow the distributions and correlations between stellar population parameters, positions, and kinematics to be studied. Since GCs are essentially single age, single metallicity (but not single-abundance pattern, see reviews by Gratton, Carretta & Bragaglia 2012; Bastian & Lardo 2017) stellar populations, it is easier to measure their stellar population parameters than it is for the field star populations of galaxies which contain a range of ages and metallicities.

The calcium triplet (CaT: 8498, 8542, and 8662 Å) is one of the strongest spectral features in the optical or near-infrared spectra of stars and old stellar populations. As a triplet of widely spaced lines, it is ideal for kinematic studies of individual stars and galaxies (e.g. Katz et al. 2004; Steinmetz et al. 2006; Pota et al. 2013; Arnold et al. 2014). The CaT wavelength region is less affected by extinction and is closer to the blackbody peak of the cool stars that dominate the luminosity of old or metal-rich stellar populations than commonly studied features such as Hβ and Mg*b* (e.g. Worthey et al. 1994). At almost all redshifts, the singly ionized calcium lines of the CaT overlap with a multitude of strong sky emission lines. Hence, measuring the equivalent width of the CaT requires careful sky subtraction. Thankfully, the CaT wavelength region is relatively unaffected by telluric absorption lines at low redshift.

In individual stars, the CaT displays complex behaviour with temperature, surface gravity, and metallicity. For both dwarfs and giants of all metallicities, the CaT strength peaks at effective temperatures between 4000 and 6000 K (e.g. Cenarro et al. 2002). In giants, the CaT strength increases strongly with decreasing surface gravity. However, at lower metallicities and in dwarf stars, the relationship between the CaT and surface gravity is weaker (e.g. Cenarro et al. 2002). The CaT is more sensitive to metallicity in giants and at temperatures between 5000 and 6000 K (e.g. Cenarro et al. 2002). As a result, the CaT has long been used to measure metallicities of individual giant stars in GCs (e.g. Armandroff & Da Costa 1991; Rutledge, Hesser & Stetson 1997; Da Costa 2016), in open clusters (e.g. Cole et al. 2004; Carrera et al. 2015) and in dwarf galaxies in the Local Group (e.g. Olszewski et al. 1991; Tolstoy et al. 2004; Cole et al. 2005; Battaglia et al. 2008; Leaman et al. 2013).

For integrated light (i.e. combined spectra of entire stellar populations), the depth of the CaT features increases with metallicity. The reasons for the CaT's sensitivity to metallicity are twofold. First, higher [Ca/H] directly leads to stronger CaT lines. Second, the effects of higher metallicity on stellar evolution push stars to cooler temperatures and giants to lower surface gravities. Both lower temperatures and lower surface gravities produce stronger CaT lines.

Early work (e.g. Spinrad & Taylor 1971; Cohen 1978) on integrated light spectra of stellar populations focused on the sensitivity of the CaT to surface gravity to study the initial mass function (IMF). The dependence of the CaT on metallicity was empirically confirmed by Bica & Alloin (1987) using spectra of 30 star clusters in the Milky Way (MW) and in the Large Magellanic Cloud (LMC). Armandroff & Zinn (1988) were the first to use the CaT as a metallicity indicator. They measured the strength of the CaT in 27 GCs in the MW and found a strong relationship between the CaT and metallicity.

The CaT has been used in studies of integrated galaxy light (e.g. Cenarro, Cardiel & Gorgas 2008; Foster et al. 2009; Pastorello et al. 2014), but its use as a metallicity indicator has been questioned due to the potentially significant effect of the IMF and [Ca/Fe] abundance (e.g. Saglia et al. 2002; Cenarro et al. 2003). In parallel, interest in the CaT as an IMF indicator has grown in recent years with claims (e.g. van Dokkum & Conroy 2010; Cappellari et al. 2012; La Barbera et al. 2013; Martín-Navarro et al. 2015; van Dokkum et al. 2017) that the centres of massive ETGs have relatively more dwarf stars (i.e. a bottom-heavy IMF) than the MW.

Recent studies (e.g. Goudfrooij et al. 2001; Brodie et al. 2012; Usher et al. 2013) have used the CaT to measure metallicities in unresolved extragalactic GCs. However, the reliability of the CaT as a GC metallicity indicator has been debated due to its known, but as yet unquantified, dependence on other parameters. For example, Foster et al. (2010, 2011) and Usher et al. (2012) found different relationships between GC colour and CaT strength in different galaxies. Different sets of stellar population synthesis models (Vazdekis et al. 2012; Chung et al. 2016) predict different CaT–metallicity relations. Sakari & Wallerstein (2016) observed a clear correlation between the CaT strength and [Fe/H] in M31 GCs. However, they also found that [Ca/Fe] affects the CaT strength in GCs. Despite this, Usher et al. (2012) found good agreement between Lick index (Worthey et al. 1994) based metallicities and their CaT-based metallicities.

As discussed by Sakari & Wallerstein (2016), there are advantages and disadvantages to studying GC stellar populations in the MW. As they are the closest GCs to us, they are the best studied, with measured detailed abundances, ages, and mass functions from studies of resolved stars (e.g. Carretta et al. 2009b; Dotter et al. 2010; VandenBerg et al. 2013; Paust et al. 2010; Mészáros et al. 2015; Sollima & Baumgardt 2017). However, due to their relative nearness, MW GCs appear more extended on the sky than their extragalactic counterparts with a median half-light radius of 1 arcmin [the Harris (1996, 2010) catalogue]. It is thus challenging to obtain integrated spectra that sample adequately all phases of stellar evolution. Additionally, the MW hosts a limited range of GC ages and abundances with no easily observable, massive star cluster younger than ∼8 Gyr (the age of Terzan 7, e.g. Dotter et al. 2010). By also studying star clusters in the MW's satellite galaxies, which span a wide range of ages (e.g. Geisler et al. 1997) and have smaller angular sizes (median half-light radius = 14 arcsec for old GCs in the LMC, McLaughlin & van der Marel 2005) but still have abundances from individual stars (e.g. Mucciarelli et al. 2008; Mucciarelli, Origlia & Ferraro 2010; Dalessandro et al. 2016) and ages from *Hubble Space Telescope (HST)* colour–magnitude diagrams (e.g. Olsen et al. 1998; Glatt et al. 2008b; Niederhofer et al. 2015), we can empirically address concerns about the dependence of the CaT on e.g. horizontal branch (HB) morphology, the IMF, age, and elemental abundances presented above.

In this paper, we use integrated spectra of GCs in the MW and its satellite galaxies (the LMC, Small Magellanic Cloud – SMC and the Fornax dSph) from the WiFeS Atlas of Galactic Globular cluster Spectra (WAGGS, Usher et al. 2017) to study the behaviour of the CaT spectral feature as a function of a number of stellar population and GC parameters. As in Usher et al. (2017), we will use the term GC to refer to all massive ($>10^4$ $M_\odot$) star clusters irrespective of age. However, we make a distinction between GCs and objects which likely are or once were the nuclei of galaxies, namely NGC 5139 (ω Cen) and NGC 6715 (M54) (e.g. Hilker & Richtler 2000; Ibata et al. 1997). We note that both NGC 5139 and NGC 6715 have extended star formation histories (>2 Gyr, e.g. Layden & Sarajedini 2000; Stanford et al. 2006; Siegel et al. 2007; Villanova et al. 2014) and large-metallicity spreads (>1 dex, e.g. Johnson & Pilachowski 2010; Carretta et al. 2010a), while GCs have little or no age or metallicity spreads.

This paper is organized as follows. In Section 2, we present our sample and describe our observations and data reduction. In Section 3, we present our measurement technique and discuss some of the systematics affecting our measurements. In Section 4, we

discuss the effects of metallicity, age, Ca abundance, stochasticity, HB morphology, and the present-day mass function on the CaT. In Section 5, we summarize our results.

2 SAMPLE AND OBSERVATIONS

The first WAGGS paper (Usher et al. 2017) provided a detailed discussion of the scientific aims of the project, described the initial WAGGS sample and observations, detailed our data reduction, presented a comparison of repeated observations and described the publicly released spectra. Here, we provide a brief outline of our sample, observations, and data reduction, highlighting any additions or differences from what was presented in Usher et al. (2017).

We use an expanded version of the WAGGS sample presented in Usher et al. (2017). As detailed in that paper, the sample consists of GCs in the MW, the LMC, the SMC, and the Fornax dSph observable from the Siding Spring Observatory with central surface brightness brighter than $\mu_V \sim 20$ mag arcsec^2. We favoured GCs with high-quality *HST* photometry (e.g. Sarajedini et al. 2007) and GCs with abundances from high-resolution spectroscopy. Our sample contains both NGC 5139 and NGC 6715 which are thought to be the nuclear remnants of tidally disrupted (or disrupting) dwarf galaxies accreted by the MW (e.g. Hilker & Richtler 2000; Ibata et al. 1997). We will refer two these two objects as nuclear remnants throughout the rest of the paper. Additions from the previously presented sample in Usher et al. (2017) are 11 GCs in the LMC and SMC and 16 GCs in the MW including a handful of well studied but lower surface brightness MW GCs (e.g. NGC 288 and NGC 6496) and a number of relatively bright but poorly studied MW GCs (e.g. NGC 6626, NGC 6638, and NGC 6642). We also observed additional pointings of a handful of nearby GCs (NGC 3201, NGC 6121, and NGC 6397) with low masses in our field of view in an attempt to decrease stochastic effects due to the low number of giant stars in the field of view. Our sample spans a wide range of ages and metallicities (see Fig. 1). Details of our sample are given in Table 1. In total, we analyse 138 spectra of 113 GCs.

We used the surface brightness profiles calculated from the structural parameters given in Table 1 using the LIMEPY code (Gieles & Zocchi 2015) and a King (1966) profile to calculate the *V*-band luminosity within the field of view of each datacube. We converted these enclosed luminosities into masses using the same mass-to-light ratios (*M*/*L*) used in Usher et al. (2017), namely a constant $M/L = 2$ for all GCs older than 10 Gyr and predictions of the Bruzual & Charlot (2003) stellar population synthesis models for the younger clusters.

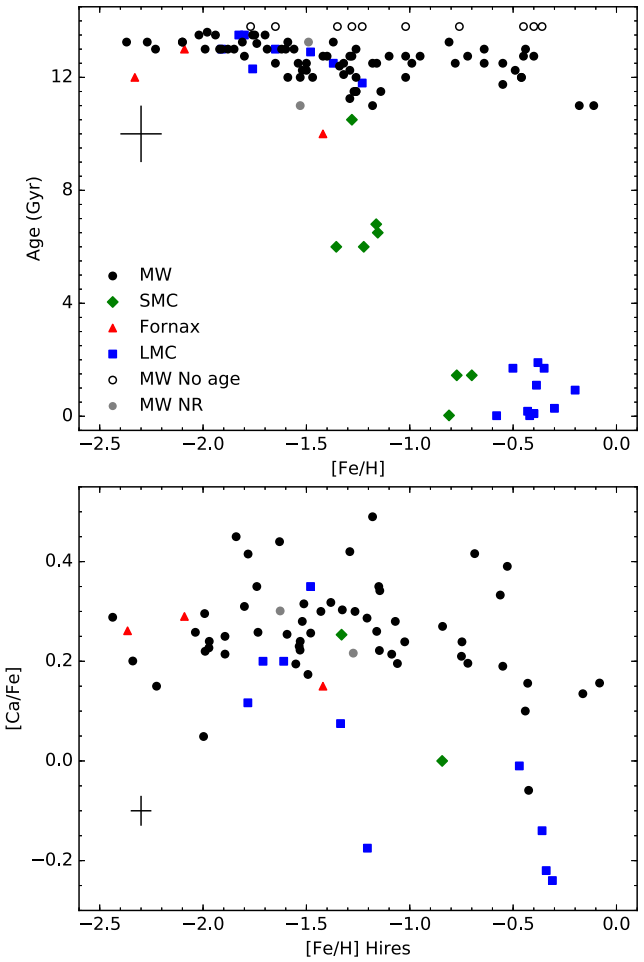

Figure 1. Literature stellar population parameters used in this work. MW GCs are marked as black circles, LMC GCs as blue squares, SMC GCs as green diamonds, and Fornax dSph GCs as red triangles. The nuclear remnants in the MW (NGC 5139 and NGC 6715) are denoted as grey circles. We note that metallicities, ages, and abundances are from a wide range of sources detailed in the text. Typical uncertainties are provided as a cross in each panel. Top: GC age as a function of metallicity. MW GCs with no literature ages have been assigned an age of 13.8 Gyr and are marked as hollow circles. Each galaxy follows its own age–metallicity relation. Bottom: [Ca/Fe] versus [Fe/H] from high-resolution spectroscopy. At fixed metallicity our sample covers a wide range of [Ca/Fe] values.

2.1 Observations and data reduction

As described in Usher et al. (2017), we used the WiFeS instrument (Dopita et al. 2007, 2010) on the Australian National University 2.3 m telescope at the Siding Spring Observatory. WiFeS is a dual-arm, 38 × 25 arcsec image slicer integral field spectrograph. For this work, we only used data observed with the RT615 beam splitter and I7 000 grating. This setup covers 6800–9050 Å with 0.57 Å pixel^{-1} and provides a spectral resolution of $\delta\lambda/\lambda \sim 6800$ in the region of the CaT. To perform reliable sky subtraction, we used WiFeS in nod-and-shuffle mode (Glazebrook & Bland-Hawthorn 2001). We used the PyWiFeS (Childress et al. 2014a,b) pipeline to reduce the data. We used 2MASS (Skrutskie et al. 2006) *J*-band images to perform astrometry on our datacubes before applying heliocentric velocity corrections. To create integrated light spectra, we simply summed the spatial pixels of each datacube ignoring the first two and last two rows due to their significantly higher noise. Unlike in Usher et al. (2017), we used the entire datacube for all GCs and did not limit our extracted spectra to within 1 half-light radius for GCs with more extended spatial coverage. Please see Usher et al. (2017) for a detailed description of the observations and data reduction.

Our observations span nights from 2015 January to 2017 July. In addition to the observations presented in Usher et al. (2017), we utilized a series of observations of NGC 104 from our 2016 September 29 to October 3 observing run. We also utilized observations from our 2017 April 3 to April 6 run and our 2017 June 29 to July 2 run. These observations were conducted using the same instrument set up as those presented in Usher et al. (2017) and reduced in an identical manner. We plot examples of our spectra in the region of the CaT in Figs 2 and 3.

Table 1. Sample properties.

ID	Galaxy	[Fe/H]	Age	R_c	R_h	μ_{V0}	A_V	GC Mass	FoV Mass	[Fe/H] Source	Age Source	Structural Source
		(dex)	(Gyr)	(arcsec)	(arcsec)	(mag arcsec^{-2})	(mag)	(log $M_\odot$)	(log $M_\odot$)			
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
NGC 104	MW	−0.72	12.8	21.6	190	14.4	0.12	6.0	4.7	Harris (2010)	Dotter et al. (2010)	Harris (2010)
Kron 3	SMC	−1.15	6.5	21.1	39.9	20.1	0.05	5.2	4.4	Da Costa & Hatzidimitriou (1998)	Glatt et al. (2008a)	McLaughlin & van der Marel (2005)
NGC 121	SMC	−1.28	10.5	9.6	19.0	18.3	0.45	5.6	5.2	Dalessandro et al. (2016)	Glatt et al. (2008a)	McLaughlin & van der Marel (2005)
NGC 288	MW	−1.32	12.5	81.0	134	20.0	0.09	4.9	3.2	Harris (2010)	Dotter et al. (2010)	Harris (2010)
NGC 330	SMC	−0.81	0.03	8.1	21.0	16.5	0.20	4.6	4.1	Hill (1999)	Sirianni et al. (2002)	McLaughlin & van der Marel (2005)
...	...	...	...	...	...	...	...	...	...	...	...	...

Notes: Column (1): GC name. Column: (2) host galaxy. Column (3): metallicity in dex. Column (4): age in Gyr. Column (5): projected core radius in arcseconds. Column (7): *V*-band central surface brightness in mag arcsec^{-2}. Column (8): *V*-band extinction in mag. Column (9): GC log stellar mass in solar masses calculated from *V*-band luminosity. Column (10): log stellar mass in solar masses enclosed by the WiFeS field of view calculated from the surface brightness profile. Columns (11)–(13): sources for [Fe/H], age, and structural parameters, respectively. Harris (2010) refers to the Harris (1996, 2010) catalogue. The full version of this table is provided in a machine readable form in the online Supporting Information.

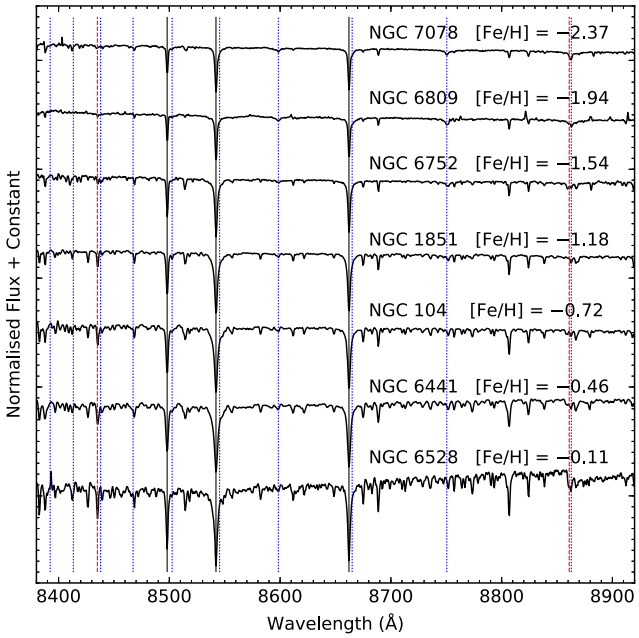

Figure 2. Effects of metallicity on integrated GC spectra in the CaT region. The GCs plotted are all old (age > 11 Gyr) and increase in metallicity from top to bottom. The CaT lines are identified by solid black vertical lines at the top, the hydrogen Paschen lines as dotted blue lines, and two TiO bandheads as red dashed lines. The CaT lines increase with metallicity as do the other metal lines. Weak Paschen absorption is only present in the most metal-poor GCs ([Fe/H] < −2) and significant TiO absorption is only seen at near solar metallicity. We note that fitted velocity dispersions of these GC range from ∼5 km s^{-1} (NGC 6528) to ∼17 km s^{-1} (NGC 6441) which is comparable to the instrument resolution (a velocity dispersion of 19 km s^{-1}).

2.2 Literature metallicities and abundances

We use three sets of metallicities and abundances in this work. First, we find a literature metallicity for every GC in our sample. For the MW, we adopt the metallicities provided by the Harris (1996, 2010) catalogue. While the sources of metallicities in the Harris catalogue are heterogeneous, an attempt has been made to place them on the Carretta et al. (2009b) scale. For the three GCs in the Fornax dSph, we adopt the metallicities derived by Larsen, Brodie & Strader (2012). For GCs in the LMC and SMC, we preferred metallicities from high-resolution spectroscopy to those from the resolved star CaT strengths, and CaT-based metallicities to those based on resolved colour–magnitude diagrams. We adjusted the metallicity measurements based on the CaT strengths of red giant branch (RGB) stars to the Carretta et al. (2009b) metallicity scale. For NGC 1466, we follow the recommendation of Walker (1992) and adopt a metallicity for NGC 1466 based on the more metal-rich star ([Fe/H] = −1.91 using the Carretta et al. 2009b calibration) from Olszewski et al. (1991). Da Costa & Hatzidimitriou (1998) measured the strength of the CaT for four RGB stars in NGC 361 but did not have sufficiently high-quality photometry to convert these measurements into metallicities. Using the *HST* WFPC2 photometry of Mighell, Sarajedini & French (1998) and the CaT−metallicity calibration of Da Costa (2016), we derive a metallicity of [Fe/H] = −1.16 for NGC 361. A challenge for all LMC- and SMC-based metallicities is that their ages and [α/Fe] can be quite different to MW GCs of the same metallicity. Metallicities derived using techniques calibrated on MW GCs (i.e. CaT strengths of RGB stars) may be less reliable.

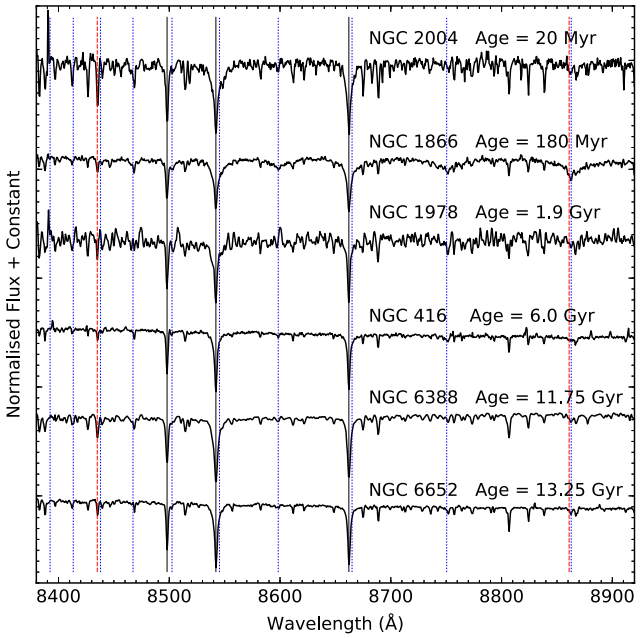

Figure 3. Effects of age on integrated GC spectra in the CaT region. The GCs increase in age from top to bottom. The GCs are generally LMC and SMC metallicity ([Fe/H] = −0.8 to −0.4) save for NGC 416 which is significantly more metal poor ([Fe/H] = −1.2). As in Fig. 2, the CaT lines are identified by solid black vertical lines at the top, the hydrogen Paschen lines as dotted blue lines and two TiO bandheads as red dashed lines. At old ages (>5 Gyr), age seems to have little effect on the CaT spectral region as the light is dominated by the RGB. At younger ages (∼2 Gyr), the thermally pulsing AGB becomes important and the spectra can be dominated by carbon stars. At even younger ages (∼200 Myr), hotter stars from the MSTO play an important role and strong Paschen line absorption occurs. At the youngest ages (∼20 Myr), the CaT region is dominated by red supergiants which show stronger metal line absorption compared to RGB-dominated old populations with the same metallicities.

Second, we assembled a sample of Fe abundances ([Fe/H]) and Ca abundances [Ca/Fe]) based on high-resolution ($R > 20\,000$) spectroscopy of individual stars. We gave preference to studies of the RGB to avoid any biases from comparing measurements from different evolutionary stages. We also gave preference to studies with large numbers of stars. We shifted all the abundances to the Asplund et al. (2009) abundance scale. For GCs with multiple studies, we took the error-weighted mean of the available studies. When no mean abundance for the GC was provided by the authors, we took the simple mean of the abundances of the individual stars as the abundance of the cluster. If uncertainties for each individual star were provided we used these to calculate the uncertainty in the mean; otherwise the statistical uncertainty in the mean is given by $\sigma/\sqrt{N-1}$, where σ is the standard deviation of the individual star measurements, and N is the number of stars. We were careful to add in quadrature the systematic uncertainties provided by the authors to the statistical uncertainty; if no systematic uncertainties were provided we assumed a 0.05 dex systematic uncertainty in [Fe/H] and 0.02 dex in [Ca/Fe] as found by Carretta et al. (2009b) and Carretta et al. (2010b). While we aimed to be comprehensive, our high-resolution sample should not be seen as complete.

In the case of the Carretta et al. (2010c) study of NGC 6715, we use both the stars identified by the authors as members of NGC 6715 and as members of the nucleus of the Sagittarius dwarf galaxy to calculate mean [Fe/H] and [Ca/Fe] abundances as both are captured by our observations. Gratton et al. (2015) studied red HB stars in NGC 6723. They note that Ca abundances derived from HB stars are systematically higher than those from the RGB. As such we lower their [Ca/Fe] measurement by their recommended 0.2 dex. We took APOGEE abundances from Mészáros et al. (2015) and Schiavon et al. (2017). The Gaia–ESO (Gilmore et al. 2012) [Fe/H] measurements are from Pancino et al. (2017), while Gaia–ESO [Ca/Fe] measurements were taken from the ESO Science Archive Facility. We supplement this high-resolution sample with abundances from high-resolution integrated studies for Fornax 3, Fornax 4, and Fornax 5 (Larsen et al. 2012) as well as for NGC 1916 (Colucci et al. 2011, 2012). While we do include the Johnson & Pilachowski (2010) study of NGC 5139 and the Carretta et al. (2010c) study of NGC 6715 in our plots, we do not include these nuclear remnants in our analysis due to their significant metallicity spreads. We plot our adopted high-resolution [Fe/H] and [Ca/Fe] values in Fig. 1 and list our adopted values in Table 2.

Third, we use the sample of Carretta et al. (2009b) to provide a homogeneous sample of abundances. We use the ESO VLT UVES-based values for [Fe/H] from Carretta et al. (2009b) and for [Ca/Fe] from Carretta et al. (2010b). To enlarge the sample, we include subsequent studies by the same authors using the same instrument and analysis. We also include the Carretta et al. (2001) study of NGC 6528 and the Cohen et al. (1999) study of NGC 6553 as corrected by Carretta et al. (2007a). We note that these two GCs were studied using Keck HIRES observations of red HB stars, rather than the VLT UVES observations of RGB stars used in the rest of the Carretta et al. sample. We do not include NGC 6715 in our Carretta et al. sample due to the significant metallicity spread in what is likely the nucleus of the Sagittarius dE.

Table 2. Adopted high-resolution abundances.

Name	[Fe/H] (dex)	[Ca/Fe] (dex)	Sources
(1)	(2)	(3)	(4)
NGC 104	−0.74 ± 0.02	0.23 ± 0.01	Carretta et al. (2004)
			Alves-Brito et al. (2005)
			Koch & McWilliam (2008)
			Carretta et al. (2009b)
			Carretta et al. (2010b)
			Thygesen et al. (2014)
			Cordero et al. (2014)
			Pancino et al. (2017)
NGC 121	−1.35 ± 0.03	0.24 ± 0.04	Johnson et al. (2004)
			Dalessandro et al. (2016)
NGC 288	−1.30 ± 0.07	0.41 ± 0.03	Carretta et al. (2009b)
NGC 330	−0.84 ± 0.04	0.00 ± 0.05	Hill (1999)
NGC 362	−1.14 ± 0.03	0.28 ± 0.04	Carretta et al. (2013)
			Pancino et al. (2017)
...	...	...	...

Note: Column (1): GC name. Column (2): [Fe/H] abundance in dex. Column (3): [Ca/Fe] abundance ratio in dex. Column (4): source(s). The full version of this table is provided in a machine readable form in the online Supporting Information.

3 MEASURING THE CALCIUM TRIPLET

3.1 Measurement technique

We measured the strength of the CaT using the template fitting method of Foster et al. (2010) and Usher et al. (2012). In this method,

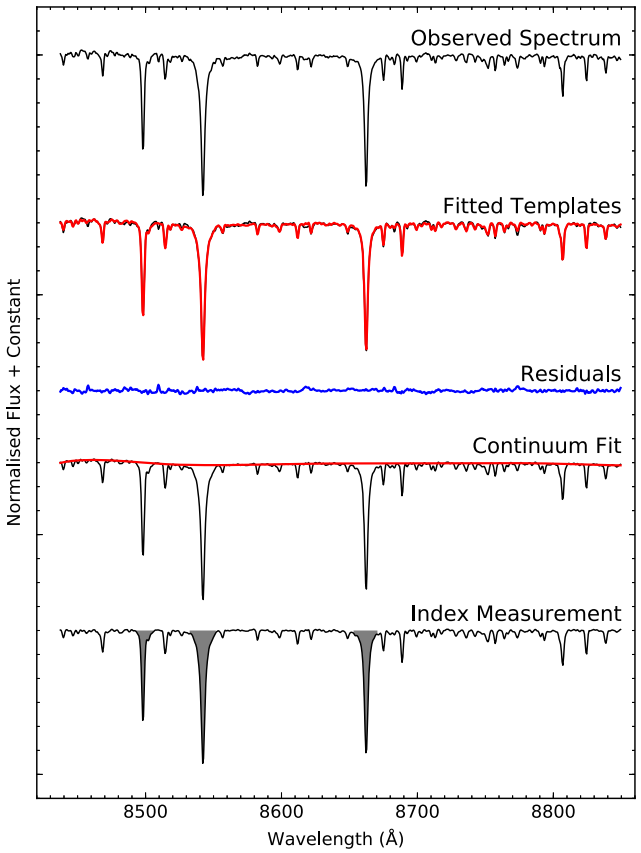

Figure 4. Measuring the CaT by template fitting. First from the top is our observed spectrum of NGC 2808 shifted to the rest frame. Second from the top, we plot the fitted combination of stellar templates in red over our observed spectrum in black. Third from the top, in blue we plot the difference of our fitted spectrum and our input spectrum. Fourth from the top, in red we plot the continuum fitted to the templates in black. Fifth from the top, we plot the continuum normalized templates in black and the region between the normalized templates and unity in our CaT bandpasses in grey.

wavelength regions affected by sky emission lines are masked before the observed spectrum is fitted by a linear combination of stellar templates. The fitted combination of templates is then continuum normalized and the equivalent width of the CaT is measured on the normalized combination of templates. This process allows the CaT to be reliably measured from low signal-to-noise (S/N) spectra with strong residuals from sky subtraction. We plot a summary of the steps in our CaT measurement method in Fig. 4. As in our previous work, we used the same index definition as Armandroff & Zinn (1988, 8490.0 to 8506.0, 8532.0 to 8552.0, and 8653.0 to 8671.0 Å) to measure the CaT.

As in Foster et al. (2010) and Usher et al. (2012), we used the PPXF pixel fitting code (version 5.1.10 Cappellari & Emsellem 2004) to fit the template stars to the observed spectra as well as the radial velocity, velocity dispersion, and a continuum normalization. As in Usher et al. (2012), we fit a seventh-order continuum polynomial and used an initial guess of 10 km s^{-1} for the velocity dispersion. For the initial guess of the radial velocity, we used the literature values presented in Table 1. For a handful of GCs (NGC 416, NGC 1850, NGC 1856, and NGC 2004) in the LMC or SMC, we could not locate literature radial velocities so we used the systemic velocities of these galaxies (278 and 158 km s^{-1} respectively, from NED[1]). A major improvement over Usher et al. (2012) is using the PYTHON version of PPXF (Cappellari & Emsellem 2004) rather than the IDL version. This change resulted in over a factor of 100 speedup. The measurement code is available from https://github.com/chrusher/measure_CaT.

3.1.1 *Effects of skyline masks*

We ran PPXF both using the same sky line masks as in Usher et al. (2012) and without masking any pixels affected by sky emission lines. We note that PPXF performs a weighted fit so the higher variance sky line regions are downweighted. Additionally, we allow PPXF to clip pixels more than 3σ from the fit. Another difference from Usher et al. (2012) was shifting the lower wavelength limit of the fitted spectra to 8437 from 8425 Å to avoid the TiO bandhead at 8420 Å as the presence of this molecular feature complicates the determination of the continuum. Our upper wavelength limit of our fit, 8850 Å, was chosen to avoid the strong TiO bandhead at 8860 Å and remains unchanged from Usher et al. (2012).

In the top panel of Fig. 5, we show comparisons of the CaT strengths measured by fitting masked and unmasked spectra. There is significant scatter (root-mean squared – rms – difference of 0.27 Å) in the difference between the CaT strength measured with the skylines masked and unmasked, but only a minor median difference (0.04 Å). In Fig. 6, we plot this difference as a function of radial velocity. The difference between the masked and unmasked measurement peaks around radial velocities −120 and +240 km s^{-1}. This is where the sky line masks overlap with the CaT feature. GCs with radial velocities between −200 and −40 km s^{-1} and between 120 and 320 km s^{-1} having an rms difference of 0.32 Å, while GCs outside of these radial velocity ranges have an rms difference of 0.20 Å between the masked and unmasked measurements.

3.1.2 *Effects of stellar templates*

We fitted our observed spectra using the same 13 stellar templates (observed using the DEIMOS spectrograph by Faber et al. 2003 on the Keck II telescope) as used by Foster et al. (2010) and Usher et al. (2012), since the spectral resolution of these templates is similar to that of WiFeS. Details of the DEIMOS templates are provided in Table 3. As the DEIMOS templates only cover a limited range of stellar parameters – for example they lack any metal-poor dwarfs – we also experimented with using stellar templates drawn from the higher resolution (FWHM ∼1 Å) Indo-US library (Valdes et al. 2004). We did not broaden the Indo-US templates to match the WiFeS spectral resolution, instead relying on the fitting code to broaden the fitted templates to match the observed spectra. We selected 40 stars from the Indo-US library that had spectral coverage of the CaT region and covered the range of effective temperatures, surface gravities, and metallicities expected for GC stars with a wide range of ages and metallicities. We also include a carbon star to check whether it would improve the fit to the spectra of intermediate-age GCs. Details of the Indo-US templates are provided in Table 4. For the DEIMOS and Indo-US templates, we included a constant flux template. For the DEIMOS templates, this constant flux template is required to fit low-metallicity spectra, while for the Indo-US templates, the inclusion of a constant flux

[1]The NASA/IPAC Extragalactic Database is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

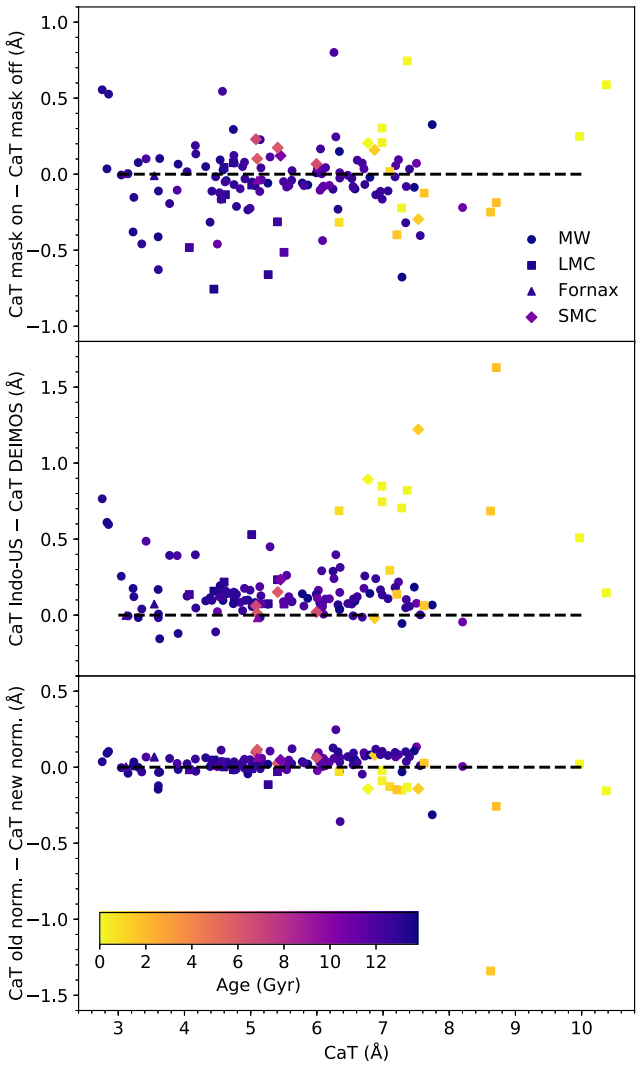

Figure 5. Comparison of the CaT measurements using different parameters. Top: difference between CaT measured with and without masking sky emission lines. Middle: difference between CaT measured using the Indo-US templates and the DEIMOS templates. Bottom: difference between the continuum normalization used in Usher et al. (2012) and the one used here. The points are colour coded by age with MW GCs denoted by circles, LMC GCs by squares, SMC GCs by diamonds, and Fornax dSph GCs by triangles. We see good agreement between the masked and unmasked CaT measurements but see large scatter. Modulo a small constant offset, we see agreement between the DEIMOS and Indo-US templates for most old GCs. However, the most metal-poor GCs and the youngest GCs show relatively stronger Indo-US based CaT measurements. We see little difference between the old and new continuum method save at the highest CaT values where older GCs show slightly stronger CaT values using the new method and young GCs show slightly weaker values.

template has no measurable effect on the fits. We did not use the Cenarro et al. (2001) library of stellar spectra due to their lower spectral resolution (FWHM $= 1.5$ Å versus 1.26 Å for WiFeS).

In the middle panel of Fig. 5, we show a comparison of the DEIMOS and Indo-US templates. The CaT measurements using the Indo-US templates are larger by a median difference 0.12 Å (rms 0.31 Å) although the difference and scatter is smaller for GCs 3 Gyr and older (median difference of 0.11 and rms difference of 0.19 Å). However, at the lowest CaT strengths, the Indo-US-based

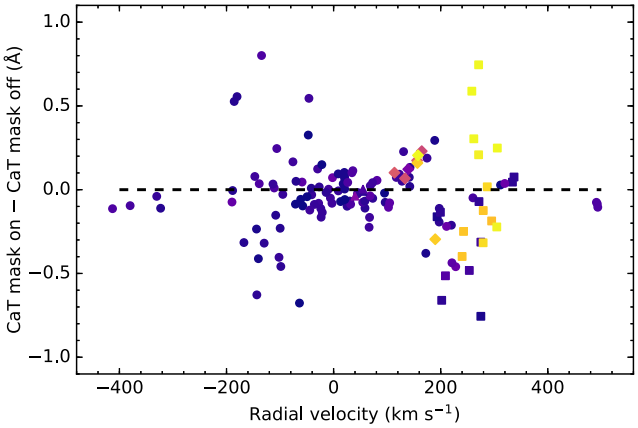

Figure 6. Differences between CaT measured with and without masked sky emission lines as a function of radial velocity. As in Fig. 5, the points are colour coded by age with MW GCs denoted by circles, LMC GCs by squares, SMC GCs by diamonds, and Fornax dSph GCs by triangles. The CaT strength difference between the masked and unmasked spectra increases at radial velocities where skylines overlap with the CaT lines.

CaT strengths are stronger than the DEIMOS ones by ∼0.6 Å. Most younger GCs (<2 Gyr) also show stronger Indo-US-based values than DEIMOS-based ones. For both the lowest metallicity GCs and for most of the younger GCs, this difference is due to the Indo-US templates better fitting the Paschen lines of hotter stars, since the DEIMOS templates lack any stars hotter than the 6100 K HD 136202 (Maldonado et al. 2015).

To assess the effects of the template fitting process on our measurements, we also directly ran our normalization code on the raw spectra. We plot a comparison of the CaT measured from the directly normalized spectra with the CaT values from the fitted templates in Fig. 7. At high S/N (>250 Å^{-1}), the CaT values directly measured from the spectra are offset to larger CaT values by 0.21 ± 0.03 Å independent of CaT strength. At lower S/N, the directly measured CaT values are offset to even larger values, since at lower S/N our continuum fitting process overestimates the height of the continuum when applied to noisier spectra. This is not an issue for our template fitting process as it is the high S/N fitted templates that are continuum-normalized.

3.1.3 Effects of continuum determination

In this work, we refined the continuum determination method used in Usher et al. (2012). We note that at the spectral resolution considered in this work, there are no true regions of continuum in the spectral range we study. As such, we fit a pseudo-continuum to regions of low line opacity by first masking wavelength regions with strong absorption before fitting a linear combination of polynomials with least squares to the masked templates. Pixels more than 0.4 per cent below or 2 per cent above the fitted polynomials were rejected and the polynomials were refitted. This continuum fitting process was repeated until no more pixels were rejected. We note that the wavelength mask is only used in the first iteration of the fitting.

In Usher et al. (2012), a wavelength mask based on the Vazdekis et al. (2003) stellar population models was used. We defined a new wavelength mask in the following manner. We used synthetic spectra calculated with ATLAS12 (Kurucz 1970, 2013) and SYNTHE (Kurucz 1993) of a $T_{\rm eff} = 4300$, $\log g = 1.5$, [Fe/H] $= -0.5$ giant

Table 3. DEIMOS templates.

Star	Type	$T_{\rm eff}$ (K)	log g (log cm g s^{-2})	[Fe/H] (dex)	Reference
(1)	(2)	(3)	(4)	(5)	(6)
HD 4388	K3III				
HD 35410	G9III	4800	2.6	−0.33	Liu et al. (2014)
HD 36079	G5II	5200	2.5	−0.25	Liu et al. (2014)
HD 44131	K5III				
HD 65934	G8III				
HD 74377	K3V	4670	4.5	−0.37	Prugniel, Vauglin & Koleva (2011)
HD 107328	K1III	4340	1.7	−0.52	Boeche & Grebel (2016)
HD 136202	F8IV	6050	3.8	−0.10	Boeche & Grebel (2016)
HD 161096	K2III	4590	2.6	+0.10	Boeche & Grebel (2016)
HD 237903	K7V	4080	4.7	−0.26	Prugniel et al. (2011)
NGC 1904−S71		5140	2.0	−1.81	Kirby, Guhathakurta & Sneden (2008)
NGC 1904−S174		4450	1.1	−1.77	Kirby et al. (2008)
NGC 2419−S223		4250	0.5	−2.09	Cohen, Huang & Kirby (2011)

Notes: Column (1): name of star. Column (2): spectral type according to SIMBAD (Wenger et al. 2000). Column (3): effective temperature in K. Column (4): log surface gravity in cm g s^{-2}. Column (5): [Fe/H] in dex. Column (6): source for atmospheric parameters.

and a $T_{\rm eff} = 4240$, log $g = 4.7$, [Fe/H] $= -0.5$ dwarf to define a new wavelength mask, stellar parameters typical of an old (12.6 Gyr) stellar population with this metallicity (e.g. Choi et al. 2016). As in Kirby et al. (2008), we identified continuum regions by smoothing the model spectra to the resolution of WiFeS ($R = 6800$) and selecting wavelength regions where the ratio of the smoothed flux to the continuum was greater than 0.95. To ensure that the continuum regions, and the sections between them, were at least 1.2 Å wide (2 pixels for WiFeS), we adjusted their limits. We then compared the new continuum mask with our integrated spectra of NGC 2808, NGC 6388, NGC 6528, NGC 7078, and NGC 7089 to identify absorption features either present in the observed spectra but not the model spectra or vice versa. Most notably, we masked the cores of the hydrogen Paschen lines which are visible in metal-poor GCs. We change the order of the polynomial used to fit the continuum to 7 to match the order of the polynomial used by PPXF for fitting the continuum. We also changed the limits for pixel rejection during normalization to include pixels up to 2 per cent above the model rather than 1 per cent above as in Usher et al. (2012). This resulted in faster and more stable determinations of the continuum level.

In the bottom panel of Fig. 5, we show comparison Usher et al. (2012) normalization with the updated normalization adopted here. Most GCs show almost no difference between the CaT strengths measured using old and new normalizations (median difference of 0.03 Å and rms difference of 0.14 Å). However, at high CaT strengths (>6.5 Å), the old normalization shows slightly higher CaT values for old GCs (median difference 0.07 Å and rms difference 0.10 Å) and slightly lower CaT strengths for young GCs (median difference −0.13 Å and rms difference 0.37 Å).

3.2 CaT measurement systematics

We investigated the effects of radial velocity, velocity dispersion, and S/N on our measurement process by using our spectra of NGC 2808, NGC 6388, NGC 6528, NGC 7078, and NGC 7089. These GCs were selected as the highest S/N spectra spanning the range of observed metallicities. To simulate the effect of observing GCs at different radial velocities, we shifted these GC spectra to radial velocities between −500 and 1500 km s^{-1} in steps of 25 km s^{-1}. At each of these velocities, we repeated our CaT measurement process. The effects of these radial velocity shifts on our CaT measurements are shown in Fig. 8. When the radial velocity shifts the sky line wavelength mask on to the CaT lines, such as at −125 and +225 km s^{-1}, the CaT measurement becomes biased. Our choice of templates has no effect on this radial velocity bias. As can be seen in Fig. 6, we observe this bias in the difference between our masked and unmasked CaT measurements. These variations are smaller at radial velocities above 500 km s^{-1} where extragalactic GCs are typically studied. Even when we do not mask the spectra, we still expect the CaT measurement to be noisier at radial velocities where the CaT spectral features overlap with skylines.

We velocity broaden these GC spectra to a range of 20 velocity dispersions evenly spaced in logarithmic space between 20 and 400 km s^{-1} accounting for the resolution of our spectra (18.8 km s^{-1} in velocity dispersion). This is equivalent to lowering the spectral resolution of our data from $R = 6800$ to a range of resolutions between $R = 6400$ and 320. This velocity dispersion range spans the observed range of not just GCs and nuclear remnants but also galaxies (e.g. Norris et al. 2014). The effects of velocity dispersion (or equivalently spectral resolution) are shown in Fig. 9. For low and intermediate metallicities, the CaT strength is relatively insensitive to velocity dispersion up to ∼100 km s^{-1} ($R \sim 1200$) before declining with velocity dispersion. For higher metallicity GCs, the CaT declines at all velocity dispersions. For velocity dispersions above ∼100 km s^{-1}, we can not effectively distinguish between the CaT strengths of [Fe/H] $= -0.55$ (NGC 6388) and [Fe/H] $= -0.11$ (NGC 6528) using our technique. Whether we mask the sky lines or not and our choice of templates has virtually no influence on the effect of velocity dispersion on CaT measurements. We note that the velocity dispersion of GCs – 0.6 to 23 km s^{-1} with a median of 6 km s^{-1} in the MW (Baumgardt & Hilker 2018) – is smaller or comparable to the instrument resolution. Thus, the velocity dispersion of our GCs has no effect on our measured CaT strengths except for our most metal rich and most massive GCs where the effect is still relatively small (∼0.1 Å). The effects of velocity dispersion would be important in studies of the galaxy light and the effects of spectral resolution would be important for studies carried out with a significantly lower resolution spectrograph.

Table 4. Indo-US templates.

Star	Type	$T_{\rm eff}$ (K)	log g (log cm g s^{-2})	[Fe/H] (dex)
(1)	(2)	(3)	(4)	(5)
G 37−26	A4p	6016	4.4	− 1.95
G 48−29	sd:A2	6295	4.0	− 2.66
G 102−27	G0	5423	4.0	− 1.05
G 163−78	K0	5400	4.2	+0.45
G 165−39	A4	6330	4.0	− 1.96
G 176−11	M0.5	3544	4.9	+0.00
G 245−32	sd:F2	6346	4.5	− 1.62
HD 5916	G8III	4755	2.0	− 0.80
HD 11636	A5V	9000	4.0	+0.16
HD 19476	K0III	4940	3.1	+0.04
HD 19510	A0	6109	2.6	− 2.50
HD 25329	K1V	4840	4.9	− 1.68
HD 39587	G0V	5953	4.5	− 0.03
HD 44007	G5IV	4850	2.0	− 1.71
HD 44478	M3III	3450	1.0	+0.00
HD 45282	G0	5280	3.1	− 1.52
HD 78479	K3III	4509	2.5	+0.57
HD 79028	F9V	5881	4.2	− 0.08
HD 87141	F5V	6403	4.1	+0.04
HD 92588	K1IV	5044	3.6	− 0.10
HD 92839	CII	2847		+0.10
HD 95849	K3II	4430	2.3	+0.30
HD 106516	F5V	6247	4.4	− 0.70
HD 109995	A0p	8262	3.5	− 1.99
HD 110281	K5	3950	0.2	− 1.56
HD 111721	G6V	4995	2.5	− 1.26
HD 116976	K1III	4550	2.0	+0.01
HD 120933	K5III	3820	1.5	+0.50
HD 122563	F8IV	4500	1.3	− 2.74
HD 126327	M7.5	3000	0.0	− 0.58
HD 144585	G5V	5831	4.0	+0.23
HD 148783	M6III	3250	0.2	+0.00
HD 158148	B5V			
HD 160933	F5V	5765	4.0	− 0.39
HD 165195	K3p	4450	1.1	− 2.24
HD 175545	K2III	4429	2.9	+0.29
HD 180928	K4III	4000	1.2	− 0.55
HD 187111	G8	4429	1.2	− 1.54
HD 218935	G8III	4819	2.5	− 0.31
HD 338529	B5	6100	3.6	− 2.41

Notes: Column (1): name of star. Column (2): spectra type. Column (3): effective temperature in K. Column (4): log surface gravity in cm g s^{-2}. Column (5): [Fe/H] in dex. All parameters are from Valdes et al. (2004).

We added Gaussian noise to each of these GC spectra to obtain a range of 20 spectra with S/N evenly spaced in logarithmic space with between 5 and 300 Å^{-1}. The effects of S/N on these five spectra are plotted in Fig. 10. As previously reported by Usher et al. (2012), at low S/Ns, template-based CaT measurements of low-metallicity GCs are biased to higher values, while the most metal-rich GC has its CaT strength biased to slightly smaller values. This effect is less than 0.2 Å for GCs with S/N > 20 Å^{-1} making it unimportant for this study as the median S/N of our spectra is 116 Å^{-1} and 127 of our 138 spectra have S/N > 20 Å^{-1}. However, it is important for extragalactic studies such as those of Usher et al. (2012) where most spectra have low S/N.

We repeated this S/N analysis for both masked and unmasked sky lines and for the Indo-US templates. Measurements with masked sky lines showed larger biases with S/N (by up to a factor of 2) and larger uncertainties at fixed S/N (by up to 70 per cent) compared to

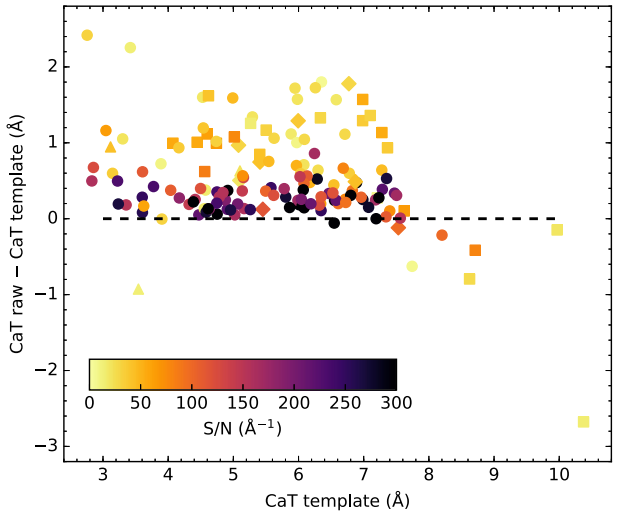

Figure 7. Differences between CaT measurement using templates and direct measurement using same normalization and index definition. The points are colour coded by S/N with MW GCs denoted by circles, LMC GCs by squares, SMC GCs by diamonds, and Fornax dSph GCs by triangles. At high S/N, the directly measured CaT strengths are offset by a constant 0.21 ± 0.03 Å, while at low S/N, the directly measured CaT values are stronger as the continuum normalization procedure clips off statistical fluctuations as if they were absorption lines.

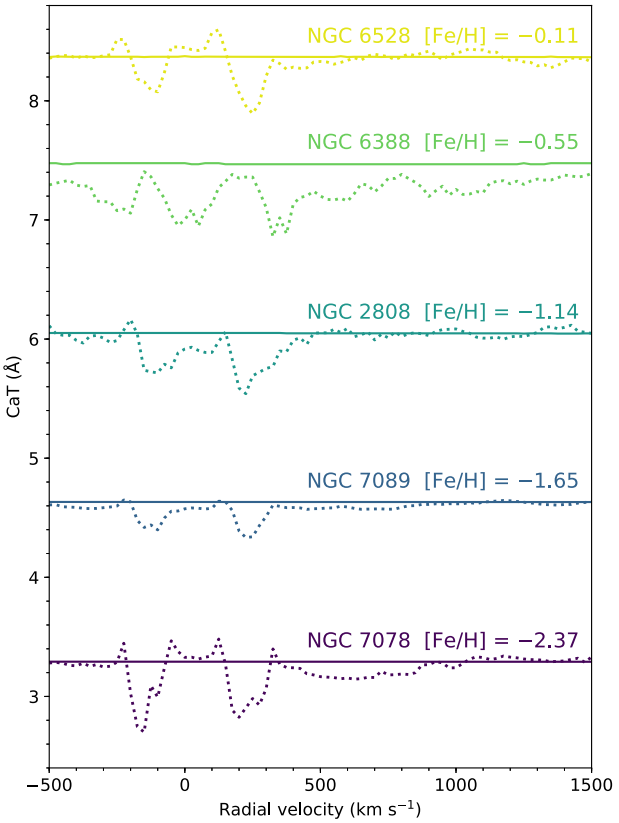

Figure 8. Effects of radial velocity on CaT measurement. The solid lines show the CaT strength of 5 GCs measured without masking sky emission lines when shifted to a range of radial velocities, while the dotted lines show the CaT strengths measured when masking the sky lines. The lines are colour coded by the metallicity of the GCs. Masking sky lines introduces significant biases in the CaT measurements at radial velocities such as ±200 km s^{-1} where the sky lines overlap with the CaT lines.

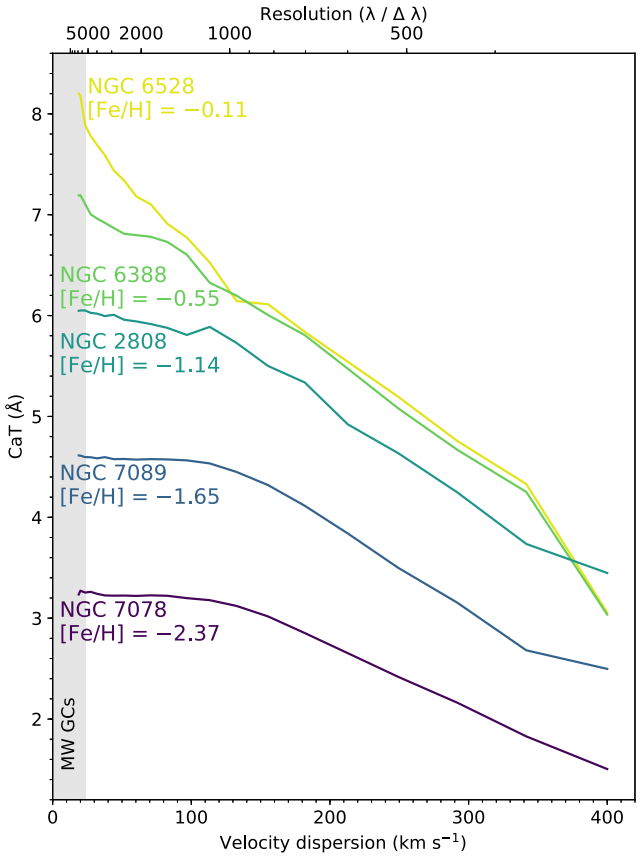

Figure 9. Effects of velocity dispersion (or equivalently spectral resolution) on CaT measurement. Like Fig. 8, the lines are colour coded by the metallicity of the GCs. Higher velocity dispersions lead to weaker CaT values with more metal-rich GCs being more strongly affected as well as being affected at lower velocity dispersions. Our WiFeS spectra have a spectral resolution of $R = 6800$ (19 km s^{-1}) in the region of the CaT, while literature central velocity dispersion measurements for MW GCs range from 0.6 to 23 km s^{-1} with a median of 6 km s^{-1} (Baumgardt & Hilker 2018). The lowest velocity dispersion plotted is the WiFeS resolution and the range of MW velocity dispersion is given by the shaded region.

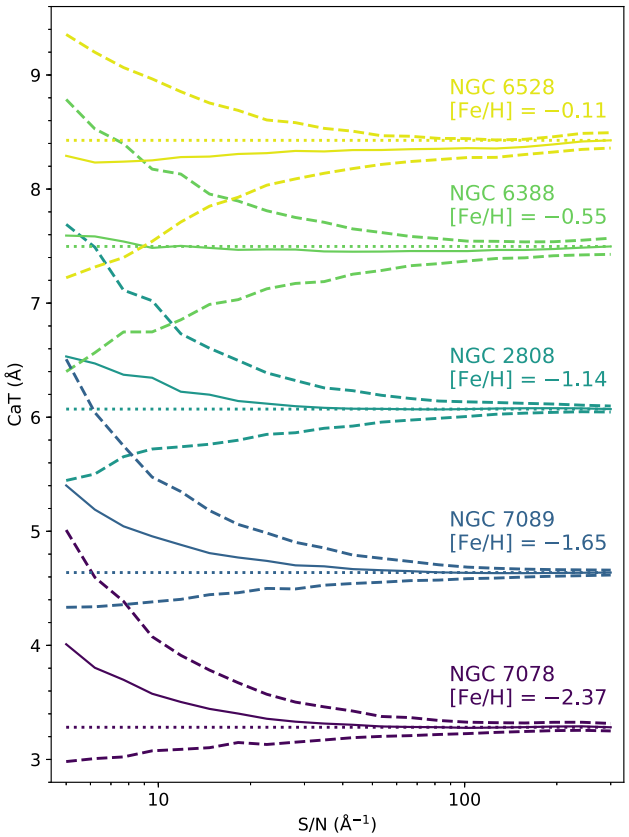

Figure 10. Effects of S/N on CaT measurement. For each GC, the solid line is the mean CaT value measured as a function of S/N, the dashed lines give the 68 percentiles of the measured CaT, and the dotted line gives the CaT strength measured on the original spectra. Like Fig. 8, the lines are colour coded by the metallicity of the GCs. At low S/N, measurements of low metallicity GCs are biased towards higher values, while the most metal-rich GCs are biased towards lower metallicities. At fixed S/N, measurements of higher metallicity GCs have higher uncertainties. As the median S/N of our sample is 116 Å^{-1}, this S/N bias does not have a major effect on our measurements.

the unmasked CaT measurements. Using the Indo-US templates also results in larger biases and uncertainties although to a lesser extent than the effects of masking sky lines. Given the biases introduced by masking sky lines due to their overlap with the CaT lines at some radial velocities and the larger S/N-based biases and uncertainties with both masking sky lines and the Indo-US templates, we use the unmasked, DEIMOS template-based CaT measurements in our analysis. We give our unmasked, DEIMOS template-based CaT measurements in Table 5.

3.3 Classical indices

We measured the CaT strengths on the observed spectra using a number of classical index definitions from the literature. We used the Armandroff & Zinn (1988) definition (CaT A&Z), the Cenarro et al. (2001) definition (CaT Cenarro), the Cenarro et al. (2001) definition with the correction for Paschen line absorption (CaT* Cenarro), and the Pastorello et al. (2014) SKiMS definition. Besides the CaT indices, we also measured the Cenarro et al. (2001) PaT Paschen line index. For each index, we adopted each author's index and continuum definitions and calculated the index strength and

Table 5. CaT measurements.

Name	RV (km s^{-1})	CaT (Å)	[Ca/H] (dex)	Notes
(1)	(2)	(3)	(4)	(5)
NGC 104	$-21.4^{+0.8}_{-0.6}$	$6.90^{+0.11}_{-0.10}$	$-0.55^{+0.04}_{-0.04}$	
Kron 3	$133.3^{+0.0}_{-1.8}$	$5.99^{+0.17}_{-0.22}$	$-0.91^{+0.07}_{-0.08}$	
NGC 121	$139.4^{+0.5}_{-0.6}$	$5.45^{+0.34}_{-0.29}$	$-1.12^{+0.14}_{-0.11}$	
NGC 288	$-44.9^{+0.8}_{-0.9}$	$4.53^{+0.19}_{-0.14}$	$-1.49^{+0.08}_{-0.06}$	LM
NGC 330	$157.6^{+4.5}_{-5.3}$	$6.77^{+0.43}_{-0.21}$	$-0.60^{+0.17}_{-0.08}$	Y
...	...	...	...	...

Notes: Column (1): GC name. Column (2): radial velocity in km s^{-1} measured using DEIMOS templates on the unmasked spectra. Column (3): CaT strength in Å measured using DEIMOS templates on the unmasked spectra. For both the radial velocity and the CaT strength, the uncertainty has been estimated using a block bootstrap of the datacubes. Column (4): [Ca/H] abundance in dex measured from the CaT strength using equation (4). Column (5): notes on [Ca/H] measurement. 'LM' indicates that the [Ca/H] measurement should be considered unreliable due to the low ($<5 \times 10^3$ M$_\odot$) mass within the observed field of view, while 'Y' indicates that the [Ca/H] measurement should be considered less reliable since the GC is younger than 3 Gyr. The full version of this table is provided in a machine readable form in the online Supporting Information.

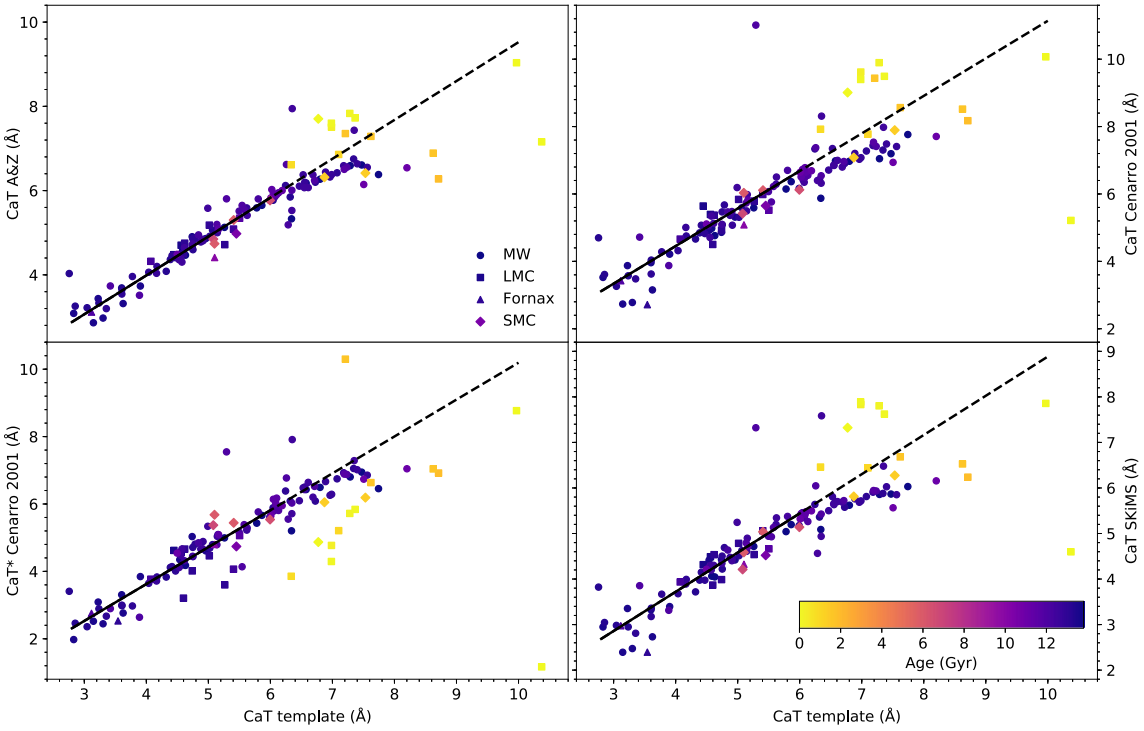

Figure 11. Comparison of classical CaT index measurements and the template-based CaT measurements. We compare our template-based CaT measurement with the CaT strength measured using the index definition of Armandroff & Zinn (1988, top left), the CaT index of Cenarro et al. (2001,top right), the CaT* index of Cenarro et al. (2001,bottom left), and the CaT SKiMS index used by Foster et al. (2009) and Pastorello et al. (2014, bottom right). As in Fig. 5, in each panel, the points are colour coded by age with MW GCs denoted by circles, LMC GCs by squares, SMC GCs by diamonds, and Fornax dSph GCs by triangles. In each panel, the black line is a fit between the classical index and the template-based index for old GCs with template-based CaT strengths of less than 6 Å with a solid line for the range of CaT values fitted and a dashed line for the extrapolation to higher CaT values. While each of the classical CaT indices follows the behaviour of the template-based CaT at low CaT strengths, at high CaT strengths old GCs show weaker classical CaT strengths relative to the trend at lower CaT strengths. Younger GCs have relatively stronger classical CaT values compared to template-based strengths for all indices except the CaT* one.

associated error using the formula provided in the appendix of Cenarro et al. (2001). We plot our classical CaT index measurements in Fig. 11.

At low to intermediate CaT strengths (<6 Å), all four of the classical indices show linear relationships with the template-based CaT measurements. However, at high CaT strengths all of the classical CaT measurements are lower than extrapolations of the linear relationship between the classical indices and the template. This is due to the stronger effects of line blanketing on the classical indices continuum passbands. We note that the feature passbands are the same for the Armandroff & Zinn (1988) as for our template-based measurements so much of the difference in behaviour is due to the differences in the way the continuum is determined. Most younger GCs have relatively stronger classical indices compared to template-based indices except for the CaT* index of Cenarro et al. (2001) which shows weaker indices. This is mostly due to the effects of the stronger Paschen line absorption at young ages on the feature passbands. The CaT* index was designed to correct for Paschen line absorption but it appears to over correct the absorption from these H lines. Unlike the template fitting method that can handle bad pixels such as those caused by poor sky line subtraction, all of the classical indices can be easily affected.

We performed the same tests on the effects of radial velocity, velocity dispersion, and S/N on the classical CaT indices as we did for our template-based CaT measurements. Unsurprisingly, measuring classical indices on spectra that have not been shifted to the rest frame can lead to catastrophically wrong index measurements. As shown by previous authors (e.g. Cenarro et al. 2001), higher velocity dispersions lead to weaker classical CaT indices and less sensitivity to metallicity although the broader Cenarro et al. (2001) indices show less dependence on the velocity dispersion (or spectral resolution). Unlike the template-based measurements, the classical indices show no bias with S/N but show larger uncertainties at fixed S/N by at least 40 per cent compared to the DEIMOS template-based CaT measurements.

The only previous study of the CaT strength of the integrated light of MW GCs is Armandroff & Zinn (1988). We smoothed our spectra to match their spectral resolution of $R = 2000$ and measured the index according to their definition on the smoothed spectra. In Fig. 12, we show the comparison of the CaT measurements by Armandroff & Zinn (1988) with our measurements using the same index definition on the smoothed spectra. We note that while the agreement is generally good, the rms difference between the two studies (0.3 Å) is significantly larger than the mean uncertainty in the differences between the studies (0.1 Å) indicating that the uncertainties of one or both of the studies are likely underestimated. These underestimated uncertainties are likely at least partially due to stochastic effects as the drift scanning by Armandroff & Zinn (1988) and our WiFeS observations likely covered different radial ranges (Armandroff & Zinn 1988 do not specify how far they drift scanned nor over what range of their long-slit observations they extracted their spectra other than saying that their scans covered at least the core radius).

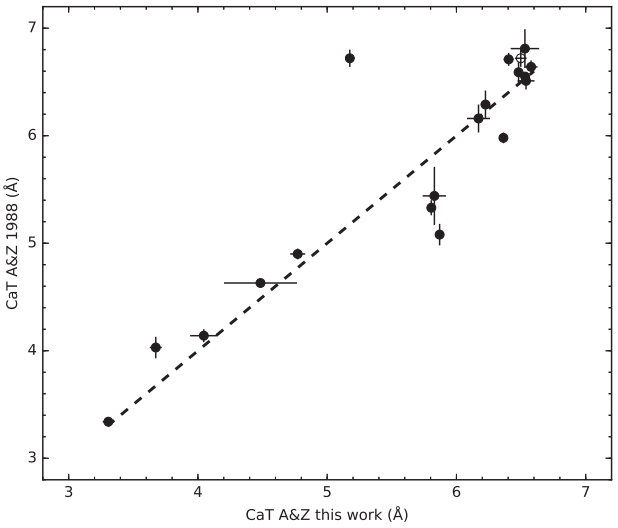

Figure 12. Comparison between our CaT measurements and those of Armandroff & Zinn (1988) using the Armandroff & Zinn (1988) index definition. For this measurement, we have smoothed our spectra to match that of Armandroff & Zinn (1988) ($R = 2000$). The empty point is the measurement of NGC 5927 with the long-period variable NGC 5927 V3 removed from the extracted spectra by masking it in the datacube. The dashed line is the 1-to-1 line. While there is general agreement between the two studies, the difference between the two is larger than the mean uncertainties.

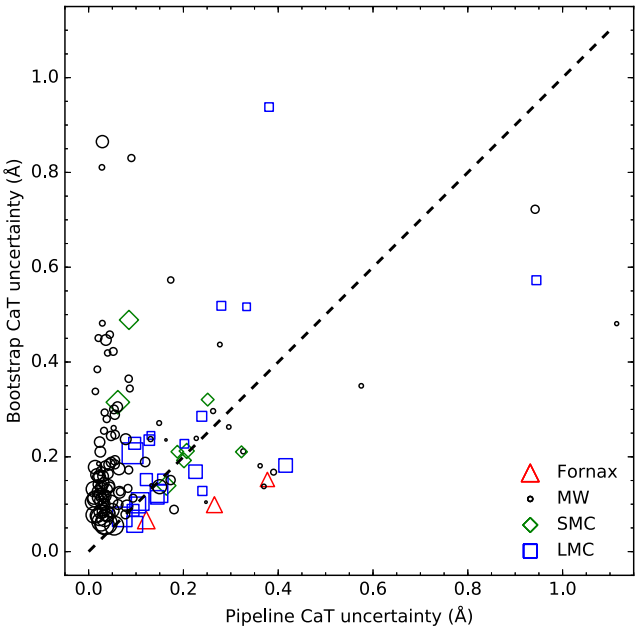

Figure 13. Comparison of the CaT strength uncertainties derived using the block bootstrap technique with the uncertainties calculated from the uncertainties provided by the data reduction pipeline. MW GCs are shown as black circles, LMC GCs as blue squares, SMC GCs as green diamonds, and Fornax GCs as red triangles. The size of the points is proportional to the mass enclosed within the field of view. Most of our GCs show significantly larger bootstrap uncertainties compared to the pipeline based uncertainties.

3.4 Uncertainties and repeated measurements

As in Usher et al. (2012), we used a Monte Carlo technique to estimate the uncertainties our CaT measurements. Using the uncertainties provided by the PyWiFeS pipeline and the best-fitting combination of templates, we created 1024 realizations of the spectra and performed the template fitting, continuum normalization and index measurement procedures on each. This is the same process as in Usher et al. (2012) although we use a factor of 10 more realizations resulting in a factor of 3 improvement in the precision of the uncertainty estimate.

We also used a block bootstrap technique to estimate the uncertainties in our CaT measurements. We spatially divided each datacube into 48 blocks and extracted a spectrum for each of the blocks. We then drew 1024 samples with replacement from these 48 block spectra. For each sample, we summed the sampled block spectra together and performed the template fitting, continuum normalization, and index measurement procedures on each. We then used the distribution of sample measurements to calculate the uncertainties.

A comparison of the uncertainties provided by these two methods is given in Fig. 13. In general, the bootstrap uncertainties are larger than those calculated using the Monte Carlo technique from the pipeline provided flux uncertainties at high S/N, but smaller at low S/N (the pipeline uncertainties strongly correlate with the inverse of the S/N). This effect is stronger for GCs with lower stellar masses in the field of view. A challenge to studying the CaT in the integrated light of GCs in the MW is that the CaT strength of individual stars depends strongly on surface gravity and hence luminosity. Since the brightest giants have CaT strengths significantly stronger than dwarfs, stochastic variations in the number of bright giants within the observed field of view can noticeably change the measured CaT strength. This significant stochastic variation within the datacubes was already noted in Usher et al. (2017, see their fig. 15). The bootstrap uncertainties include this statistical effect, with observations of GCs such as NGC 5927 whose integrated light are dominated by a small number of bright giants showing the largest bootstrap uncertainties. However, for observations such as some fields in NGC 6121 and NGC 6397 where there are no bright giants in the observed field of view, the bootstrap uncertainties underestimate the true uncertainties on the measurement.

Another way of estimating the uncertainties of our measurements is by comparing CaT indices in repeat spectra that are available for some of the GCs in our sample. Our repeated observations fall into two classes. First, we have cases where there is substantial spatial overlap between the repeated observations of the same GCs. Eight GCs (NGC 1466, NGC 1846, NGC 2210, NGC 2808, NGC 6284, NGC 6304, NGC 6342, and NGC 6717) were observed twice and two (NGC 104 and NGC 7099) were observed three times in this manner. Second, we have cases where the luminosity surface density is low and we deliberately observed multiple pointings of the same GC to obtain a more representative spectrum. We note that the two different classes represent different regimes of enclosed mass since the first class has a median observed mass of 2.7×10^4 $M_\odot$, while the second class has a median observed mass of 800 $M_\odot$.

We show the difference of the repeated measurements in the first class in Fig. 14. We find a median absolute difference of 0.12 Å between repeated measurements and median absolute difference of 0.10 Å for those measurements with at least an S/N of 25 $Å^{-1}$. We find a χ^2 value of 57.2 for the 14 repeated measurements using the pipeline-based uncertainties and a χ^2 value of 25.2 using the bootstrap uncertainties. These χ^2 values reduce to 53.2 and 13.1 for the 12 repeated measurements with S/N > 25 $Å^{-1}$. The difference in repeated CaT measurements is consistent with the bootstrap-based uncertainties but is underestimated by a factor of 2 by the pipeline-based uncertainties. For subsequent analysis of the GCs in

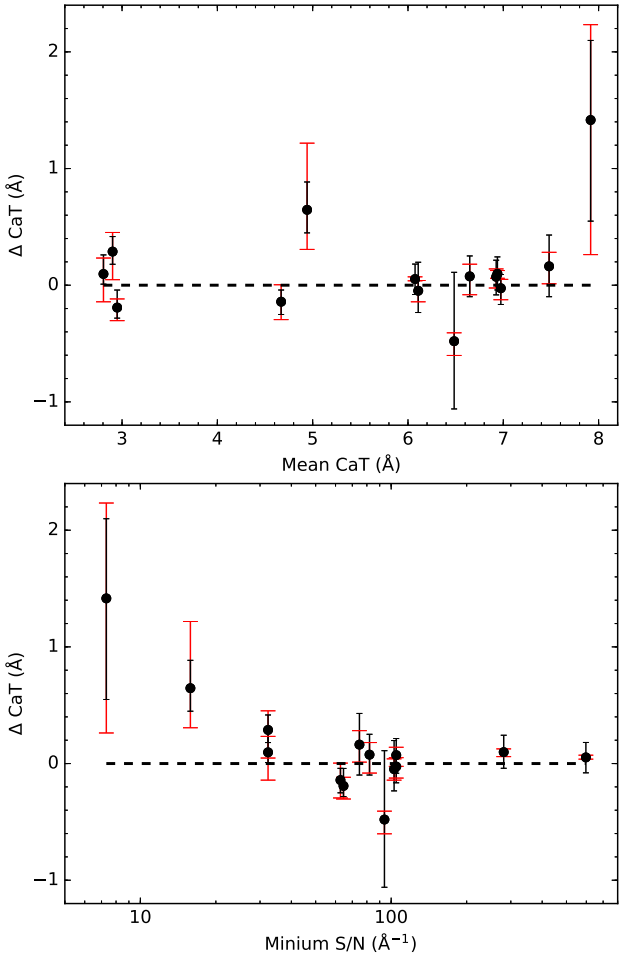

Figure 14. Differences between repeated CaT measurements for observations with substantial spatial overlap. In the top panel, we show the difference as a function of the average CaT measurement of the two observations, while in the bottom panel, we show the difference as a function of the lower S/N of the two observations. The red error bars were calculated from the uncertainties provided by the data reduction pipeline, while the black error bars were calculated using the block bootstrap technique. The pipeline uncertainties generally underestimate the differences between repeated measurements, while the bootstrap uncertainties are consistent.

the first class, we combined the repeated observations using an S/N weighted sum.

For the second class, we summed spectra of the individual pointings together with equal weights before performing our full measurement process on the summed spectra. In Fig. 15, we compare the CaT values of individual pointings with the measurements of the summed spectra. In most GCs, there are significant differences between the pointings due to the small number of giant stars in each pointing. The case of NGC 6121 is instructive as an example of this. By using catalogues from the *HST* ACS Globular Cluster Treasury Survey (Anderson et al. 2008), we identified what stars contribute to our datacubes. The NGC 6121_FIELD1 pointing is dominated by a single bright ($M_{\rm I} \sim -1.3$) RGB star and has a CaT measurement of $5.63^{+0.25}_{-0.23}$ Å, while the NGC 6121_FIELD2 pointing, which has no stars brighter than $M_{\rm I} \sim 2.1$ (corresponding to the base of the RGB), has CaT strength of $4.55^{+0.16}_{-0.12}$ Å. We note that these two pointings each contain stellar masses of about 750 $M_\odot$, whereas our median pointing covers 2.5×10^4 $M_\odot$.

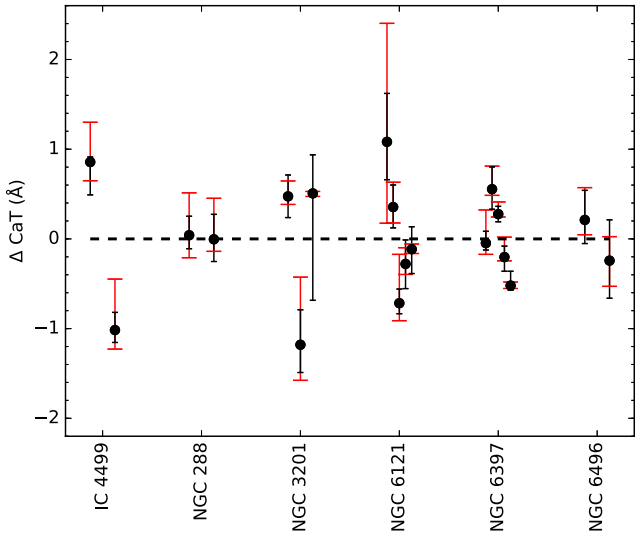

Figure 15. Differences of CaT measurements between individual pointings and the combination of all pointings for GCs with multiple spatially distinct pointings. The red error bars were calculated from the uncertainties provided by the data reduction pipeline, while the black error bars were calculated using the block bootstrap technique. Comparisons of different pointings of the same GC have been offset in the *x*-axis for legibility. In most cases, the differences between pointings of the same GC are larger than the uncertainties. These significant CaT strength spreads are due to stochastic variations in the number of bright giants in the WiFeS field.

Our observations of NGC 5927 represent an extreme case. We measure a CaT strength of $6.29^{+1.39}_{-0.34}$ Å from our single pointing which is lower but not significantly so than other GCs with similar metallicities ([Fe/H] $= -0.49$ in the Harris catalogue). Inspection of our NGC 5927 datacube reveals that our integrated spectrum is dominated by the long-period variable NGC 5927 V3. We observed NGC 5927 just after the maximum of NGC 5927 V3 (Sloan et al. 2010). The flux from within 2.5 pixels of NGC 5927 V3 contributes ∼25 per cent of the integrated flux of the datacube in the spectral region of the CaT, while if the datacube was spatially uniform, only 2.3 per cent of the flux would be within 2.5 pixels. Our spectrum of NGC 5927 V3 reveals the strong TiO absorption bands and insignificant CaT absorption characteristic of a very cool giant. The large bootstrap uncertainties of NGC 5927 show the large effect this single star has on our observations. Re-extracting the spectrum of NGC 5927 while masking spaxels within 2.5 arcsec of NGC 5927 V3, we find a CaT strength consistent with GCs with similar metallicities ($7.52^{+0.17}_{-0.19}$ Å). This large effect of a single star still occurs in the unmasked cube despite there being 4.6×10^4 $M_\odot$ of stellar mass within the WiFeS field of view. Stochastic sampling of the IMF appears to be the dominant source of uncertainty for most of our observations. We return to the impact of stochastic effects in Section 4.4.

3.5 Kinematics

As part of our template fitting process, the radial velocity and velocity dispersion of each spectra are fit. We give our radial velocity measurements in Table 5. For our repeated measurements, we find an rms difference of 5.2 km s^{-1} which is significantly larger than either the pipeline- or bootstrap-based uncertainties (both ∼1 km s^{-1}). We find similar sized differences in radial velocities of the same GC measured from different WiFeS gratings (Dalgleish et al.

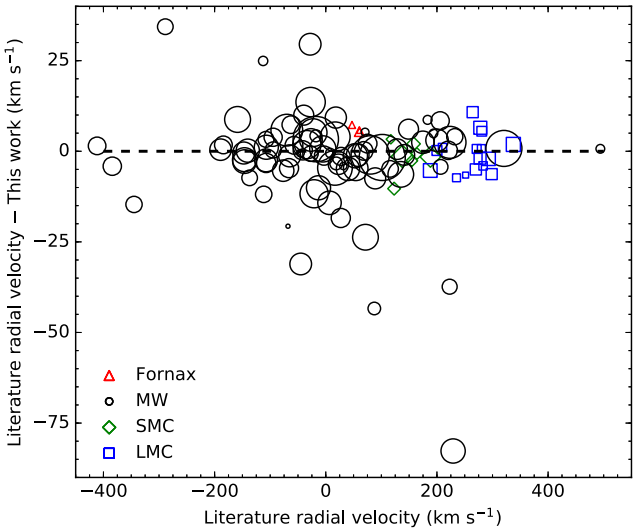

Figure 16. Comparison between our mean radial velocity and literature measurements. The size of the points is inversely proportional to the S/N, and as in Fig. 13, MW GCs are shown as black circles, LMC GCs as blue squares, SMC GCs as green diamonds, and Fornax GCs as red triangles. While differences between our measurements and literature measurements are generally small (a median absolute difference of 4 km s^{-1}), there are a number of MW GCs with significant discrepancies.

in preparation). The origin of these systematic differences is currently under investigation. As the radial velocity is a nuisance parameter in our analysis, these systematic differences have no effect on our index measurements.

We compare our radial velocity measurements with literature in Fig. 16. We are unaware of literature radial velocity measurements of NGC 411, NGC 1850, NGC 1856, and NGC 2004. In general our measurements are in good agreement with literature values (the median of the absolute velocity differences is 4 km s^{-1}), but for a small number of MW GCs we find significant differences (>25 km s^{-1}) between our measurements and those in the Harris (1996, 2010) catalogue. As also noted by Baumgardt & Hilker (2018), several of the radial velocity measurements in the Harris catalogue are based on low-resolution spectroscopy of small numbers of stars. Of the GCs with larger than a 15 km s^{-1} difference between our work and the Harris catalogue, two of our measurements agree with the values from high-resolution spectroscopy (NGC 5634, Sbordone et al. 2015 and NGC 6287, Lee & Carney 2002), one is consistent within uncertainties (Pal 11) and the remainder (NGC 6235, NGC 6316, NGC 6333, NGC 6356, NGC 6496, NGC 6569, and NGC 6584) have values in the Harris catalogue based on older, low-resolution work (e.g. Webbink 1981; Zinn & West 1984; Hesser, Shawl & Meyer 1986). We note that for these GCs our measurements are in general agreement with the recent low-resolution measurements of Saviane et al. (2012), Dias et al. (2016), and Vásquez et al. (2018). We find a median absolute difference of 3 km s^{-1} between our measurements and those of Baumgardt & Hilker (2018). The two GCs with more than a 15 km s^{-1} difference between our work and Baumgardt & Hilker (2018) either are consistent within uncertainties (Pal 11) or have a velocity based on only 13 stars in Baumgardt & Hilker (2018, NGC 6642).

Our velocity dispersion measurements are in general agreement with literature values (mostly from the Harris 1996, 2010 catalogue) for velocity dispersions greater than ∼5 km s^{-1}. A detailed study of the systematic uncertainties affecting our kinematics measurements is beyond the scope of this work and will be considered in future work.

4 ANALYSIS

In Fig. 17, we plot our CaT measurements of all 113 GCs as a function of both age and metallicity. Example spectra ordered by metallicity may be seen in Fig. 2 and ordered by age in Fig. 3. GCs 6 Gyr and older follow a roughly linear relationship between our template-based CaT strength and metallicity in line with previous studies (Armandroff & Zinn 1988; Usher et al. 2012; Sakari & Wallerstein 2016). GCs younger than 2 Gyr, however, show more complicated behaviour. GCs with ages of 1.4–2 Gyr in the LMC and SMC show a range of CaT strengths with some showing elevated CaT strengths compared to old MW GCs with the same metallicity, while other GCs show similar CaT strengths. GCs in the age range 100 Myr–1.1 Gyr also show a range of CaT strengths with some (but not all) GCs showing significantly weaker CaT strengths than GCs with the same metallicity. However, the youngest GCs (∼20 Myr) in our sample show the strongest CaT strengths of any GC. Before discussing the effects of age on the CaT in more detail, we will discuss the effects of metallicity and Ca abundance on the CaT.

4.1 An empirical CaT–metallicity relationship

Due to the lack of strong spectral features sensitive to their abundances in the optical, it is challenging (e.g. O) or impossible (e.g. Ne) to measure from stellar spectra the abundances of all elements that significantly contribute to the total metallicity Z. Thus, the abundance of a single element is often used as a proxy for the total metallicity. Typically, the Fe abundance ([Fe/H]) is used as a proxy for stellar metallicity but [Fe/H] is only a direct proxy for metallicity for populations with the same abundance pattern.

Given the common use of [Fe/H] to measure metallicity, we first compare our CaT measurements with our three samples of literature [Fe/H] measurements in Fig. 18. As mentioned before, many of the GCs younger than 3 Gyr show a larger range in CaT strengths than GCs with the same [Fe/H]. We defer a detailed discussion of the effects of age on the CaT to Section 4.2 and focus only on GCs older than 10 Gyr in this subsection. Unsurprisingly, there is less scatter in the CaT–[Fe/H] relation when using high-resolution abundances.

We used the Markov Chain Monte Carlo (MCMC) code PYMC3 (Salvatier, Wiecki & Fonnesbeck 2016) to fit [Fe/H] as a linear function of CaT strength accounting for the uncertainty in both variables. We used weakly informative Gaussian priors on both the slope and the intercept ($\mu = 0$ and $\sigma = 10$ for both). We performed fits for both our sample of high-resolution [Fe/H] measurements

$$[\mathrm{Fe/H}]_{\mathrm{High\ Res}} = 0.438^{+0.007}_{-0.007}\mathrm{CaT} - 3.696^{+0.037}_{-0.040} \quad (1)$$

and for the sample of [Fe/H] measurements by Carretta et al.

$$[\mathrm{Fe/H}]_{\mathrm{Carretta}} = 0.446^{+0.013}_{-0.013}\mathrm{CaT} - 3.720^{+0.080}_{-0.072}. \quad (2)$$

To reduce stochastic effects (see Section 4.4), we restricted our fits to GCs with stellar masses greater than 5×10^3 M$_\odot$ within the field of view. This gives us 54 GCs in our high-resolution sample and 18 GCs in the Carretta sample. Using the high observed mass, high-resolution sample, we find an rms difference of 0.14 dex between the literature [Fe/H] and those calculated using equation (1) and a χ^2 value of 131.5 for 52 degrees of freedom; for the high observed mass Carretta sample, we find an rms difference of 0.12 dex and a χ^2 value of 36.7 for 16 degrees of freedom. The χ^2 value is

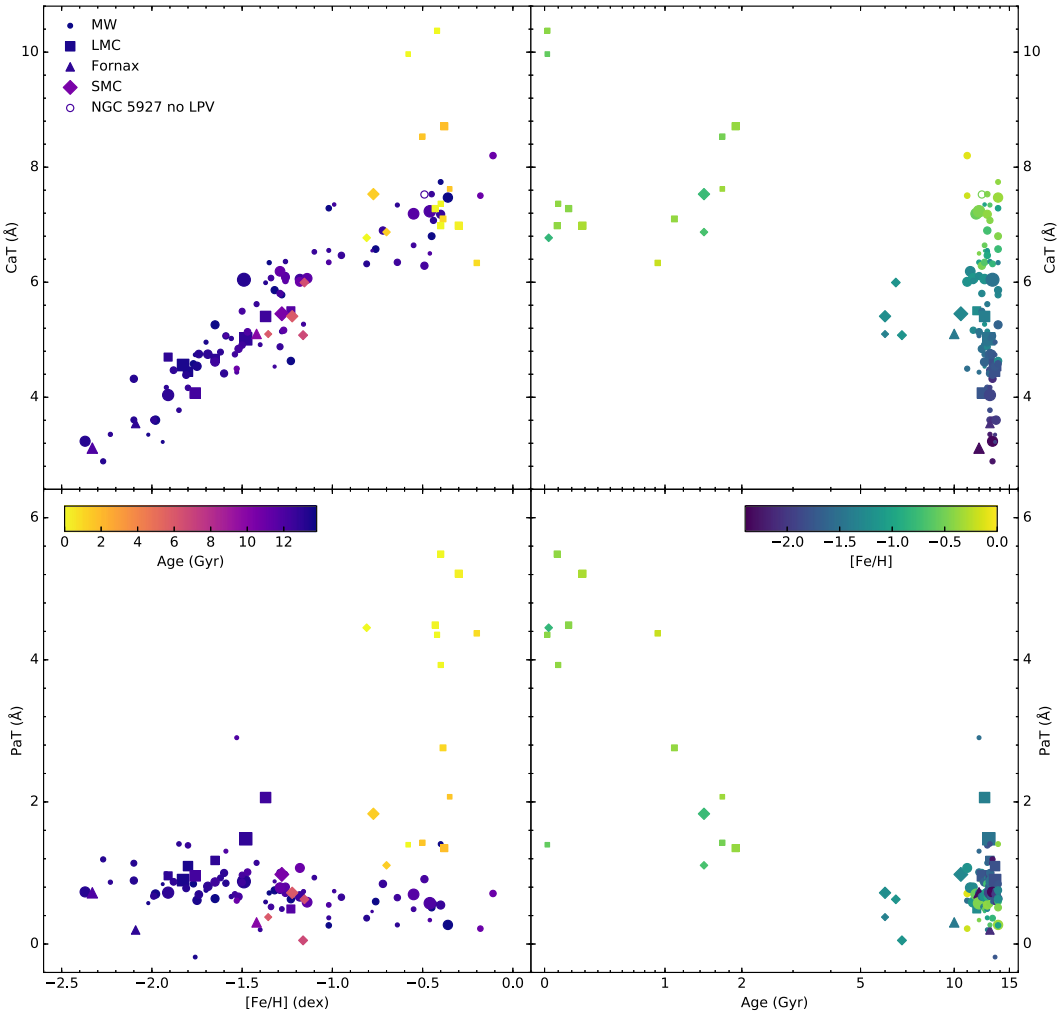

Figure 17. Top left: CaT strength as a function of metallicity. The points show our measurements colour coded by age with sizes proportional to the mass enclosed within the WiFeS field of view. Top right: CaT strength as a function of age. The points and models are colour coded by metallicity. Bottom left: strength of the Paschen lines using the index definition of Cenarro et al. (2001) as a function of metallicity. Bottom right: strength of the Paschen lines as a function of age. In both panels, the empty circle is the CaT measurement of NGC 5927 with the long-period variable NGC 5927 V3 removed from the extracted spectra. CaT strength increases with metallicity and is largely independent of age apart from at ages younger than 2 Gyr. PaT strength decreases with increasing age. The PaT indices of old GCs slightly decrease with increasing metallicity. The PaT indices of ∼6 Gyr GCs in the SMC are similar to old GCs with the same metallicities.

calculated as

$$\chi^2 = \sum \frac{([\mathrm{Fe/H}]_{\mathrm{Calc.}} - [\mathrm{Fe/H}]_{\mathrm{Lit.}})^2}{\sigma^2_{\mathrm{Calc.}} + \sigma^2_{\mathrm{Lit.}}} \quad (3)$$

and the degrees of freedom as the difference between the number of GCs and the number of model parameters (2 for a linear fit and 1 for a constant value). This compares to median uncertainties in our calculated [Fe/H] of 0.05 dex and 0.07 dex for the high-resolution and Carretta samples respectively. The high χ^2 value for the high-resolution sample is partially driven by the unrealistically low uncertainties on some of the literature [Fe/H] measurements. Given the similarities between the two relations, we adopt the high-resolution sample for the rest of this work due to its lower uncertainty.

4.1.1 Effects of calcium abundance

Most but not all GCs in the MW and other galaxies show enhanced [Ca/Fe] abundance ratios (e.g. Fig. 1, Puzia et al. 2005; Colucci et al. 2013; Sakari et al. 2016). Since the CaT spectral feature is dominated by atomic Ca lines, we expect that CaT strength should depend on the Ca abundance as well as on the metallicity. Unlike the light α-elements (O, Ne, and Mg), which are primarily formed via hydrostatic burning in high-mass stars, but like the other heavy α-elements (Si, S, Ar, and Ti), Ca is mainly formed by explosive nucleosynthesis, mostly in core collapse supernovae but with a significant contribution from type Ia supernovae (e.g. Woosley & Weaver 1995; Wiersma et al. 2009; Hopkins et al. 2018). As such the Ca abundance should more closely trace the total metallicity than the Fe abundance but not as well as the O or Mg abundances. Stellar population synthesis modelling by Brodie et al. (2012) showed that the relationship between Z and CaT has little dependence on [α/Fe] (equivalent to the relationship between [Fe/H] and CaT strongly depending on [α/Fe]), while the models of Conroy & van Dokkum (2012) also predict a significant dependence on [Ca/Fe] at fixed [Fe/H]. Observations by Sakari & Wallerstein (2016) showed that GCs in M31 with low [Ca/Fe] have lower CaT strengths compared to GCs with the same [Fe/H] but higher [Ca/Fe]. We also note that

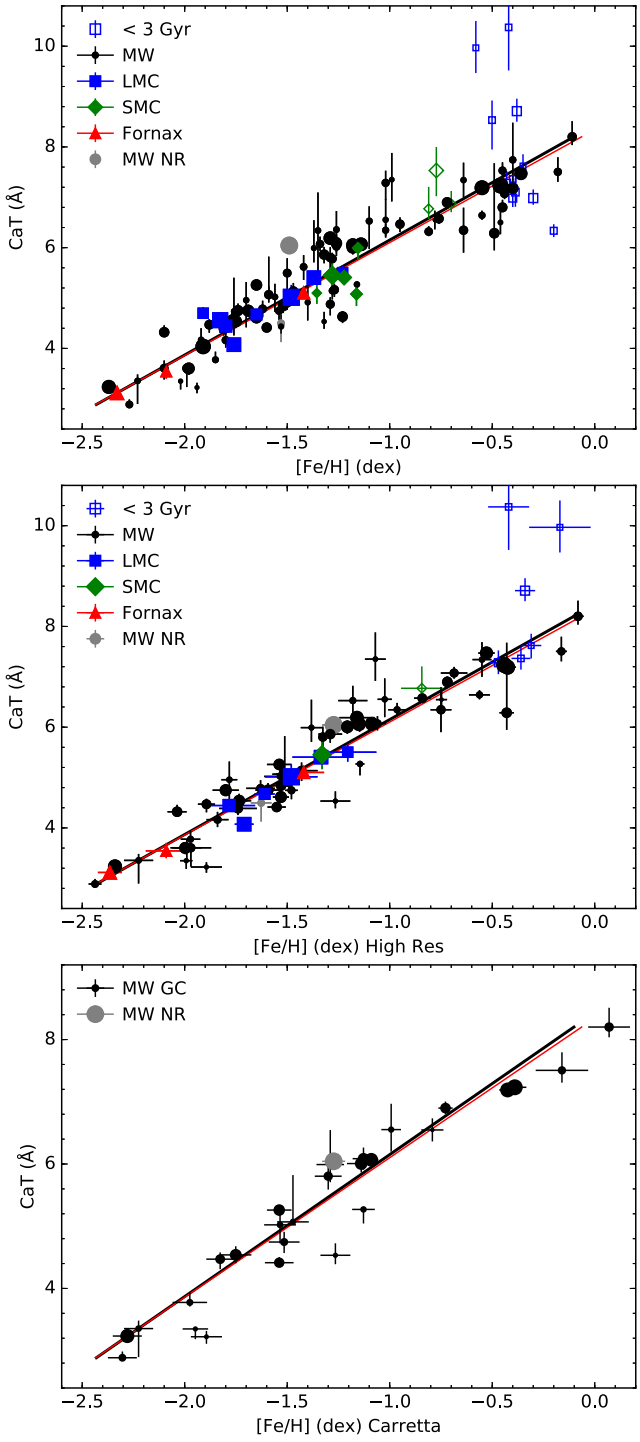

Figure 18. CaT–[Fe/H] relationship. Top: CaT strength versus all [Fe/H] measurements as in the upper left panel of Fig. 17. Middle: CaT strength as a function of [Fe/H] from high-resolution spectroscopy. Bottom: CaT strength versus [Fe/H] from the Carretta et al. group. In all three panels, black circles are MW GCs, grey circles are MW nuclear remnants, blue squares are LMC GCs, green diamonds are SMC GCs, and red triangles are Fornax dSph GCs, while empty points are GCs younger than 3 Gyr. The size of each point is proportional to the mass observed within the WiFeS field of view. The thick black solid line is the CaT–[Fe/H] relationship fit to the high-resolution [Fe/H] measurements of old GCs (equation 1), while the thin red solid line is the CaT–[Fe/H] fit to the Carretta et al. [Fe/H] measurements (equation 2).

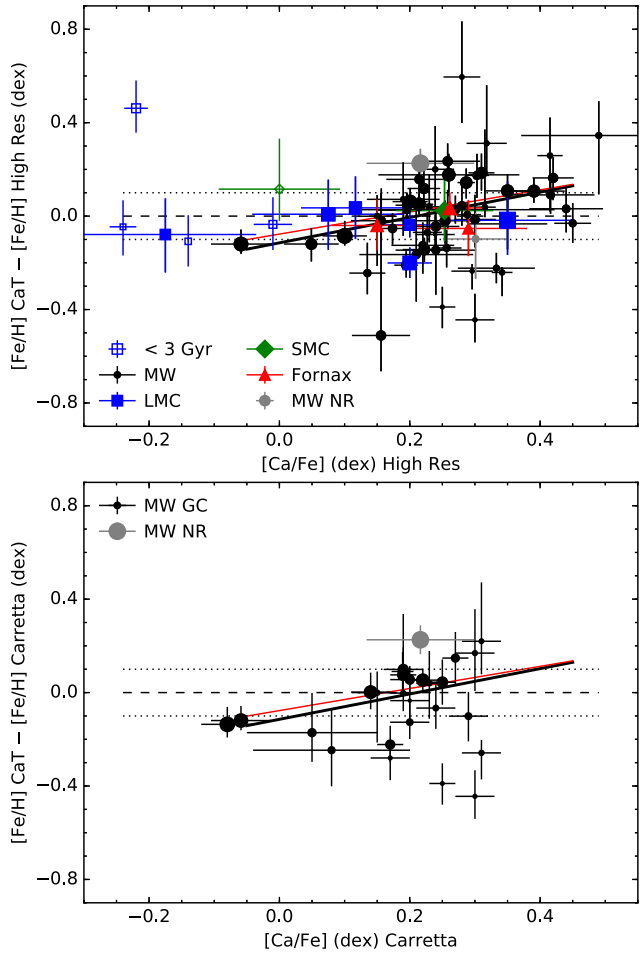

Figure 19. Comparison of the $[Fe/H]_{CaT} - [Fe/H]_{lit.}$ difference and [Ca/Fe]. Top: [Fe/H] as a function of [Ca/Fe] based on high-resolution spectroscopy. Bottom: [Fe/H] strength versus [Ca/Fe] based on abundances from the Carretta et al. group. The colours, shapes, and size of the points are the same as in Fig. 18. The thin red line is the Δ[Fe/H]–[Ca/Fe] relationship fit with the Carretta et al. abundances, while the thick black line is the Δ[Fe/H]–[Ca/Fe] fit to the high-resolution abundances of old GCs. The dashed line shows Δ[Fe/H] = 0 ,while the dotted lines show Δ[Fe/H] ± 0.1. GCs with lower [Ca/Fe] ratios have weaker CaT strengths than GCs with higher [Ca/Fe]. The majority of GCs with high [Ca/Fe] ratios but weaker than expected CaT strengths have low stellar masses within the observed field of view and hence less reliable CaT measurements due to stochastic effects.

Ca does not show abundance spreads within GCs unlike the spreads in the lighter α-process elements O and Mg observed in almost all GCs (e.g. Carretta et al. 2010b).

In Fig. 19, we show the difference between [Fe/H] measurements from high-resolution spectroscopy and the [Fe/H] calculated from our CaT triplet measurements as a function of the [Ca/Fe] ratio. Although there is a lot of scatter, we do see a trend in that GCs with lower [Ca/Fe] values have lower CaT-based [Fe/H] values and GCs with higher [Ca/Fe] have higher CaT based [Fe/H]. Using the high observed mass, high-resolution sample, Kendell's correlation test gives evidence for a significant correlation between Δ[Fe/H] and [Ca/Fe] (a correlation coefficient of $\tau = 0.30$ corresponding to a significance of $p = 0.001$), while the smaller, high-mass Carretta sample does not possess a significant correlation ($\tau = 0.25$ and $p = 0.16$).

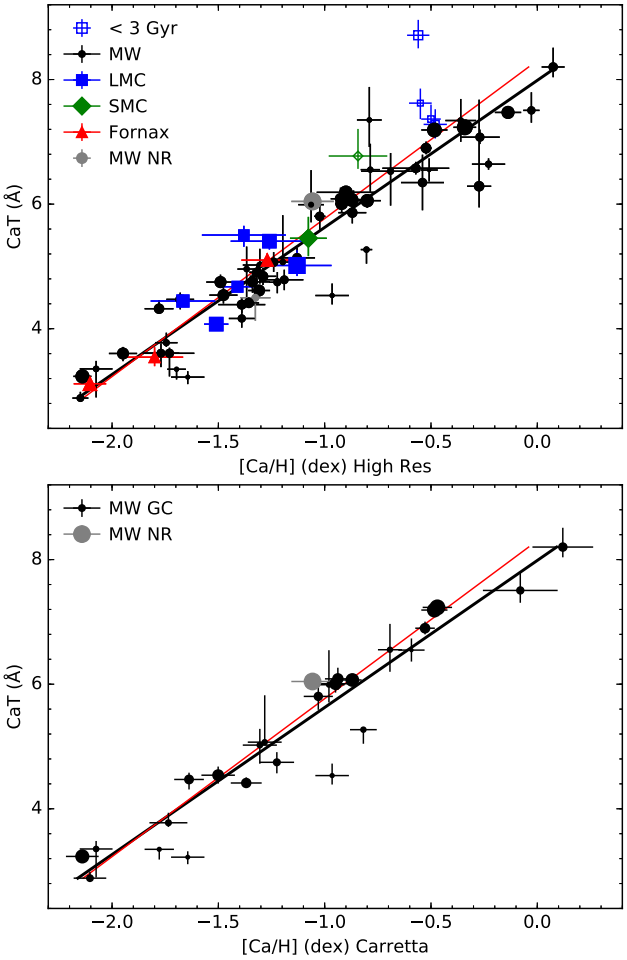

Figure 20. CaT–[Ca/H] relationship. Top: CaT strength as a function of [Ca/H] from high-resolution spectroscopy. Bottom: CaT strength versus [Ca/H] from the Carretta et al. group. The colours, shapes, and size of the points are the same as in Fig. 18. The thick black solid line is the CaT–[Ca/H] relationship fit to the high-resolution [Fe/H] measurements of old GCs (equation 4), while the thin red solid line is the CaT–[Ca/H] fit to the Carretta et al. [Fe/H] measurements (equation 5).

To further assess whether a relation exists between [Ca/Fe] and the CaT strength, we used PYMC3 to fit the difference between the CaT-based [Fe/H] and the literature [Fe/H] as a function of [Ca/Fe] for the high observed mass, high-resolution sample and for the high-mass Carretta sample. As before we used weakly informative Gaussian priors. We find significant slopes for both our high-resolution sample ($0.54^{+0.10}_{-0.13}$ from 52 GCs) and the Carretta sample ($0.47^{+0.19}_{-0.19}$ from 17 GCs). We plot these relations in Fig. 19. The Bayesian information criterion (BIC) strongly favours a linear relationship between the [Fe/H] difference and [Ca/Fe] with a difference in BIC of 32.1 for the high-mass, high-resolution sample and a difference of 5.8 for the high-mass Carretta sample. Both favour the linear relationship over a constant difference. We note that the relation between the residuals about the CaT–[Fe/H] fit and [Ca/Fe] is not driven by the decrease in [Ca/Fe] with increasing metallicity. Our sample shows a range of [Ca/Fe] at most metallicities (Fig. 1) and if the sample is split by metallicity, weak evidence for a relationship with [Ca/Fe] is seen at low, intermediate, and high metallicities. Our relationship between the difference in CaT-based [Fe/H] and high-resolution-based [Fe/H] and [Ca/Fe] is broadly similar to the relationship that Da Costa (2016) found for the CaT strengths of individual giant stars.

Given that the CaT strength depends on the Ca abundance, we performed fits using PYMC3 of [Ca/H] as a linear function of CaT strength for both our high-resolution sample

$$[\mathrm{Ca/H}]_{\mathrm{High\ Res}} = 0.434^{+0.008}_{-0.009}\mathrm{CaT} - 3.384^{+0.047}_{-0.044} \quad (4)$$

using the 52 high observed mass GCs with [Ca/H] measurements and for the Carretta sample

$$[\mathrm{Ca/H}]_{\mathrm{Carretta}} = 0.394^{+0.014}_{-0.013}\mathrm{CaT} - 3.274^{+0.080}_{-0.083} \quad (5)$$

using the 17 high observed mass Carretta GCs with [Ca/H] measurements. We plot these fits in Fig. 20. As for the [Fe/H]–CaT relationships, we use weakly informative Gaussian priors and only fit GCs with at least a mass of 5×10^3 $M_\odot$ in the observed field of view. Using the high observed mass, high-resolution sample, we find an rms difference of 0.14 dex between the literature [Ca/H] and those calculated using equation (1) with a χ^2 value of 94.9 for 50 degrees of freedom; for the high observed mass Carretta sample, we find an rms difference of 0.10 dex and a χ^2 value of 17.0 for 15 degrees of freedom. These compare with median uncertainties in [Ca/H] of 0.07 dex for both literature samples, respectively. The lower χ^2 values for [Ca/H] compared to [Fe/H] show that there is a stronger relationship between [Ca/H] and CaT strength than between [Fe/H] and CaT. This conclusion is further supported by the fact that [Ca/H] abundances are more uncertain than the [Fe/H] ones. Unsurprisingly, the BICs for the [Ca/H] fits are significantly smaller than for the [Fe/H] fits (differences of 15.3 and 12.6 for the high-resolution and Carretta samples, respectively). This is also supported by the slightly lower rms value for [Ca/H] for the Carretta sample although the high-resolution sample shows similar rms differences for [Fe/H] and [Ca/H]. Using bootstrapping, we can estimate the uncertainty on these differences. Drawing 1024 samples from the high observed mass GCs, the χ^2, BIC, and rms values are smaller for [Ca/H] relation than for the [Fe/H] relation 96.6, 85.6, and 54.4 per cent of the time for the high-resolution sample and 97.3, 96.4, and 82.3 per cent for the Carretta sample.

If we restrict our high-mass, high-resolution sample to the MW, we see larger differences between the [Ca/H] and [Fe/H] in rms (0.15 dex for [Fe/H] versus 0.14 dex for [Ca/H]), reduced χ^2 (a difference in χ^2 of 29.9 for 40 degrees of freedom) and BIC (a difference of 20.0) compared to the entire high-mass, high-resolution sample. The sample of 10 GCs in satellite galaxies is consistent with both the [Ca/H] and [Fe/H] relations but does not show a preferences for either relation. We note that the median uncertainties in literature [Ca/H] measurements in the satellite galaxies are twice as large in the satellite galaxy GCs as in the MW GCs (0.06 versus 0.13 dex). If we split the MW sample in half by [Fe/H], by the observed stellar mass or by the fraction of the GC observed, we see a preference for a relationship with [Ca/H] for each of the subsets at lower significance than for the entire MW sample. Given the similarities between the two relations, we adopt the high-resolution sample relation for the rest of this work due to its lower uncertainty. We give [Ca/H] values based on equation (4) for each of our GCs in Table 5.

The two nuclear remnant in our sample, NGC 5139 and NGC 6715, show significant metallicity spreads (>1 dex, e.g. Johnson & Pilachowski 2010; Carretta et al. 2010c). Both of these objects have CaT strengths in general agreement with their mean [Fe/H] and [Ca/H] abundances suggesting that the CaT can be used to measure the mean metallicities of stellar populations with significant metallicity spreads.

The CaT also depends on more than just the total metallicity and the Ca abundance. The spectral range we fit with stellar templates and use to determine the continuum contains atomic lines of a number of elements including Mg, Si, Al, Ti, and Fe as well as CN and TiO molecular bands. As such the abundances of all these elements can have effects our CaT strength measurements. We note that the effects of these other elements on the CaT spectral region are stronger at high metallicity than at low metallicity (see figs 22 and 23 of Conroy et al. 2018). Abundance variations may also have minor effects the CaT strength via the effects of changing opacity on stellar structure and evolution (e.g. Dotter et al. 2007; VandenBerg et al. 2012) or via the dependence of the singly ionized CaT on the electron pressure (e.g. Conroy & van Dokkum 2012). However, the effects of varying abundances on spectral indices is usually dominated by effects of varying line opacity within the index definition rather than via the effects of varying abundances on the isochrones (Schiavon 2007; Lee et al. 2009). We note that the dependence of the CaT on the abundance of other elements is a common problem for spectral indices (e.g. Korn, Maraston & Thomas 2005; Tripicco & Bell 1995; Lee et al. 2009; Conroy & van Dokkum 2012).

As noted earlier, the abundances of α-elements do not always vary in lockstep due to the different nucleosynthetic origins. For example, stars in dwarf galaxies and some stars in the MW halo follow different [Mg/Fe]–[Ca/Fe] relationships than the MW disc or bulge (e.g. Van der Swaelmen et al. 2013; McWilliam, Wallerstein & Mottini 2013; Lemasle et al. 2014; Hayes et al. 2018). However, any abundance effect on the CaT by the [Mg/Ca] ratio is likely small since the Fornax GCs in our sample, which show low [Mg/Ca] ratios (Letarte et al. 2006; Larsen et al. 2012) compared to MW or LMC GCs with similar [Ca/Fe] ratios (e.g. Mucciarelli et al. 2010; Carretta et al. 2010b), follow the same CaT–[Fe/H] and CaT–[Ca/H] relations as the MW or LMC GCs within uncertainties. Given the lack of a significant difference between CaT strengths of the Fornax GCs and MW GCs with the same [Fe/H] or [Ca/H] abundances, we expect the CaT strengths of extragalactic GCs to follow the same CaT–[Fe/H] and CaT–[Ca/H] relations as for our sample.

Finally, we note that all our GCs equal in age or older than NGC 1978 ($\sim$1.9 Gyr, Mucciarelli et al. 2007) that have been studied show evidence for multiple populations (Bastian & Lardo 2017; Martocchia et al. 2018) such that a significant fraction of stars in the GCs in our sample show enhanced abundances of He, N, Na, or Al and depleted C, O, or Mg. Changes in the abundances of these elements could affect the CaT via their effects on the HB morphology (He), on the pseudo-continuum (C, N, and O) or on electron pressure (Na, Mg, and Al). These abundance changes also implies the abundances of GCs in our sample is not the same as field stars of the same age, metallicity or α-element enhancement. We return to the effects of multiple populations in Section 4.3 where we study the effects of HB morphology on the CaT.

4.1.2 The CaT at high metallicity

The behaviour of the CaT at high metallicity has been debated. The relationship between CaT strength and [Fe/H] observed by Armandroff & Zinn (1988) flattens at high metallicity. The model predictions of Vazdekis et al. (2003) and Chung et al. (2016) reproduce this behaviour. Both authors explained this flattening as the effect of a cooler RGB at higher metallicity as CaT strength decreases rapidly with declining temperature for stars cooler than 3600 K (e.g. Cenarro et al. 2002). Vazdekis et al. (2003) and Chung et al. (2016) both explain this weakening of the CaT in cooler stars as due to the lower fraction of ionized Ca at lower temperatures. However, the lower temperatures and higher metallicities also produce stronger line opacity in the CaT spectral region. This lowers the flux in the pseudo-continuum passbands used to calculate the CaT index, weakening the measured CaT strength. As seen in the top left panel of Fig. 11, this effect is stronger for the classical indices observed by Armandroff & Zinn (1988) and modelled by Vazdekis et al. (2003) and Chung et al. (2016) than for our template-based CaT measurements where the pseudo-continuum is calculated by iteratively fitting a polynomial to wavelength regions with low line opacity.

On the basis of a non-linear colour–CaT relation for the GCs in NGC 1407, Foster et al. (2010) questioned whether the CaT saturates at high metallicity. However, the GC system of NGC 4494 shows a linear relationship between colour and the CaT (Foster et al. 2011). Usher et al. (2012) found a wide range of colour–CaT relationships for their galaxies but found good agreement between metallicities calculated from CaT strengths measured using the template fitting technique and a linear fit to the Vazdekis et al. (2003) stellar population models and metallicities calculated using Lick indices (Worthey et al. 1994; Worthey & Ottaviani 1997) and models such as those of Thomas, Maraston & Bender (2003) up to [Z/H] = 0.2 dex once the effects of [α/Fe] in the Cenarro et al. (2001) spectral library [used by the Vazdekis et al. (2003) models] were accounted for. Usher et al. (2015) using stacked spectra of extragalactic GCs found a linear relationship between CaT strength and the sum of weak Fe lines in the CaT region in the range $-1.7 <$ [Z/H] < -0.2 suggesting that the CaT strength is a reliable metallicity indicator up to near solar metallicities. In their study of M31 GCs, Sakari & Wallerstein (2016) found a clear linear relationship between CaT and [Fe/H] for GCs less metal rich than [Fe/H] $= -0.4$ but their single higher metallicity GC (B193-G244, [Fe/H] $= -0.16$) showed weaker than expected CaT absorption, though they note that they had difficulty determining the continuum for this GC.

Our study shows a clear, linear relationship between CaT strength and [Ca/H] up to solar [Ca/H]. The GCs more metal rich than [Fe/H] $= -0.5$ show weaker CaT strengths than expected for their [Fe/H] due to the effects of lower [Ca/Fe] ratios. The mean difference of the [Fe/H] abundances calculated from the CaT and literature [Fe/H] is -0.12 ± 0.05 dex for the four GCs with [Fe/H] $= -0.5$ (excluding NGC 5927 due to its large observational errors), while the difference in [Ca/H] is 0.00 ± 0.07 dex for the same GCs. The Carretta sample of abundances shows this more dramatically due to the higher [Fe/H] value for NGC 6528 ([Fe/H] = 0.07 for Carretta et al. 2001 versus [Fe/H] $= -0.08$ for average of literature studies) and the low [Ca/Fe] ratio for NGC 6388 and NGC 6441 (Carretta et al. 2007a). The two most metal-rich GCs in our study, NGC 6528 ([Fe/H] $= -0.11$; Harris 1996, 2010 catalogue) and NGC 6553 ([Fe/H] $= -0.18$; Harris 1996, 2010 catalogue), have CaT strengths consistent with our [Fe/H]–CaT and [Ca/H]–CaT relationships within uncertainties. The saturation of the CaT at high [Fe/H] in observations of Armandroff & Zinn (1988) and in the models of Vazdekis et al. (2003) and Chung et al. (2016) is likely due to the effects of line blanketing at high metallicity on the classical indices used as well as the effects of lower [Ca/Fe] ratios at high metallicity in the MW.

We note that GC stellar population studies in the MW at high metallicity are a challenge as only NGC 6528 and NGC 6553 have near solar metallicities and manageable amounts of foreground reddening. There is still considerable uncertainty in the abundances for these two GCs, with [Ca/Fe] measurements for NGC 6528 rang-

ing from [Ca/Fe] = −0.4 (Zoccali et al. 2004) to [Ca/Fe] = 0.34 (Origlia, Valenti & Rich 2005). Observing these bulge GCs is also a challenge due to the high field star surface density and reddening in the bulge. It is more difficult to reliably determine the continuum at solar metallicity than at lower metallicities and the CaT measurement is significantly more sensitive to the spectral resolution (or velocity dispersion) at high metallicity (Fig. 9). As is highlighted by the effect of a single cool giant in NGC 5927 (Section 3.4), stochastic effects are likely stronger at higher metallicities due to the higher temperature sensitive of the CaT at the lower temperatures reached by the most luminous metal-rich giants. Taken together with observations in M31 and in more distant galaxies, our observations show that the CaT can be used as indicator of [Ca/H] up to at least solar metallicity, although the reliability of the CaT declines at near solar metallicity and higher. The effects of Ca abundance limits the usefulness of the CaT in measuring [Fe/H] in populations with low [Ca/Fe] ratios such as those found at high metallicities.

4.2 Behaviour of the CaT with age

On the basis of our observations and stellar evolution calculations (e.g. Bressan et al. 2012; Choi et al. 2016), we can break the behaviour of the CaT into four age ranges. To estimate the contribution of different stellar evolutionary phases to the CaT wavelength region, we used the predictions of the MIST isochrones (Choi et al. 2016) with a metallicity of [Fe/H] = −0.4 (similar to the young and intermediate-age GCs in the LMC) for the *HST* WFC3 F845M filter and a Kroupa (2001) IMF. In the oldest age range, with GCs older than ∼3 Gyr, the CaT is essentially insensitive to age (top right panel of Figs 17 and 21). As seen in Fig. 17, no significant change in Paschen absorption as measured by the PaT index with age is seen in this age range. Here, the CaT spectral region is dominated by RGB stars with smaller contributions from the main sequence (MS), the core He-burning phase (CHeB – the red clump or HB) and the asymptotic giant branch (AGB). Since the effective temperatures and surface gravities of the post-MS stellar evolution phases vary little within this age range and the relative contributions of different stellar evolution phases is nearly constant with age in this range, the CaT has little or no age dependence for populations older than ∼3 Gyr.

The age range of ∼1.3 to ∼3 Gyr is one of transitions. The importance of the RGB declines as the CHeB comes to dominate the CaT spectral region while the contribution of the AGB, especially the thermally pulsing AGB, reaches its maximum in this age range (Girardi & Bertelli 1998; Girardi et al. 2013). Additionally, the temperature of the MS turnoff (MSTO) increases towards younger ages. Higher temperatures lead to weaker CaT absorption and stronger Paschen line absorption, especially above ∼6500 K (Cenarro et al. 2002), so as the contribution of MSTO stars hotter than this increases, the Paschen line strength of the integrated population increases and the CaT strength decreases. This increase in Paschen line absorption at younger ages can be seen via the increase of the PaT index in Fig. 17.

The SMC GCs NGC 411 and NGC 419 have similar ages and metallicities (1.45 Gyr, [Fe/H] = −0.7, Goudfrooij et al. 2014) but different CaT values ($6.87^{+0.25}_{-0.17}$ Å versus $7.53^{+0.46}_{-0.51}$ Å). Likewise the three LMC GCs NGC 1783, NGC 1846, and NGC 1978 have similar ages (1.7–2.0 Gyr, Goudfrooij et al. 2014; Mucciarelli et al. 2007) and metallicities ([Fe/H] ∼ − 0.4, Grocholski et al. 2006; Mucciarelli et al. 2008) but have dramatically different CaT strengths ($7.62^{+0.23}_{-0.26}$, $8.53^{+0.39}_{-0.58}$, and $8.71^{+0.25}_{-0.21}$ Å, respectively). Careful inspection of the spectra of these GCs (Fig. 22) reveals that the

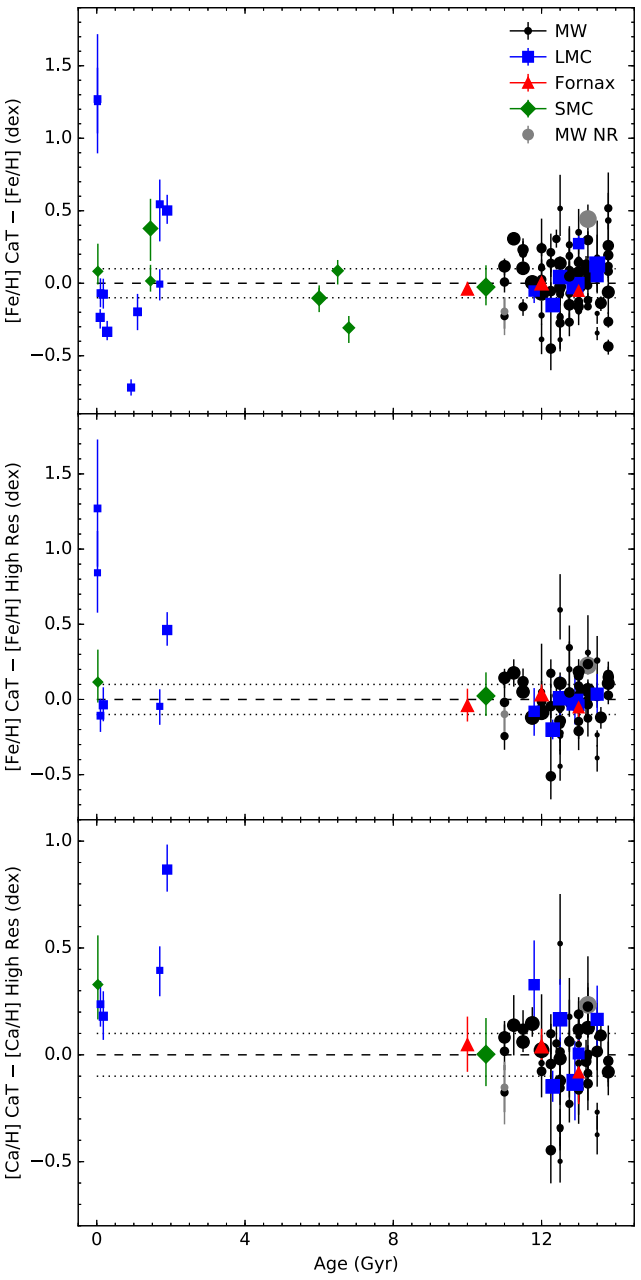

Figure 21. Top: differences in CaT-based [Fe/H] with all literature [Fe/H] as a function of age. Middle: differences in [Fe/H] between CaT-based and high-resolution spectroscopy values. Bottom: differences in [Ca/H] between CaT-based and literature values. The dashed line shows Δ[Fe/H] or Δ[Ca/H] = 0, while the dotted lines show Δ[Fe/H] or Δ[Ca/H] = ±0.1. For ages older than 3 Gyr, the strength of the CaT is independent of age. GCs between 1.4 and 2 Gyr with significant contributions from C stars show significantly stronger CaT strengths than predicted by our CaT–metallicity relations but GCs without significant numbers of C stars in this age range show the same CaT strengths as old GCs with the same [Fe/H]. GCs younger than 1.4 Gyr show CaT strengths slightly weaker than old GCs with the same [Fe/H] except for the youngest GCs (<25 Myr) which are dominated by RSGs. The CaT values of young (<3 Gyr) GCs are larger than their [Ca/H] would predict.

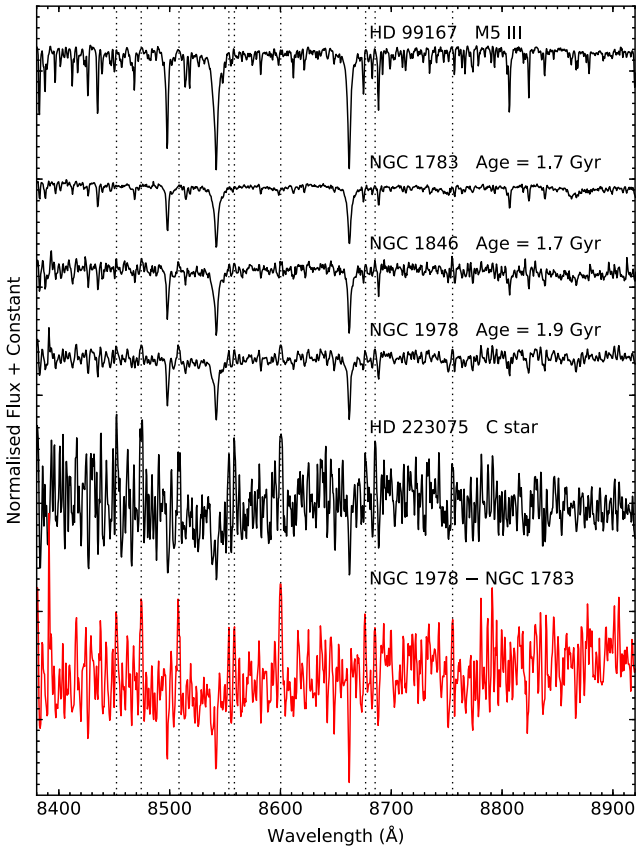

Figure 22. Effects of C stars on the spectra of intermediate-age GC spectra in the CaT region. From top to bottom, we show the spectrum of an M giant from the Indo-US library (Valdes et al. 2004), the spectra of three intermediate-age GCs with similar metallicities, the Indo-US spectrum of a C star and the scaled difference of NGC 1978 and NGC 1783 in red. We note that the S/N of the Indo-US C star spectrum is ∼450 Å^{-1}; the spectrum is dominated by real molecular features and not noise. The dotted vertical lines show some of the stronger C star spectral features. The spectrum of NGC 1978 clearly shows the molecular absorption characteristic of a carbon star, while the NGC 1846 spectrum shows weaker carbon star features. The spectrum of NGC 1783, however, shows no evidence for carbon stars. Despite similar ages and metallicities for all three GCs, we measure significantly higher CaT strengths for NGC 1846 and NGC 1978 ($8.53^{+0.39}_{-0.58}$ and $8.71^{+0.25}_{-0.21}$ Å, respectively) compared to NGC 1783 ($7.62^{+0.23}_{-0.26}$ Å) due our inability to correctly place the continuum when C star molecular features are present.

GCs with stronger than expected CaT strengths show the molecular spectral features distinctive to carbon stars (C stars). If we subtract a scaled spectrum of NGC 1783 from NGC 1978, we get a spectrum that resembles that of a C star. The presence of more C stars in NGC 1978 and to a lesser extent NGC 1846 (compared to NGC 1783) is supported by Frogel, Mould & Blanco (1990) who found that 7 of their 16 AGB stars in NGC 1978 were C stars, 9 out of the 21 stars in NGC 1846, and 5 out of the 17 stars in NGC 1783. As can be seen in Fig. 21, the GCs in this age range without major contributions of light from C stars have CaT based [Fe/H] values consistent with literature values but NGC 1783, the only such GC with abundances from high-resolution spectroscopy, shows a stronger CaT-based [Ca/H] than expected for its literature [Ca/H] abundance by 0.4 dex. The GCs with C stars show CaT-based [Fe/H] ∼ 0.5 dex higher and CaT based [Ca/H] ∼1 dex than their literature values.

The strong C-dominated molecular absorption present in C stars is unlike the atomic and TiO-dominated absorption present in M giants (Fig. 22). The strong molecular absorption present in C stars lowers the pseudo-continuum but for our template-based CaT measurements, the extra absorption in the CaT feature passbands leads to stronger CaT measurements than expected from old GCs with the same metallicity. Since our set of Indo-US templates includes a C star but our DEIMOS templates do not, our Indo-US template-based measurements are more strongly affected by C stars. For the four classical indices we consider, the presence of C stars leads to lower CaT indices strengths compared to the GCs with similar ages and metallicities. We note that there is an observed lack of C stars at ages younger than 300 Myr or older than 4 Gyr in the LMC and SMC (e.g. Frogel et al. 1990) in line with the stellar evolution models of Marigo et al. (2017) who predict that the number of C stars per unit stellar mass peaks at an age of 1.6 Gyr, declining to virtually none for ages younger than 200 Myr or older than 2.5 Gyr. Both observations (Boyer et al. 2013) and models (Marigo et al. 2017) find relatively few C stars at supersolar metallicities.

In the age range of ∼100 Myr to ∼1.2 Gyr, the CaT spectral region is dominated by CHeB and the MSTO is hot (effective temperatures of $\gtrsim$10 000 K). Our GCs in this age range have slightly weaker CaT strengths than old GCs with the same metallicity and strong Paschen absorption. The two GCs at 100–200 Myr with high-resolution abundances however have [Fe/H] values consistent with those calculated from their CaT values (Fig. 21); their calculated [Ca/H] are too high by ∼0.2 dex. Using the classical indices to measure the CaT strength in this age range results in stronger measured CaT strengths due to the contribution from the Paschen 13, 15, and 16 lines which overlap in wavelength with the CaT lines. The CaT* index of Cenarro et al. (2001) attempts to correct for this Paschen absorption but appears to overcorrect at ages less than 1 Gyr as the CaT* indices in this age range are weaker than those of older GCs with similar metallicities. The lack of hot stars in the DEIMOS templates means that only the CaT lines themselves are fit and thus our DEIMOS-based CaT measurements are less affected by Paschen absorption. The Indo-US templates, however, do include hot stars and hence the associated CaT measurements are affected by the Paschen lines in a similar manner to the classical index measurements. The smaller effect of Paschen absorption on the DEIMOS templates motivates our use of them rather than the Indo-US templates.

In the youngest age bin, with ages of only a couple tens of million years, the light in the CaT spectral region is dominated by supergiant stars. The two youngest GCs in our sample – NGC 2004 and NGC 2100 – have ages of 20 Myr (Niederhofer et al. 2015). At this age >90 per cent of the light in the CaT region (Gazak et al. 2014) comes from red supergiant stars. The light of these cool, very low surface gravity stars gives these GCs the strongest CaT values in our sample. The light of the slightly older NGC 330 (30 Myr, Sirianni et al. 2002) is dominated by a mix of blue and red supergiants. The presence of the blue supergiants gives NGC 330 similar Paschen absorption to older (100 Myr–1 Gyr) GCs. For NGC 330, we measure a broadly similar CaT strength to an old population with a similar [Fe/H] using the DEIMOS templates, but a stronger CaT strength than expected from its [Ca/H].

The only previous study to consider the effects of age on the CaT is that of Sakari & Wallerstein (2016) who found no strong effect down to ages of ∼2 Gyr. While this is in line with our observations, we note that Sakari & Wallerstein (2016) only studied three GCs younger than 8 Gyr and the ages used by Sakari & Wallerstein (2016) are mostly based on integrated light spectroscopy that provides

much less reliable ages than those derived by fitting isochrones to photometry of resolved stars.

We find no age effect on our CaT strengths for GCs 6 Gyr old and older. For GCs 2 Gyr old and younger, the CaT is a less reliable metallicity indicator. While intermediate-age GCs with significant numbers of C stars and young GCs dominated by red supergiants have significantly stronger (>0.5 dex) CaT-based [Fe/H] values compared to their literature abundances, the difference between CaT-based and literature [Fe/H] values is generally less than 0.3 dex for the other young- and intermediate-age GCs in our sample. The difference in [Ca/H] is higher due to the low [Ca/Fe] abundance in these GCs. We note that we are limited in our ability to quantify how little the CaT changes between 10 and 6 Gyr by the lack of abundances from high-resolution spectroscopy for the ∼6 Gyr SMC GCs. We also note that we are limited to a relatively small range of ages and metallicities by the availability of high apparent surface brightness GCs in the MW and its satellite galaxies. The CaT is much less sensitive to age than spectral indices commonly used to study metallicity such as the Mg_b and Fe5270 Lick indices (e.g. Worthey 1994; Schiavon 2007; Vazdekis et al. 2010; Conroy & van Dokkum 2012).

4.3 Effects of the horizontal branch morphology

The HB morphology can have a significant effect on the integrated light of a stellar population. In particular, an old population with a particularly blue HB can mimic the appearance of a younger population in the optical (e.g. Lee, Yoon & Lee 2000; Schiavon et al. 2004; Trager et al. 2005). GCs in the MW and other galaxies show a wide range of HB morphologies even at constant metallicity (the second parameter effect, e.g. Sandage & Wallerstein 1960; van den Bergh 1965). Recent work (Dotter et al. 2010; Milone et al. 2014) suggest that much of the variation can be explained by age differences between GCs but a third parameter is also required to explain the range of observed HB morphologies.

In principle, a hot HB can affect the CaT through the effect of stronger Paschen lines contaminating the CaT. In their study of GCs in NGC 1407, Foster et al. (2010) measured significant Paschen absorption in their fitted spectra and suggested that extremely blue HBs would lead to stronger CaT measurements. However, the re-analysis of the same spectra by Usher et al. (2012) found no evidence for significant Paschen absorption, while Usher et al. (2015) found no evidence for strong Paschen absorption in their stacked GC spectra.

In our study, we do not see clear evidence for old GCs with enhanced Paschen line absorption relative to other GCs at the same metallicity. The handful of old GCs with PaT index measurements in Fig. 17 above others with similar metallicity are due to low S/N or poorly subtracted sky lines. To study whether there is a more subtle effect of the HB morphology on the CaT, in Fig. 23 we compare the difference between our CaT-based [Ca/H] with the [Ca/H] from high-resolution spectroscopy with two measures of HB morphology from Milone et al. (2014). The first, the L1 parameter of Milone et al. (2014) measures the colour difference between the HB and the RGB. This parameter depends on the age and metallicity of a GC. The lack of a relationship between this parameter (Kendell's $\tau = -0.07$, $p = 0.63$) and the [Ca/H] difference indicates that we are capturing the behaviour of the CaT with age and metallicity with our fitted [Ca/H]-CaT relation at old ages. The second, the L2 parameter of Milone et al. (2014) measures the colour range of the HB and correlates with the GC mass. A weak correlation between the L2 parameter and the [Ca/Fe] difference is observed (Kendell's $\tau = 0.35$, $p = 0.01$). However, this relation is driven by the highly uncertain measurements of NGC 5927 and NGC 6637. Without GCs with uncertainties in calculated [Ca/H] larger than 0.2 dex, the correlation is non-significant (Kendell's $\tau = 0.25$, $p = 0.09$).

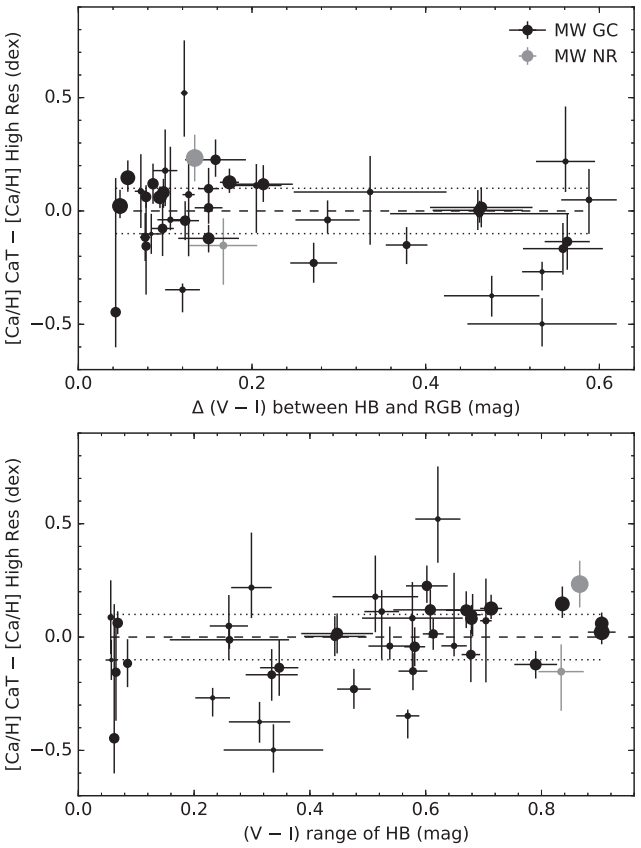

Figure 23. Effects of HB morphology on the CaT. Top: difference between literature and our CaT-based [Ca/H] as a function of the $(V - I)$ colour difference between the red edge of HB and the RGB (the L1 parameter from Milone et al. 2014). Bottom: difference between literature and our CaT-based [Ca/H] as a function of the $(V - I)$ colour range of the HB (the L2 parameter from Milone et al. 2014). The colours, shapes, and size of the points are the same as in Fig. 18. As in Fig. 19, the dashed line shows Δ[Ca/H] =0, while the dotted lines show Δ[Ca/H] =±0.1. The strength of the CaT does not significantly depend on the HB morphology.

We can use the colour range of the HB as a proxy for the size of light element abundance variations within a GC since the colour range of the HB strongly correlates with the size of the He spread (Milone et al. 2014) and with the range of the Na-O anticorrelation (Carretta et al. 2007b; Gratton et al. 2010). The lack of any relationship between the strength of the CaT and the colour range of the HB indicates that the CaT does not strongly vary with the strength of the multiple population phenomena.

4.4 Stochastic and aperture effects

Due to the limited field of view of WiFeS (25 × 38 arcsec), we only observe a fraction of each GC. While for our most distant GCs – those in the Fornax dwarf – we observe ∼75 per cent of the light, for the median MW GC we only observe 10 per cent of the light (see fig. 2 in Usher et al. 2017). This is a concern if the observed (central) part of the GC is not representative of the whole GC.

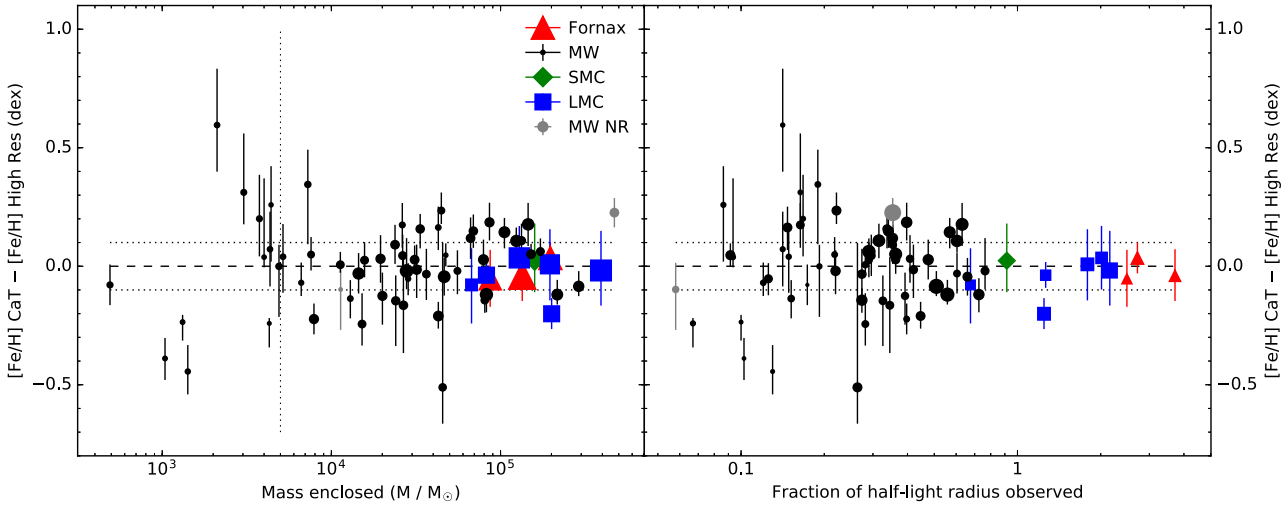

Figure 24. Observational effects on the CaT measurement for old GCs. Left: difference between literature and our CaT-based [Fe/H] as a function of the stellar mass within the field of view of WiFeS. Right: difference between literature and our CaT-based [Fe/H] as a function of the fraction of the half-light radius observed with WiFeS. The colours, shapes, and size of the points are the same as in Fig. 18 save for the left-hand panel where the point size is proportional to the fraction of the half-light radius within the field of view. As in Fig. 19, the dashed line shows Δ[Fe/H] = 0, while the dotted lines show Δ[Fe/H] = $\pm$0.1. For old GCs, the scatter in Δ[Fe/H] increases below an enclosed mass of $\sim 10^4$ M$_\odot$. For this reason we restrict our fits to GCs with more than 5×10^3 M$_\odot$ (the dotted vertical line). Δ[Fe/H] is largely independent of the fraction of the GC observed save for an increased scatter at small fractions due to the low enclosed mass.

The observed part of the GC can be unrepresentative for two reasons. First is stochastic sampling of all stellar evolutionary stages. If the stellar mass sampled by the observations is relatively low, then the number of stars from short lived but bright stellar evolutionary phases can vary dramatically. This is especially a concern for any spectral features in the far red, because of the important contribution to the integrated light by bright short-lived stars. This effect is illustrated in Section 3.4 and in Fig. 15 through the difference in CaT strength in different pointings as well as by the effects of the long-period variable NGC 5927 V3 on the spectrum of NGC 5927 and of variable numbers of C stars on the spectra of intermediate-age GCs (Fig. 22). We note that since stochasticity is a function of the total stellar mass sampled by the observations, regardless of what fraction of GC mass that represents, observations of low-mass GCs observed in their entirety are affected in the same way as those sampling small fractions of massive GCs.

Secondly, if there is a radial stellar population gradient in the GC, the centre is no longer representative of the whole. While normal GCs do not possess significant age (Marino et al. 2012; Nardiello et al. 2015) or metallicity spreads (e.g. Carretta et al. 2009b), GCs do show star-to-star variations in the abundances of various light elements including He, N, O, and Na (See reviews by Gratton et al. 2012; Bastian & Lardo 2017). Radial gradients in abundances of these elements are observed in several GCs (e.g. Carretta et al. 2009a; Lardo et al. 2011; Larsen et al. 2015). More importantly for the CaT, GCs are mass segregated due to dynamical effects (e.g. Rood et al. 1999; Lee et al. 2003; Andreuzzi et al. 2004; Beccari et al. 2015). Since the strength of the CaT strongly depends on surface gravity, a radial variation in the ratio of giant-to-dwarf stars should introduce a radial gradient in the CaT strength.

In Fig. 24, we plot the difference between our CaT-based [Fe/H] and that from literature as both a function of the mass within the observed field of view and as a function of the fraction of the half-light radius within the field of view. The scatter in Δ[Fe/H] is higher at low enclosed mass than at high enclosed mass as well as for GCs with poor radial coverage. As noted earlier, to try to reduce stochastic effects of low stellar mass, we only include GCs more massive than 5×10^3 M$_\odot$ in our fits and comparisons. This limit is broadly similar to the predicted mass for stochastic effects to dominate as calculated by Cerviño & Luridiana (2004) for the *I*-band. We do not see any clear evidence for a relationship between the fraction of the half-light radius observed and the difference between the CaT-based [Fe/H] and the literature values. Kendall's correlation test finds no evidence for a correlation between either the observed mass ($\tau = 0.06$ and $p = 0.49$) or the fraction of the half-light radius observed ($\tau = -0.05$ and $p = 0.59$) and [Fe/H]. This indicates that the CaT value we measure from the centres are representative of the GCs overall.

Extragalactic observations of GCs should be less affected by both of these effects. Due to their much larger distances, it is much easier to observe the entire extent of a GC (in M31 the median GC half-light radius is 0.7 arcsec, e.g. Peacock et al. 2010, at the distance of the Virgo cluster the median radius is 0.03 arcsec, e.g. Masters et al. 2010). Additionally, extragalactic studies typically target GCs in the brighter half of the GC luminosity function and as such observe GCs more massive than $\sim 2 \times 10^5$ M$_\odot$(e.g. Jordán et al. 2007; Villegas et al. 2010). While most spectroscopic observations of galaxies enclose enough mass for stochastic effects not to be an issue, care should be taken with observations of nearby low surface brightness galaxy light. We encourage the effects of properly sampling all stellar evolutionary phase to be considered in both stellar population modelling and observations, especially when testing these methods on GCs in the Local Group. Indeed, valuable insights into stellar populations can be gained by studying how the spectral energy distribution varies from spatial resolution element to element (e.g. Blakeslee, Vazdekis & Ajhar 2001; Jensen et al. 2003; van Dokkum & Conroy 2014; Conroy & van Dokkum 2016).

4.5 Effects of stellar mass function and dynamical evolution

As mentioned earlier, GCs are mass-segregated, with higher mass stars being more centrally concentrated (e.g. Spitzer & Harm 1958). GCs lose mass via tidal interactions and two body relaxation (e.g. Baumgardt & Makino 2003) with stars at larger radii being preferentially lost. Thus, GCs more quickly lose their lower-mass stars giving the present-day mass function of GCs, both globally and in their centres, fewer low-mass stars than their IMFs (e.g. De Marchi, Paresce & Pulone 2007; Paust et al. 2010; Webb & Leigh 2015).

Since stellar population models predict that the CaT is sensitive to the IMF (e.g. Vazdekis et al. 2003; Conroy & van Dokkum 2012), we must consider to what extent the dynamical evolution of GCs affects our measurements. Using the stellar mass function slopes from Sollima & Baumgardt (2017), Kendall's correlation test gives little evidence for a correlation between the mass function slope α (defined such that a Salpeter 1955 IMF has a slope of $\alpha = -2.35$) and the difference between our CaT based [Ca/H] and literature [Ca/H] for the high-mass, high-resolution sample ($\tau = -0.29$ and $p = 0.10$) As for the relationship between [Ca/Fe] and CaT (Section 4.1.1), we use PYMC3 to fit linear relationship between the [Ca/H] difference as a function of α for the 17 GCs in our high-mass, high-resolution sample with mass function slope measurements. We find a slope of $-0.225^{+0.059}_{-0.060}$ and a difference in the BIC of 12, both supporting significant relationship between the slope of the mass function and the strength of the CaT. However, the sign of this relation is in the opposite direction to what is commonly expected for the behaviour of the CaT with the slope of the mass function (the CaT is usually expected to be stronger with a higher fraction of giants).

If we split the sample with mass function measurements up by metallicity, a more complicated picture emerges. In Fig. 25, we plot the difference between our CaT [Ca/H] and literature [Ca/Fe] measurements versus the stellar mass function slope from Sollima & Baumgardt (2017) for GCs with [Fe/H] < -1.5 and > -1.5. We note that the most metal-rich GC in common with our high-mass, high-resolution sample and that of Sollima & Baumgardt (2017) only has a metallicity of [Fe/H] $= -1.2$ (NGC 1851). Using Kendall's correlation test, we see no evidence for an anticorrelation ($\tau = -0.31$ and $p = 0.19$) in the 11 high enclosed-mass GCs in the low-metallicity subsample nor in the higher metallicity subsample ($\tau = -0.20$ and $p = 0.57$) which admittedly only contains six high-mass GCs. Using PYMC3 to fit relations between α and the [Ca/H] difference we find a significant slope for the low-metallicity subsample ($-0.254^{+0.045}_{-0.045}$) but not for the higher metallicity subsample ($0.06^{+0.36}_{-0.40}$). Likewise, the difference in BIC favours a linear relationship over a constant value for the low-metallicity subsample (ΔBIC $= 11.2$) but not for the higher metallicity GCs (ΔBIC $= -0.5$). The correlation between α and the CaT strength seems to be driven by the lower metallicity GCs.

Unfortunately, direct measurements of the stellar mass function are only available for a small subset of our GCs. Since there is a strong anticorrelation between the half-mass radius relaxation time and the stellar mass function (Sollima & Baumgardt 2017), we also consider how the CaT depends of the half-mass radius relaxation time. Following Harris (1996), we calculate the half-mass radius relaxation time using equation (11) of Djorgovski (1993) from the mass and half-light radii in Table 1. For our entire high enclosed mass, high-resolution abundances sample, Kendall's correlation test finds a significant correlation between the log relaxation time and the [Ca/H] difference ($\tau = 0.26$ and $p = 0.005$). Using PYMC3, we find a significant slope ($0.146^{+0.036}_{-0.037}$ for a linear log relaxation time–Δ [Ca/H] relation and strong evidence that a linear relation

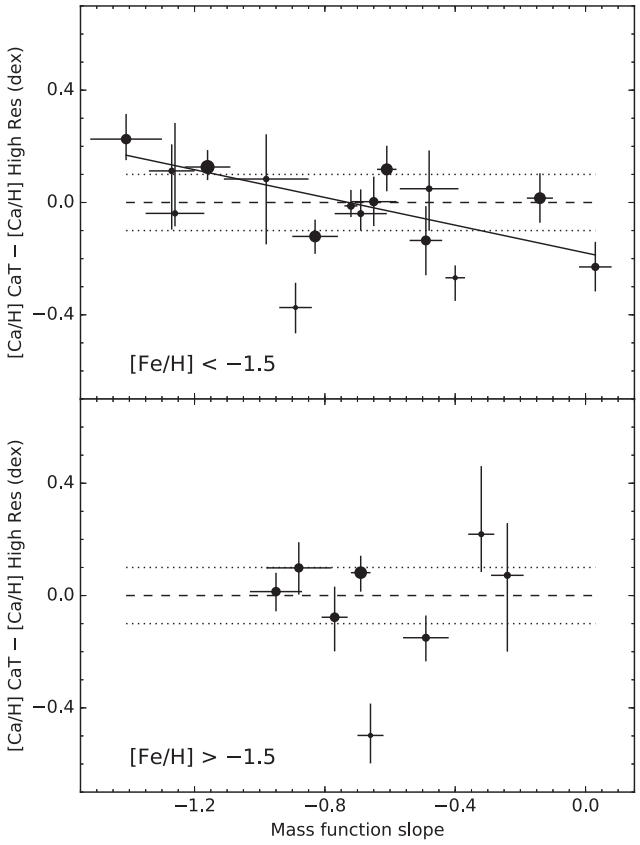

Figure 25. Differences between literature and our CaT-based [Ca/H] as a function of the stellar mass function slope from Sollima & Baumgardt (2017). The top panel shows GCs with [Fe/H] < -1.5, while the bottom panel shows GCs with [Fe/H] > -1.5. The black line in the top panel shows our fitted relationship between Δ[Ca/H] and the mass function slope for the lowest metallicity GCs. The colours, shapes, and size of the points are the same as in Fig. 18. As in Fig. 19, the dashed line shows Δ[Ca/H] $= 0$, while the dotted lines show Δ[Ca/H] $= \pm 0.1$. A Salpeter (1955) IMF would have a slope of $\alpha = -2.35$. The most metal-poor GCs with steeper mass functions (more low-mass stars) show higher CaT values than GCs with similar metallicities but flatter mass functions. No relationship between CaT strength and mass function slope is seen for the more metal-rich GCs though we note the most metal-rich GC in common between our sample and that of Sollima & Baumgardt (2017) has [Fe/H] $= -1.19$ (NGC 1851).

is favoured over a constant value (ΔBIC $= 11.8$). This behaviour is consistent with what is seen with the direct measurements of the mass function slope.

We break our high enclosed mass, high-resolution abundances sample into three metallicity bins ([Fe/H] < -1.5, $-1.5 <$ [Fe/H] < -0.7, and [Fe/H] > -0.7) and plot the difference between CaT based and literature [Ca/H] as a function of half-mass relaxation time for each bin in Fig. 26. For the low-metallicity bin, Kendall's test provides weak evidence ($\tau = 0.26$ and $p = 0.07$) for a correlation. Using PYMC3, we fit a significant slope ($0.202^{+0.051}_{-0.058}$) and find that a linear relationship is favoured over a constant value (ΔBIC $= 9.2$) for 24 GCs in the low-metallicity bin. The nuclear remnant NGC 5139 appears to be a significant outlier from the low-metallicity GC relationship. For both the 20 GCs in the intermediate-metallicity bin and the eight GCs in the high-metallicity bin, Kendall's test provides no evidence ($\tau = 0.18$, $p = 0.27$ and $\tau = 0.00$, $p = 1.0$, respectively) for a correlation. The PYMC3 fits finds an insignificant slope for the intermediate-

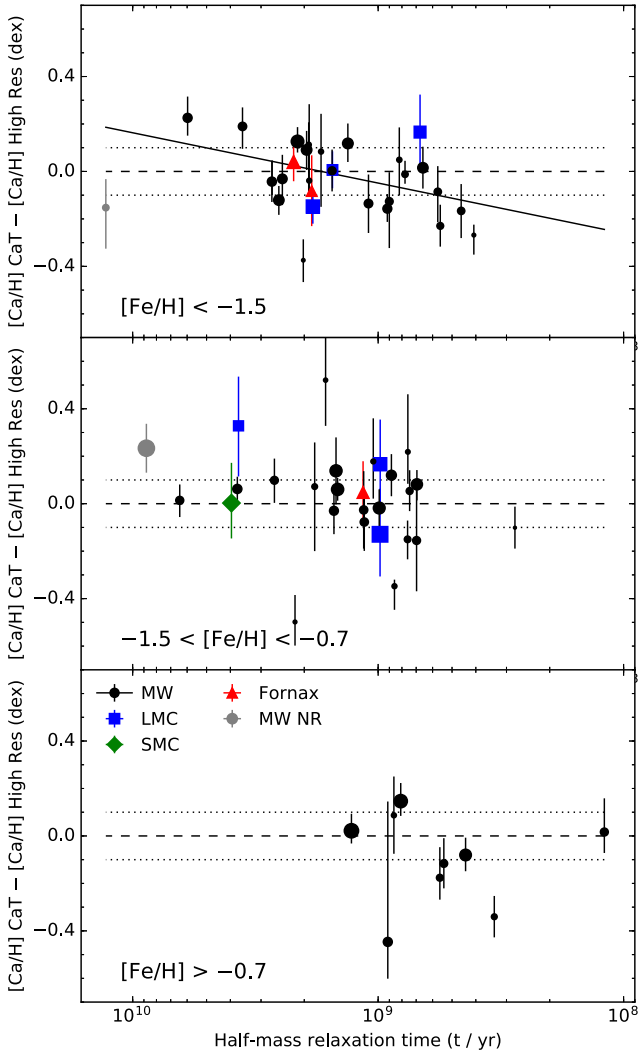

Figure 26. Differences between literature and our CaT-based [Ca/H] as a function of the half-mass radius relaxation time. From top to bottom, we show the difference for GCs with [Fe/H] < -1.5, $-1.5 <$ [Fe/H] < -0.7, and [Fe/H] < -0.7. The black line in the top panel shows our fitted relationship between Δ[Ca/H] and the relaxation time for the lowest metallicity GCs. The colours, shapes, and size of the points are the same as in Fig. 18. Since the slope of the mass function and the relaxation time are anticorrelated, the relaxation time increases from right to left to match the trend seen between the CaT and the mass function slope. As in Fig. 19, the dashed line shows Δ[Ca/H] $= 0$, while the dotted lines show Δ[Ca/H] $= \pm 0.1$. The lowest metallicity GCs with longer relaxation times show stronger CaT strengths than GCs with shorter relaxation times and the same metallicity. However, intermediate- and high-metallicity GCs show no relationship between CaT strength and relaxation time at fixed metallicity.

metallicity subsample ($0.035^{+0.061}_{-0.60}$), but a marginally significant slope for the metal-rich subsample ($0.226^{+0.113}_{-0.103}$). Neither subsample shows strong evidence that a linear relationship is preferred over a constant value (ΔBIC $= -5.7$ and 2.3, respectively). A significant relationship between the CaT and the half-mass relaxation time only at low metallicity is consistent with the behaviour of the CaT with the slope of the stellar mass function. For both the mass function slope and the relaxation time relations, the relationship for all GCs in our sample is driven by the low-metallicity GCs.

The observed qualitative behaviour of the CaT with the stellar mass function slope and metallicity is consistent with the predictions of stellar population synthesis models. The Vazdekis et al. (2003, 2012) models predict that the slope of the relationship between the CaT strength and the slope of the stellar mass function as a strong function of metallicity, with a strong anticorrelation between the two at low metallicities, little or no mass function dependence at intermediate metallicities and a strong correlation at solar metallicities. In the models of Schiavon, Barbuy & Bruzual (2000), the CaT shows no dependence on the mass function slope at subsolar metallicities ([Fe/H] $= -0.5$), but a weak correlation at supersolar metallicities. We note that most of the work on using the CaT as an IMF indicator (e.g. Conroy & van Dokkum 2012; La Barbera et al. 2013) has focused on metallicities close to solar. These models predict that the CaT responds more weakly to the mass function slope for models shallower than that of a Salpeter (1955, $\alpha = -2.35$) IMF, although we note that none makes predictions for mass functions as shallow as observed in some MW GCs ($\alpha \sim 0$, e.g. Paust et al. 2010; Sollima & Baumgardt 2017). The predicted anticorrelation at low metallicities and the lack of significant correlation at intermediate metallicities is consistent with our observations.

The lack of an observed correlation between the mass function and the strength of the CaT at near-solar metallicity is likely due to the small number of high-metallicity GCs in our sample (only NGC 6528 and NGC 6553 are close to solar) and the challenges associated with reliably measuring both the integrated CaT strengths and the chemical abundances of these GCs. We note that the metal-rich GCs have on average shorter relaxation times than the metal-poor GCs due to their smaller mean half-light radii. The lack of a significant relationship between the slope of the stellar mass function and the CaT for most metallicities may explain why we did not find a significant dependence on the fraction of half-light radius observed (Section 4.4).

Dynamical effects should be less serious for extragalactic studies as they typically target the most massive GCs that have longer relaxation times. Galaxies are unaffected by internal dynamical effects due to their much longer relaxation time-scales. However, the effects of a varying IMF remain a concern when using the CaT to measure the metallicity of galaxies although the IMF likely only varies significantly in the centres of some massive galaxies (e.g. Martín-Navarro et al. 2015; van Dokkum et al. 2017). We note that almost all of the GCs in our sample show mass functions shallower than the low-mass slope of a Kroupa (2001) IMF much less the slope of a Salpeter (1955) IMF so that the observed behaviour of the CaT with mass function slope in GCs should not be blindly extended to galaxies. In summary, we find evidence for a weak dependence of the CaT on the mass function slope only at low metallicities.

4.6 Comparison with stellar population models

In Fig. 27, we also plot the predictions of version 11.0 of the Vazdekis et al. (2003, 2010) models using the Girardi et al. (2000) isochrones. For these models, we downloaded the predicted spectral energy distribution and measured the CaT strengths using the same techniques and code as for our spectra. We note that spectral resolutions of the Vazdekis et al. models ($\delta\lambda/\lambda = 5700$) and our observations ($\delta\lambda/\lambda = 6800$) are quite similar so no correction is required to compare them. For old GCs, there is good agreement between the Vazdekis et al. (2003, 2010) models and observations at the $\sim$0.1 dex level. These models predict the observed weak evolution in CaT strength between $\sim$13 and 6 Gyr but underpredict the CaT strength at 2 Gyr and younger. We also show in the left-hand panel of Fig. 27 the CaT–metallicity relationship used by Usher et al. (2012) which is also based on the Vazdekis et al. models. For extragalactic GCs,

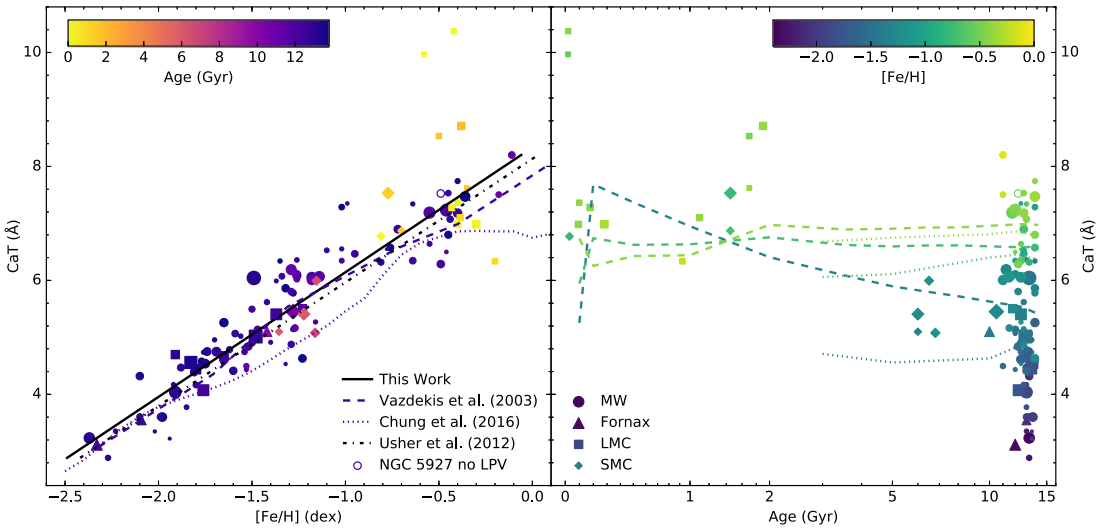

Figure 27. Comparison of CaT strengths with SSP model predictions. Left: CaT strength as a function of metallicity. The points show our measurements colour coded by age with sizes proportional to the mass enclosed within the WiFeS field of view. The dashed line is the 12.6 Gyr model predictions of Vazdekis et al. (2003), the dotted line is the 12.5 Gyr model predictions of Chung et al. (2016), the black dashed–dotted line is the relationship adopted by Usher et al. (2012) based on the Vazdekis et al. (2003) models, and the black solid line is the empirical relationship found in this work (equation 2). Right: CaT strength as a function of age. The points and models are colour coded by metallicity. The dashed lines show the Vazdekis et al. (2003) models and the dotted lines show the Chung et al. (2016) models for metallicites of [Fe/H] = −1.3, −0.7, and −0.4 and a range of ages. In both panels, the empty circle is the CaT measurement of NGC 5927 with the long-period variable NGC 5927 V3 removed from the extracted spectra.

Usher et al. (2012) found good agreement between their metallicities measured using this relation and metallicities measured using Lick indices (Worthey et al. 1994; Worthey & Ottaviani 1997) and stellar population models such as those of Thomas et al. (2003). Unsurprisingly, the Usher et al. (2012) relation shows good agreement with our observations of Local Group GCs.

We also plot the model predictions of Chung et al. (2016). We use their predictions for the index definition of Foster et al. (2010) using the stellar library of Cenarro et al. (2001). For old GCs, the Chung et al. (2016) models predict lower CaT values at higher metallicities which is likely due to the puzzling choice of Chung et al. (2016) to smooth their models to the low spectral resolution ($\delta\lambda/\lambda \lesssim 1000$) of the Lick index system (Worthey et al. 1994) prior to index measurement. The lack of significant CaT strength variation with age is in line with the predictions of the models of both Conroy & van Dokkum (2012) and Chung et al. (2016).

5 CONCLUSION

Using spectra from the WAGGS project (Usher et al. 2017, Section 2), we have measured the strength of the near-infrared CaT spectral feature in a sample of 113 GCs with a wide range of ages and metallicities in the MW and its satellite galaxies (Section 3). After refining the template-based measurement technique of Foster et al. (2010) and Usher et al. (2012), we studied the systematic effects affecting our measurement of the CaT (Section 3.2). We find that masking sky lines in the template fitting process can introduce a strong bias when the radial velocity of the GC causes the CaT lines to overlap with the sky line mask (Figs 6 and 8). This effect is more important in the Local Group than in more distant galaxies because of their recessional velocities. We also find that velocity dispersion or spectral resolution affects the CaT measurement process with our CaT measurements being less reliable above a velocity dispersion of 100 km s^{-1} (or below a spectral resolution of $R = 1200$, Fig. 9). The effects of velocity dispersion are stronger for metal-rich GCs, but are unimportant for the current work given the relatively low-velocity dispersions of GCs (1–20 km s^{-1} e.g. Baumgardt & Hilker 2018) and the intermediate resolution of our spectra ($R = 6800$). In line with previous work (Usher et al. 2012), we find that our CaT measurements are biased to higher values at low metallicity and low S/N. While a concern in extragalactic studies, this effect is not a concern for this study due to our generally high S/N (90 per cent of our spectra have S/N > 22 Å^{-1} and 50 per cent S/N > 116 Å^{-1}).

We also measured the CaT using classical spectral indices (Section 3.3). While below a metallicity of [Fe/H] = −1, the CaT strength measured using the classical indices closely traces our template-based measurements, above this metallicity the classical indices become less sensitive to metallicity. Classical index measurements also show a stronger dependence on age than template-based measurements. Based on these tests, it is clear that the way the CaT is measured affects its behaviour with metallicity and age. We strongly encourage others wanting to use the CaT to measure metallicity to download our spectra (available from https://datacentral.org.au/docs/pages/waggs/) to test their CaT measurement process and their adopted CaT–metallicity relation. We conclude that our template-based CaT measurement technique is the preferable method for measuring the metallicities of extragalactic GCs as it shows higher sensitivity to metallicity at high metallicity, less dependence on age, lower metallicity uncertainties at fixed S/N and less sensitivity to sky subtraction residuals.

We estimated the uncertainties of our measurements using both a Monte Carlo technique based on the uncertainties provided by the data reduction pipeline and by performing a block bootstrap on our datacubes (Section 3.4). At high stellar masses ($\gtrsim 10^4$ M$_\odot$) within the observed field of view, there is good agreement between repeated observations using the bootstrap-based uncertainties, while the pipeline-based uncertainties appear to underpredict the differences between observations. At low stellar observed masses ($\lesssim 10^3$ M$_\odot$), there is significant variation between different pointings of the same GCs. Taken together, this demonstrates that stochastic varia-

tions in the number of bright giant stars is the dominant source of uncertainty for most of our CaT measurements.

In Section 4.1, we derived a relationship between CaT strength and [Fe/H] using our measurements and iron abundances from high-resolution spectroscopy of individual RGB stars (Fig. 18, equation 2). We find that most of the residuals in this relation can be explained by the [Fe/Ca] ratio (Fig. 19) with the relationship between [Ca/H] and CaT strength showing less scatter than the one between [Fe/H] and the CaT (Fig. 20, equation 5). Together with previous work (Usher et al. 2012, 2015; Sakari & Wallerstein 2016), the CaT appears to be a reliable metallicity indicator up to solar metallicity.

We find that our CaT-based metallicity measurements are insensitive to age for ages older than 6 Gyr (Fig. 17, Section 4.2). Younger than 6 Gyr, the effects of age on the CaT based metallicity are relatively small ($\lesssim$0.3 dex in [Fe/H], Fig. 21) except when there is a substantial contribution of light from C stars or red supergiants. The observed behaviour of the CaT with age and metallicity is generally consistent with the predictions of the Vazdekis et al. (2003, 2012) stellar population models but not the Chung et al. (2016) models. Given the CaT's relatively low sensitivity to age and since our observations of ω Cen and M54 show that the CaT can be used to measure the mean metallicities of populations with significant metallicity spreads, the CaT can be reliably used to measure the mean metallicity of galaxy light.

HB morphology has no effect on the CaT (Fig. 23, Section 4.3) implying that the CaT does not vary strongly with the size of the spread in light element abundances (multiple populations) in GCs. We find no significant aperture bias in our observations but find that our CaT measurements are less reliable when there is a low amount of stellar mass in our observed field of view (Fig. 24, Section 4.4). The effects of the stellar mass function on the CaT strength are more complicated (Section 4.5). Below a metallicity of [Fe/H] $= -1.5$, the CaT strength decreases with the slope of the stellar mass function while at intermediate metallicities, the CaT strength is independent of the mass function slope (Fig. 25). At metallicities close to solar, we do not have enough GCs in our sample with abundances from high-resolution spectroscopy to properly assess what effect the mass function slope has on the CaT. That the CaT is stronger in dwarf star-rich stellar populations at low metallicity but not at intermediate metallicities is consistent with the predictions of stellar population models.

We note that our results are only for GCs in one galaxy and its satellites. GCs in other galaxies or integrated galaxy light may have different chemistries or dynamical states at a given metallicity or age and thus slightly different CaT values for the same metallicity. For populations with similar chemistry to the GCs in the MW and its satellite galaxies, we find that the CaT can be used to measure [Ca/H] reliably at the 0.1 dex level from intermediate-resolution integrated spectra of stellar populations older than 3 Gyr. At younger ages, the CaT remains sensitive to metallicity but appears to be biased to higher [Ca/H] values. The CaT is less reliable at high metallicities and low spectral resolutions/high velocity dispersions.

ACKNOWLEDGEMENTS

The authors wish to thank the referee for their useful comments and suggestions which helped to improve this paper. We wish to thank Nate Bastian, Elena Pancino, Joel Pfeffer, Ted Mackereth, and Lee Kelvin for helpful discussions and useful suggestions. We wish to thank the staff at Siding Spring Observatory for their assistance with our observations with the ANU 2.3 m telescope and the WiFeS instrument. CU and SK gratefully acknowledges financial support from the European Research Council (ERC-CoG-646928, Multi-Pop). PC acknowledges the support provided by FONDECYT postdoctoral research grant no. 3160375. SB acknowledges the support of the AAO PhD Topup Scholarship. This work was partially performed on the swinSTAR supercomputer at Swinburne University of Technology. Part of this research was conducted by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013. Based on data products from observations made with ESO Telescopes at the La Silla Paranal Observatory under programme ID 188.B-3002.

This work made use of NUMPY (van der Walt, Colbert & Varoquaux 2011), SCIPY (Jones et al. 2001), and MATPLOTLIB (Hunter 2007) as well as ASTROPY, a community-developed core PYTHON package for astronomy (Astropy Collaboration et al. 2013). Additionally, this work made use of TOPCAT (Taylor 2005) and ALADIN (Bonnarel et al. 2000). This research has made use of the NASA/IPAC Extragalactic Database (NED), which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. This research has made use of the SIMBAD data base, operated at CDS, Strasbourg, France.

REFERENCES

Alamo-Martínez K. A. et al., 2013, ApJ, 775, 20
Alves-Brito A. et al., 2005, A&A, 435, 657
Alves-Brito A. et al., 2006, A&A, 460, 269
Anderson J. et al., 2008, AJ, 135, 2055
Andreuzzi G., Testa V., Marconi G., Alcaino G., Alvarado F., Buonanno R., 2004, A&A, 425, 509
Armandroff T. E., Da Costa G. S., 1991, AJ, 101, 1329
Armandroff T. E., Zinn R., 1988, AJ, 96, 92
Arnold J. A. et al., 2014, ApJ, 791, 80
Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARA&A, 47, 481
Astropy Collaboration et al., 2013, A&A, 558, A33
Barbuy B., Zoccali M., Ortolani S., Hill V., Minniti D., Bica E., Renzini A., Gómez A., 2009, A&A, 507, 405
Bastian N., Lardo C., 2018, ARA&A, 56, 83
Bastian N., Silva-Villa E., 2013, MNRAS, 431, L122
Battaglia G., Irwin M., Tolstoy E., Hill V., Helmi A., Letarte B., Jablonka P., 2008, MNRAS, 383, 183
Baumgardt H., Hilker M., 2018, MNRAS, 478, 1520
Baumgardt H., Makino J., 2003, MNRAS, 340, 227
Beccari G., Dalessandro E., Lanzoni B., Ferraro F. R., Bellazzini M., Sollima A., 2015, ApJ, 814, 144
Bica E., Alloin D., 1987, A&A, 186, 49
Blakeslee J. P., Vazdekis A., Ajhar E. A., 2001, MNRAS, 320, 193
Boeche C., Grebel E. K., 2016, A&A, 587, A2
Bonnarel F. et al., 2000, A&AS, 143, 33
Boyer M. L. et al., 2013, ApJ, 774, 83
Bragaglia A., Carretta E., Sollima A., Donati P., D'Orazi V., Gratton R. G., Lucatello S., Sneden C., 2015, A&A, 583, A69
Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, MNRAS, 427, 127
Brodie J. P., Strader J., 2006, ARA&A, 44, 193
Brodie J. P., Usher C., Conroy C., Strader J., Arnold J. A., Forbes D. A., Romanowsky A. J., 2012, ApJ, 759, L33
Brodie J. P. et al., 2014, ApJ, 796, 52
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Cappellari M., Emsellem E., 2004, PASP, 116, 138
Cappellari M. et al., 2012, Nature, 484, 485
Carrera R., Casamiquela L., Ospina N., Balaguer-Núñez L., Jordi C., Monteagudo L., 2015, A&A, 578, A27

Carretta E., 2015, ApJ, 810, 148
Carretta E., Bragaglia A., 2018, A&A, 614, A109
Carretta E., Cohen J. G., Gratton R. G., Behr B. B., 2001, AJ, 122, 1469
Carretta E., Gratton R. G., Bragaglia A., Bonifacio P., Pasquini L., 2004, A&A, 416, 925
Carretta E. et al., 2007a, A&A, 464, 967
Carretta E., Recio-Blanco A., Gratton R. G., Piotto G., Bragaglia A., 2007b, ApJ, 671, L125
Carretta E. et al., 2009a, A&A, 505, 117
Carretta E., Bragaglia A., Gratton R., D'Orazi V., Lucatello S., 2009b, A&A, 508, 695
Carretta E. et al., 2010a, A&A, 520, A95
Carretta E., Bragaglia A., Gratton R., Lucatello S., Bellazzini M., D'Orazi V., 2010b, ApJ, 712, L21
Carretta E. et al., 2010c, ApJ, 714, L7
Carretta E. et al., 2010d, ApJ, 722, L1
Carretta E., Lucatello S., Gratton R. G., Bragaglia A., D'Orazi V., 2011, A&A, 533, A69
Carretta E. et al., 2013, A&A, 557, A138
Carretta E. et al., 2014, A&A, 564, A60
Carretta E. et al., 2015, A&A, 578, A116
Carretta E., Bragaglia A., Lucatello S., D'Orazi V., Gratton R. G., Donati P., Sollima A., Sneden C., 2017, A&A, 600, A118
Cavallo R. M., Suntzeff N. B., Pilachowski C. A., 2004, AJ, 127, 3411
Cenarro A. J., Cardiel N., Gorgas J., Peletier R. F., Vazdekis A., Prada F., 2001, MNRAS, 326, 959
Cenarro A. J., Gorgas J., Cardiel N., Vazdekis A., Peletier R. F., 2002, MNRAS, 329, 863
Cenarro A. J., Gorgas J., Vazdekis A., Cardiel N., Peletier R. F., 2003, MNRAS, 339, L12
Cenarro A. J., Cardiel N., Gorgas J., 2008, in Knapen J. H., Mahoney T. J., Vazdekis A., eds, ASP Conf. Ser. Vol. 390, Pathways Through an Eclectic Universe, Astron. Soc. Pac., San Francisco. p. 292
Cerviño M., Luridiana V., 2004, A&A, 413, 145
Childress M., Vogt F., Nielsen J., Sharp R., 2014a, Astrophysics Source Code Library, record ascl:1402.034
Childress M. J., Vogt F. P. A., Nielsen J., Sharp R. G., 2014b, Ap&SS, 349, 617
Choi J., Dotter A., Conroy C., Cantiello M., Paxton B., Johnson B. D., 2016, ApJ, 823, 102
Chung C., Yoon S.-J., Lee S.-Y., Lee Y.-W., 2016, ApJ, 818, 201
Cohen J. G., 1978, ApJ, 221, 788
Cohen J. G., Meléndez J., 2005, AJ, 129, 303
Cohen J. G., Gratton R. G., Behr B. B., Carretta E., 1999, ApJ, 523, 739
Cohen J. G., Huang W., Kirby E. N., 2011, ApJ, 740, 60
Cole A. A., Smecker-Hane T. A., Tolstoy E., Bosler T. L., Gallagher J. S., 2004, MNRAS, 347, 367
Cole A. A., Tolstoy E., Gallagher J. S., III, Smecker-Hane T. A., 2005, AJ, 129, 1465
Colucci J. E., Bernstein R. A., Cameron S. A., McWilliam A., 2011, ApJ, 735, 55
Colucci J. E., Bernstein R. A., Cameron S. A., McWilliam A., 2012, ApJ, 746, 29
Colucci J. E., Fernanda Durán M., Bernstein R. A., McWilliam A., 2013, ApJ, 773, L36
Conroy C., van Dokkum P., 2012, ApJ, 747, 69
Conroy C., van Dokkum P. G., 2016, ApJ, 827, 9
Conroy C., Villaume A., van Dokkum P. G., Lind K., 2018, ApJ, 854, 139
Cordero M. J., Pilachowski C. A., Johnson C. I., McDonald I., Zijlstra A. A., Simmerer J., 2014, ApJ, 780, 94
Cordero M. J., Pilachowski C. A., Johnson C. I., Vesperini E., 2015, ApJ, 800, 3
Da Costa G. S., 2016, MNRAS, 455, 199
Da Costa G. S., Hatzidimitriou D., 1998, AJ, 115, 1934
Dalessandro E., Lapenna E., Mucciarelli A., Origlia L., Ferraro F. R., Lanzoni B., 2016, ApJ, 829, 77
De Angeli F., Piotto G., Cassisi S., Busso G., Recio-Blanco A., Salaris M., Aparicio A., Rosenberg A., 2005, AJ, 130, 116
de Boer T. J. L., Fraser M., 2016, A&A, 590, A35
De Marchi G., Paresce F., Pulone L., 2007, ApJ, 656, L65
Dias B., Barbuy B., Saviane I., Held E. V., Da Costa G. S., Ortolani S., Gullieuszik M., Vásquez S., 2016, A&A, 590, A9
Djorgovski S., 1993, in Djorgovski S. G., Meylan G., eds, ASP Conf. Ser. Vol. 50, Structure and Dynamics of Globular Clusters, Astron. Soc. Pac., San Francisco, p. 373
Dopita M., Hart J., McGregor P., Oates P., Bloxham G., Jones D., 2007, Ap&SS, 310, 255
Dopita M. et al., 2010, Ap&SS, 327, 245
Dotter A., Chaboyer B., Ferguson J. W., Lee H.-c., Worthey G., Jevremović D., Baron E., 2007, ApJ, 666, 403
Dotter A. et al., 2010, ApJ, 708, 698
Dotter A., Sarajedini A., Anderson J., 2011, ApJ, 738, 74
Faber S. M. et al., 2003, in Iye M., Moorwood A. F. M., eds, Proc. SPIE Conf. Vol. 4841, Instrument Design and Performance for Optical/Infrared Ground-based Telescopes, SPIE, Bellingham, p. 1657
Feltzing S., Primas F., Johnson R. A., 2009, A&A, 493, 913
Fischer P., Welch D. L., Mateo M., 1992, AJ, 104, 1086
Forbes D. A., Bridges T., 2010, MNRAS, 404, 1203
Forbes D. A. et al., 2018, Proc. R. Soc. Lond. Ser. A, 474, 20170616
Foster C., Proctor R. N., Forbes D. A., Spolaor M., Hopkins P. F., Brodie J. P., 2009, MNRAS, 400, 2135
Foster C., Forbes D. A., Proctor R. N., Strader J., Brodie J. P., Spitler L. R., 2010, AJ, 139, 1566
Foster C. et al., 2011, MNRAS, 415, 3393
Frogel J. A., Mould J., Blanco V. M., 1990, ApJ, 352, 96
Gazak J. Z. et al., 2014, ApJ, 787, 142
Geisler D., Bica E., Dottori H., Claria J. J., Piatti A. E., Santos J. F. C., Jr, 1997, AJ, 114, 1920
Gieles M., Zocchi A., 2015, MNRAS, 454, 576
Gilmore G. et al., 2012, The Messenger, 147, 25
Girardi L., Bertelli G., 1998, MNRAS, 300, 533
Girardi L., Bressan A., Bertelli G., Chiosi C., 2000, A&AS, 141, 371
Girardi L., Marigo P., Bressan A., Rosenfield P., 2013, ApJ, 777, 142
Glatt K. et al., 2008a, AJ, 135, 1106
Glatt K. et al., 2008b, AJ, 136, 1703
Glatt K. et al., 2009, AJ, 138, 1403
Glazebrook K., Bland-Hawthorn J., 2001, PASP, 113, 197
Goudfrooij P., Mack J., Kissler-Patig M., Meylan G., Minniti D., 2001, MNRAS, 322, 643
Goudfrooij P., Gilmore D., Kissler-Patig M., Maraston C., 2006, MNRAS, 369, 697
Goudfrooij P., Puzia T. H., Kozhurina-Platais V., Chandar R., 2009, AJ, 137, 4988
Goudfrooij P., Puzia T. H., Kozhurina-Platais V., Chandar R., 2011, ApJ, 737, 3
Goudfrooij P. et al., 2014, ApJ, 797, 35
Gratton R. G. et al., 2001, A&A, 369, 87
Gratton R. G., Carretta E., Bragaglia A., Lucatello S., D'Orazi V., 2010, A&A, 517, A81 +
Gratton R. G., Carretta E., Bragaglia A., 2012, A&AR, 20, 50
Gratton R. G. et al., 2015, A&A, 573, A92
Grocholski A. J., Cole A. A., Sarajedini A., Geisler D., Smith V. V., 2006, AJ, 132, 1630
Harris W. E., 1996, AJ, 112, 1487
Harris W. E., 2010, preprint (arXiv:1012.3224)
Hayes C. R. et al., 2018, ApJ, 852, 49
Hesser J. E., Shawl S. J., Meyer J. E., 1986, PASP, 98, 403
Hilker M., Richtler T., 2000, A&A, 362, 895
Hill V., 1999, A&A, 345, 430
Hopkins P. F. et al., 2018, MNRAS, 480, 800
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Ibata R. A., Wyse R. F. G., Gilmore G., Irwin M. J., Suntzeff N. B., 1997, AJ, 113, 634
Ivans I. I., Sneden C., Kraft R. P., Suntzeff N. B., Smith V. V., Langer G. E., Fulbright J. P., 1999, AJ, 118, 1273

Ivans I. I., Kraft R. P., Sneden C., Smith G. H., Rich R. M., Shetrone M., 2001, AJ, 122, 1438
Jensen J. B., Tonry J. L., Barris B. J., Thompson R. I., Liu M. C., Rieke M. J., Ajhar E. A., Blakeslee J. P., 2003, ApJ, 583, 712
Johnson C. I., Pilachowski C. A., 2006, AJ, 132, 2346
Johnson C. I., Pilachowski C. A., 2010, ApJ, 722, 1373
Johnson J. A., , Bolte M., Hesser J. E., Ivans I. I., Stetson P. B., 2004, in McWilliam A., Rauch M., eds, Origin and Evolution of the Elements, Cambridge University Press, Cambridge
Johnson J. A., Ivans I. I., Stetson P. B., 2006, ApJ, 640, 801
Johnson C. I., Rich R. M., Kobayashi C., Kunder A., Koch A., 2014, AJ, 148, 67
Johnson C. I., Rich R. M., Pilachowski C. A., Caldwell N., Mateo M., Bailey J. I., III, Crane J. D., 2015, AJ, 150, 63
Johnson C. I., Caldwell N., Rich R. M., Pilachowski C. A., Hsyu T., 2016, AJ, 152, 21
Johnson C. I., Caldwell N., Rich R. M., Mateo M., Bailey J. I., III, Olszewski E. W., Walker M. G., 2017, ApJ, 842, 24
Johnson C. I., Rich R. M., Caldwell N., Mateo M., Bailey J. I., III, Olszewski E. W., Walker M. G., 2018, AJ, 155, 71
Jones E., et al., 2001, SciPy: Open Source Scientific Tools for Python. Available at: http://www.scipy.org/Last Accessed 2018-06-11
Jordán A. et al., 2007, ApJS, 171, 101
Kacharov N., Koch A., McWilliam A., 2013, A&A, 554, A81
Katz D. et al., 2004, MNRAS, 354, 1223
Kerber L. O., Santiago B. X., Brocato E., 2007, A&A, 462, 139
Kerber L. O., Nardiello D., Ortolani S., Barbuy B., Bica E., Cassisi S., Libralato M., Vieira R. G., 2018, ApJ, 853, 15
King I. R., 1966, AJ, 71, 64
Kirby E. N., Guhathakurta P., Sneden C., 2008, ApJ, 682, 1217
Koch A., McWilliam A., 2008, AJ, 135, 1551
Koch A., McWilliam A., 2010, AJ, 139, 2289
Korn A. J., Keller S. C., Kaufer A., Langer N., Przybilla N., Stahl O., Wolf B., 2002, A&A, 385, 143
Korn A. J., Maraston C., Thomas D., 2005, A&A, 438, 685
Kraft R. P., Sneden C., Smith G. H., Shetrone M. D., Fulbright J., 1998, AJ, 115, 1500
Kroupa P., 2001, MNRAS, 322, 231
Kruijssen J. M. D., Pfeffer J. L., Reina-Campos M., Crain R. A., Bastian N., 2018, MNRAS, in press
Kurucz R. L., 1970, SAO Special Report, Smithsonian Astrophysical Observatory, Cambridge, MA, p. 309
Kurucz R. L., 1993, SYNTHE Spectrum Synthesis Programs and Line Data, Smithsonian Astrophysical Observatory, Cambridge, MA
Kurucz R. L., 2013, Astrophysics Source Code Library, record ascl:1303.024
La Barbera F., Ferreras I., Vazdekis A., de la Rosa I. G., de Carvalho R. R., Trevisan M., Falcón-Barroso J., Ricciardelli E., 2013, MNRAS, 433, 3017
Lagioia E. P. et al., 2014, ApJ, 782, 50
Lai D. K., Smith G. H., Bolte M., Johnson J. A., Lucatello S., Kraft R. P., Sneden C., 2011, AJ, 141, 62
Lamb M. P., Venn K. A., Shetrone M. D., Sakari C. M., Pritzl B. J., 2015, MNRAS, 448, 42
Lardo C., Bellazzini M., Pancino E., Carretta E., Bragaglia A., Dalessandro E., 2011, A&A, 525, A114
Larsen S. S., Brodie J. P., Strader J., 2012, A&A, 546, A53
Larsen S. S., Baumgardt H., Bastian N., Brodie J. P., Grundahl F., Strader J., 2015, ApJ, 804, 71
Layden A. C., Sarajedini A., 2000, AJ, 119, 1760
Leaman R. et al., 2013, ApJ, 767, 131
Lee K. H., Lee H. M., Fahlman G. G., Lee M. G., 2003, AJ, 126, 815
Lee H.-c., Yoon S.-J., Lee Y.-W., 2000, AJ, 120, 998
Lee H.-c. et al., 2009, ApJ, 694, 902
Lee J.-W., 2007, Rev. Mex. Astron. Astrofis., 27, 120
Lee J.-W., Carney B. W., 2002, AJ, 124, 1511
Lee J.-W., López-Morales M., Carney B. W., 2006, ApJ, 646, L119
Lemasle B. et al., 2014, A&A, 572, A88
Letarte B., Hill V., Jablonka P., Tolstoy E., François P., Meylan G., 2006, A&A, 453, 547
Li C., de Grijs R., Deng L., 2014, ApJ, 784, 157
Liu Y. J., Tan K. F., Wang L., Zhao G., Sato B., Takeda Y., Li H. N., 2014, ApJ, 785, 94
Liu C., Ruchti G., Feltzing S., Primas F., 2017, A&A, 601, A31
Mackey A. D., Gilmore G. F., 2003, MNRAS, 338, 85
Maldonado J., Eiroa C., Villaver E., Montesinos B., Mora A., 2015, A&A, 579, A20
Marigo P. et al., 2017, ApJ, 835, 77
Marino A. F., Villanova S., Piotto G., Milone A. P., Momany Y., Bedin L. R., Medling A. M., 2008, A&A, 490, 625
Marino A. F. et al., 2011, A&A, 532, A8
Marino A. F. et al., 2012, A&A, 541, A15
Marino A. F. et al., 2015, MNRAS, 450, 815
Marino A. F. et al., 2018, ApJ, 859, 81
Martín-Navarro I., La Barbera F., Vazdekis A., Falcón-Barroso J., Ferreras I., 2015, MNRAS, 447, 1033
Martocchia S. et al., 2018, MNRAS, 473, 2688
Massari D. et al., 2017, MNRAS, 468, 1249
Masters K. L. et al., 2010, ApJ, 715, 1419
McLaughlin D. E., van der Marel R. P., 2005, ApJS, 161, 304
McWilliam A., Wallerstein G., Mottini M., 2013, ApJ, 778, 149
Meissner F., Weiss A., 2006, A&A, 456, 1085
Meléndez J., Barbuy B., Bica E., Zoccali M., Ortolani S., Renzini A., Hill V., 2003, A&A, 411, 417
Mészáros S. et al., 2015, AJ, 149, 153
Mighell K. J., Sarajedini A., French R. S., 1998, AJ, 116, 2395
Milone A. P. et al., 2014, ApJ, 785, 21
Misgeld I., Mieske S., Hilker M., Richtler T., Georgiev I. Y., Schuberth Y., 2011, A&A, 531, A4
Mucciarelli A., Ferraro F. R., Origlia L., Fusi Pecci F., 2007, AJ, 133, 2053
Mucciarelli A., Carretta E., Origlia L., Ferraro F. R., 2008, AJ, 136, 375
Mucciarelli A., Origlia L., Ferraro F. R., 2010, ApJ, 717, 277
Mucciarelli A. et al., 2011, MNRAS, 413, 837
Mucciarelli A., Origlia L., Ferraro F. R., Bellazzini M., Lanzoni B., 2012, ApJ, 746, L19
Mucciarelli A., Bellazzini M., Catelan M., Dalessandro E., Amigo P., Correnti M., Cortés C., D'Orazi V., 2013, MNRAS, 435, 3667
Muñoz C., Geisler D., Villanova S., 2013, MNRAS, 433, 2006
Muñoz C., Villanova S., Geisler D., Saviane I., Dias B., Cohen R. E., Mauro F., 2017, A&A, 605, A12
Mura-Guzmán A., Villanova S., Muñoz C., Tang B., 2018, MNRAS, 474, 4541
Nardiello D. et al., 2015, MNRAS, 451, 312
Ness M., Asplund M., Casey A. R., 2014, MNRAS, 445, 2994
Niederhofer F., Hilker M., Bastian N., Silva-Villa E., 2015, A&A, 575, A62
Norris M. A. et al., 2014, MNRAS, 443, 1151
O'Malley E. M., Knaizev A., McWilliam A., Chaboyer B., 2017, ApJ, 846, 23
Olsen K. A. G., Hodge P. W., Mateo M., Olszewski E. W., Schommer R. A., Suntzeff N. B., Walker A. R., 1998, MNRAS, 300, 665
Olszewski E. W., Schommer R. A., Suntzeff N. B., Harris H. C., 1991, AJ, 101, 515
Origlia L., Valenti E., Rich R. M., 2005, MNRAS, 356, 1276
Origlia L., Valenti E., Rich R. M., 2008, MNRAS, 388, 1419
Pancino E. et al., 2017, A&A, 601, A112
Pastorello N., Forbes D. A., Foster C., Brodie J. P., Usher C., Romanowsky A. J., Strader J., Arnold J. A., 2014, MNRAS, 442, 1003
Patrick L. R., Evans C. J., Davies B., Kudritzki R.-P., Hénault-Brunet V., Bastian N., Lapenna E., Bergemann M., 2016, MNRAS, 458, 3968
Paust N. E. Q. et al., 2010, AJ, 139, 476
Peacock M. B., Maccarone T. J., Knigge C., Kundu A., Waters C. Z., Zepf S. E., Zurek D. R., 2010, MNRAS, 402, 803
Pota V. et al., 2013, MNRAS, 428, 389
Prugniel P., Vauglin I., Koleva M., 2011, A&A, 531, A165
Puzia T. H., Kissler-Patig M., Thomas D., Maraston C., Saglia R. P., Bender R., Goudfrooij P., Hempel M., 2005, A&A, 439, 997

Ramírez S. V., Cohen J. G., 2002, AJ, 123, 3277
Ramírez S. V., Cohen J. G., 2003, AJ, 125, 224
Recio-Blanco A. et al., 2017, A&A, 602, L14
Roederer I. U., Marino A. F., Sneden C., 2011, ApJ, 742, 37
Rood R. T. et al., 1999, ApJ, 523, 752
Rutledge G. A., Hesser J. E., Stetson P. B., 1997, PASP, 109, 907
Saglia R. P., Maraston C., Thomas D., Bender R., Colless M., 2002, ApJ, 579, L13
Sakari C. M., Wallerstein G., 2016, MNRAS, 456, 831
Sakari C. M. et al., 2016, ApJ, 829, 116
Salpeter E. E., 1955, ApJ, 121, 161
Salvatier J., Wiecki T. V., Fonnesbeck C., 2016, PeerJ Comput. Sci., 2, e55
Sandage A., Wallerstein G., 1960, ApJ, 131, 598
Sarajedini A. et al., 2007, AJ, 133, 1658
Saviane I., Da Costa G. S., Held E. V., Sommariva V., Gullieuszik M., Barbuy B., Ortolani S., 2012, A&A, 540, A27
Sbordone L. et al., 2015, A&A, 579, A104
Schiavon R. P., 2007, ApJS, 171, 146
Schiavon R. P., Barbuy B., Bruzual A. G., 2000, ApJ, 532, 453
Schiavon R. P., Rose J. A., Courteau S., MacArthur L. A., 2004, ApJ, 608, L33
Schiavon R. P. et al., 2017, MNRAS, 466, 1010
Searle L., Zinn R., 1978, ApJ, 225, 357
Siegel M. H. et al., 2007, ApJ, 667, L57
Sirianni M., Nota A., De Marchi G., Leitherer C., Clampin M., 2002, ApJ, 579, 275
Skrutskie M. F. et al., 2006, AJ, 131, 1163
Sloan G. C. et al., 2010, ApJ, 719, 1274
Sneden C., Kraft R. P., Shetrone M. D., Smith G. H., Langer G. E., Prosser C. F., 1997, AJ, 114, 1964
Sneden C., Kraft R. P., Guhathakurta P., Peterson R. C., Fulbright J. P., 2004, AJ, 127, 2162
Sollima A., Baumgardt H., 2017, MNRAS, 471, 3668
Spinrad H., Taylor B. J., 1971, ApJS, 22, 445
Spitzer L., Jr, Harm R., 1958, ApJ, 127, 544
Stanford L. M., Da Costa G. S., Norris J. E., Cannon R. D., 2006, ApJ, 647, 1075
Steinmetz M. et al., 2006, AJ, 132, 1645
Tang B. et al., 2017, MNRAS, 465, 19
Taylor M. B., 2005, in Shopbell P., Britton M., Ebert R., eds, ASP Conf. Ser. Vol. 347, Astronomical Data Analysis Software and Systems XIV, Astron. Soc. Pac., San Francisco. p. 29
Thomas D., Maraston C., Bender R., 2003, MNRAS, 339, 897
Thygesen A. O. et al., 2014, A&A, 572, A108
Tolstoy E. et al., 2004, ApJ, 617, L119
Trager S. C., Worthey G., Faber S. M., Dressler A., 2005, MNRAS, 362, 2
Tripicco M. J., Bell R. A., 1995, AJ, 110, 3035
Usher C. et al., 2012, MNRAS, 426, 1475
Usher C., Forbes D. A., Spitler L. R., Brodie J. P., Romanowsky A. J., Strader J., Woodley K. A., 2013, MNRAS, 436, 1172
Usher C. et al., 2015, MNRAS, 446, 369
Usher C. et al., 2017, MNRAS, 468, 3828
Valdes F., Gupta R., Rose J. A., Singh H. P., Bell D. J., 2004, ApJS, 152, 251
Valenti E., Origlia L., Rich R. M., 2011, MNRAS, 414, 2690
van den Bergh S., 1965, J. R. Astron. Soc. Can., 59, 151
Van der Swaelmen M., Hill V., Primas F., Cole A. A., 2013, A&A, 560, A44
van der Walt S., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22
van Dokkum P. G., Conroy C., 2010, Nature, 468, 940
van Dokkum P. G., Conroy C., 2014, ApJ, 797, 56
van Dokkum P., Conroy C., Villaume A., Brodie J., Romanowsky A. J., 2017, ApJ, 841, 68
VandenBerg D. A., Bergbusch P. A., Dotter A., Ferguson J. W., Michaud G., Richer J., Proffitt C. R., 2012, ApJ, 755, 15
VandenBerg D. A., Brogaard K., Leaman R., Casagrande L., 2013, ApJ, 775, 134
Vásquez S. et al., 2018, A&A, preprint (arXiv:1808.03834)
Vazdekis A., Cenarro A. J., Gorgas J., Cardiel N., Peletier R. F., 2003, MNRAS, 340, 1317
Vazdekis A., Sánchez-Blázquez P., Falcón-Barroso J., Cenarro A. J., Beasley M. A., Cardiel N., Gorgas J., Peletier R. F., 2010, MNRAS, 404, 1639
Vazdekis A., Ricciardelli E., Cenarro A. J., Rivero-González J. G., Díaz-García L. A., Falcón-Barroso J., 2012, MNRAS, 424, 157
Villanova S., Geisler D., 2011, A&A, 535, A31
Villanova S. et al., 2007, ApJ, 663, 296
Villanova S., Geisler D., Piotto G., 2010, ApJ, 722, L18
Villanova S., Geisler D., Gratton R. G., Cassisi S., 2014, ApJ, 791, 107
Villanova S., Monaco L., Moni Bidin C., Assmann P., 2016, MNRAS, 460, 2351
Villanova S., Moni Bidin C., Mauro F., Munoz C., Monaco L., 2017, MNRAS, 464, 2730
Villegas D. et al., 2010, ApJ, 717, 603
Walker A. R., 1992, AJ, 104, 1395
Webb J. J., Leigh N. W. C., 2015, MNRAS, 453, 3278
Webbink R. F., 1981, ApJS, 45, 259
Wenger M. et al., 2000, A&AS, 143, 9
Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009, MNRAS, 399, 574
Woosley S. E., Weaver T. A., 1995, ApJS, 101, 181
Worthey G., 1994, ApJS, 95, 107
Worthey G., Ottaviani D. L., 1997, ApJS, 111, 377
Worthey G., Faber S. M., Gonzalez J. J., Burstein D., 1994, ApJS, 94, 687
Yong D., Grundahl F., 2008, ApJ, 672, L29
Yong D., Meléndez J., Cunha K., Karakas A. I., Norris J. E., Smith V. V., 2008a, ApJ, 689, 1020
Yong D., Karakas A. I., Lambert D. L., Chieffi A., Limongi M., 2008b, ApJ, 689, 1031
Yong D. et al., 2014, MNRAS, 439, 2638
Zinn R., West M. J., 1984, ApJS, 55, 45
Zoccali M., Renzini A., Ortolani S., Bica E., Barbuy B., 2001, AJ, 121, 2638
Zoccali M. et al., 2004, A&A, 423, 507

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Table 1. Sample properties.
Table 2. Adopted high-resolution abundances.
Table 5. CaT measurements.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a TEX/LATEX file prepared by the author.