A&A, 694, A248 (2025)
https://doi.org/10.1051/0004-6361/202452420

Astronomy & Astrophysics

A stellar census in globular clusters with MUSE

Metallicity spread and dispersion among first-population stars*

M. Latour[1,**], S. Kamann[2], S. Martocchia[3], T.-O. Husser[1], S. Saracino[2], and S. Dreizler[1]

[1] Institut für Astrophysik und Geophysik, Georg-August-Universität Göttingen, Friedrich-Hund-Platz 1, 37077 Göttingen, Germany
[2] Astrophysics Research Institute, Liverpool John Moores University, IC2 Liverpool Science Park, 146 Brownlow Hill, Liverpool, L3 5RF, UK
[3] Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Mönchhofstraße 12-14, 69120 Heidelberg, Germany

Received 30 September 2024 / Accepted 15 January 2025

ABSTRACT

Context. Multiple populations are ubiquitous in the old massive globular clusters (GCs) of the Milky Way. It is still unclear how they arose during the formation of a GC. The topic of iron and metallicity variations has recently attracted attention with the measurement of iron variations among the primordial population (P1) stars of Galactic GCs.
Aims. We explore the relationship between the metallicity of the P1 stars on the red-giant branch (RGB) of Galactic GCs and their $\Delta_{F275W,F814W}$ pseudo-color. We also measure the metallicity dispersion of P1 and P2 stars.
Methods. We used the spectra of more than 8000 RGB stars in 21 Galactic GCs observed with the integral-field spectrograph MUSE to derive individual stellar metallicities, [M/H]. For each cluster, we used *Hubble* Space Telescope photometric catalogs to separate the stars into two main populations (P1 and P2). We measured the metallicity spread within the primordial population of each cluster by combining our metallicity measurements with the stars' $\Delta_{F275W,F814W}$ pseudo-color. We also derived metallicity dispersions ($\sigma_{[M/H]}$) for the P1 and P2 stars of each GC.
Results. In all but three GCs we find a significant correlation between the metallicity and the $\Delta_{F275W,F814W}$ pseudo-color of the P1 stars: stars with larger $\Delta_{F275W,F814W}$ have higher metallicities. We measure metallicity spreads that range from 0.03 to 0.24 dex and correlate with the GC masses. As for the intrinsic metallicity dispersions, when combining the P1 and P2 stars, we measure values ranging from 0.02 dex to 0.08 dex, which correlate very well with the GC masses. The two clusters that show the largest $\sigma_{[M/H]}$ are NGC 6388 and NGC 6441. The P2 stars have metallicity dispersions that are smaller than or equal to those of the P1 stars.
Conclusions. We present a homogeneous spectroscopic characterization of the metallicities of the P1 and P2 stars in a set of 21 Galactic GCs. We find that both the metallicity spreads of the P1 stars (from the $\Delta_{F275W,F814W}$ spread on the chromosome maps) and the metallicity dispersions ($\sigma_{[M/H]}$) correlate with the GC masses, as predicted by some theoretical self-enrichment models presented in the literature.

Key words. stars: abundances – stars: fundamental parameters – globular clusters: general

1. Introduction

The formation of massive star clusters has been puzzling astrophysicists over the past few decades. Despite considerable progress recently gained through observations and hydrodynamical simulations, a number of open questions remain. This is probably best illustrated by the enigma of the multiple-populations: star-to-star abundance variations in light elements (such as C, N, O, and Na) that are ubiquitous in massive clusters older than ∼2 Gyr, such as the Galactic globular clusters (GCs) or intermediate-age star clusters in the Magellanic Clouds (see Bastian & Lardo 2018; Gratton et al. 2019, for recent reviews). Despite a multitude of observational studies and a number of proposed scenarios, the mechanisms underlying the formation of multiple populations are still unknown.

From an observational perspective, the ubiquity of multiple-populations has been confirmed both photometrically and spectroscopically (e.g., Carretta et al. 2009b). The work of Milone et al. (2017), making use of UV and optical magnitudes from the *Hubble* Space Telescope (HST) in the form of chromosome maps (which are pseudo-two-color diagrams), confirmed the presence of multiple populations in a sample of 57 Galactic GCs. In all of these clusters, the red-giant branch (RGB) stars have a spread in color larger than what is expected from photometric errors only. The pseudo-color (i.e., $(m_{F275W} - m_{F336W})-(m_{F336W} - m_{F814W})$) allows an efficient separation of the two principal population of stars in GCs: the so-called first and second populations (hereafter P1 and P2). This is because this pseudo-color mainly traces nitrogen variations (Milone et al. 2017). By matching photometric and spectroscopic properties of the RGB stars of these populations, it was found that the P1 stars have an atmospheric chemistry that shows primordial abundances (i.e., a scaled solar-like abundance pattern but with some degree of α enhancement) and the P2 stars have "anomalous" abundance patterns, most notably enhancements in N and Na and depletions in O (see, e.g., Marino et al. 2019a, Milone et al. 2015, Carretta & Bragaglia 2024.) A few GCs have an additional population (P3) that is

* Based on observations collected at the European Organisation for Astronomical Research in the Southern Hemisphere, Chile (Program IDs 094.D-0142(B), 095.D-0629(A), 096.D-0175(A), 097.D-0295(A), 098.D-0148(A), 099.D-0019(A), 0100.D-0161(A), 0101.D-0268(A), 0102.D-0270(A), 0103.D-0204(A), 0103.D-0545, 0104.D-0257(B), and 105.20CR.002).
** Corresponding author; marilyn.latour@uni-goettingen.de

visible not only on the chromosome map but also on their color-magnitude diagram (CMD) as a red RGB, which, in some cases, also connects with a fainter subgiant branch (SGB). The most notorious case is ω Centauri, but other GCs such as NGC 1851 and NGC 5286 also have a distinct red RGB and a faint SGB. These clusters were termed Type II GCs (also anomalous or iron-complex GCs; Johnson et al. 2015), and the stars belonging to their red RGB are typically enriched in elements produced via the slow neutron-capture process (s elements), such as Ba and La, and possibly also in iron (see, e.g., Marino et al. 2015 their Table 10), although the latter point is still a matter of debate (see Carretta & Bragaglia 2023 and references therein).

What was unexpected was the realization that in the majority of the GCs from Milone et al. (2017), the P1 stars themselves have a spread in $F275W - F814W$ that is larger than expected from measurement errors. This means that even this "primordial" population is not consistent with a simple stellar population. Spectroscopic analyses of the P1 stars present on chromosome maps were scarce at the time, and no particular chemical species could be found to explain such a color spread. Based on the photometric properties of the P1 stars in three GCs whose chromosome maps have different color spreads, Lardo et al. (2018) found that a spread in He abundance and a small range of N abundances could explain the extended color distribution of the primordial population. An extensive comparison of abundance values from the literature with the position of stars on the chromosome maps of 29 GCs performed by Marino et al. (2019a) showed no evidence of light-element variations among the P1 stars, and the authors suggested instead that variations in iron or helium could explain the color spread. Dedicated spectroscopic observations of P1 stars in NGC 2808 also resulted in no evidence of light-element variations (Cabrera-Ziri et al. 2019). Additional spectroscopic investigations of P1 stars in NGC 3201 suggested that a small spread in iron (by ∼0.1 dex) is present among these stars and could explain their pseudo-color distribution on the chromosome map (Marino et al. 2019b). Further spectroscopic and photometric investigations supported this hypothesis (e.g., Husser et al. 2020; Lardo et al. 2022), and the presence of a small iron spread among the primordial stellar population is now strongly favored, as opposed to He variations (see also Tailo et al. 2019). Iron variations among P1 stars have been measured from high-resolution spectroscopy for a handful of stars in three GCs so far, NGC 3201, NGC 2808, and NGC 104 (47 Tuc). The iron variations were found to be in the range 0.1 to 0.15 dex (Marino et al. 2019b, 2023; Lardo et al. 2023). Abundances of up to 24 atomic species were also measured in the stars of NGC 3201 and NGC 104, and for most species a positive correlation with the pseudo-color of the star was found, suggesting that not only does iron vary, but the overall stellar metallicity as well. Finally, by comparing the photometric properties of the P1 stars, essentially the width of their pseudo-color distribution on the chromosome maps, with isochrones of varying metallicity, Legnardi et al. (2022) estimated metallicity variations in 55 GCs. They found a wide range of values: from less than 0.05 to ∼0.30 dex.

The presence of metallicity (or iron) spreads within P1 stars has implications for our view of GCs as a whole. Apart from a few particular cases of massive clusters with a clear iron spread, and possibly an age spread as well, such as ω Cen, NGC 6715 (M 54), and Terzan 5 (see, e.g., Johnson & Pilachowski 2010; Alfaro-Cuello et al. 2019; Ferraro et al. 2009; Pfeffer et al. 2021), the traditional view is that GCs do not show a significant spread in iron abundances. For example, Carretta et al. (2009a) established an upper limit of 0.05 dex for iron spread among 19 GCs based on direct iron measurements from high-resolution spectroscopy. The recent catalog of iron dispersion compiled in Bailin (2019), based on selected iron abundances from the literature, shows that iron spreads among GCs are indeed modest (<0.1 dex) but significantly different than zero in the majority of cases. Yong et al. (2013) also measured a small (0.03 dex) but significant metallicity spread (for iron and a dozen additional species) among RGB stars in NGC 6752 using high-precision differential abundance measurements. Measuring such a small-amplitude internal iron spread is notoriously difficult. At times, studies have claimed to find large iron spreads (>0.1 dex) in some GCs, but they were often not corroborated by further investigations. Even the issue of whether the P3 population in some type II GCs is enhanced in iron or not is still a matter of debate (Mucciarelli et al. 2015b; Carretta & Bragaglia 2022; Vargas et al. 2022; McKenzie et al. 2022). Artificially large iron spreads can be caused by a few factors, such as the inclusion of asymptotic giant branch (AGB) stars among a sample of RGB objects (Mucciarelli et al. 2015a), the method used to determine the surface gravity of the stars (Mucciarelli et al. 2015b), or even the presence of intrinsic luminosity variations in some stars (Albornoz et al. 2021).

The presence of intrinsic iron variations is of particular interest for the theoretical modeling of GC formation. It is commonly assumed that massive star clusters form hierarchically, via the merging of smaller subclusters (e.g., see Kruijssen 2014, and references therein). In this scenario, iron variations within clusters could point to chemical inhomogeneities of the interstellar medium within the spatial scales on which massive clusters form. Alternatively, it has been suggested that clusters self-enrich, for example via core-collapse supernovae (SNe) of the first and most massive stars formed in the cluster when the least massive stars are still in their formation process (Morgan & Lake 1989; Wirth et al. 2021). In line with the hierarchical scenario of cluster formation, it has been argued that subclusters produce various levels of iron enrichment before merging to form the final GC (Bailin 2018; McKenzie & Bekki 2021). Such modeling of GC formation can reproduce, qualitatively, the mass-metallicity relationship observed among the massive metal-poor GCs ([Fe/H] ~ -1.5) around galaxies (see, e.g., Bailin & Harris 2009; Strader & Smith 2008; but see Usher et al. 2018 for an alternative explanation). However, the extent to which even massive star clusters can self-enrich is still debated. Star formation is expected to be suppressed by stellar winds even before the onset of SNe (Smith et al. 2021), and young massive clusters in the local Universe are indeed found to be gas-free (Whitmore et al. 2014; Cabrera-Ziri et al. 2015). Recent hydrodynamical simulations (e.g., Lahén et al. 2024) find that while enrichment in light elements through stellar wind material from short-lived massive stars appears feasible, the clusters are not significantly enriched in iron or other heavy elements via SNe.

In this work we made use of our MUSE spectral database of GC stars to investigate the topic of metallicity dispersion and variations within P1 stars. We selected P1 stars from the RGB of 21 GCs based on their position on the chromosome maps built from HST photometric catalogs (Piotto et al. 2015; Nardiello et al. 2018), and we measured the metallicity of the selected RGB stars from the MUSE spectra. We want to make a clear distinction here between metallicity and iron abundance as these terms are sometimes used interchangeably in the literature. What is measured from our MUSE spectra is an overall solar-scaled

Table 2. Parameters of the $\Delta_{F275W,F814W}$ pseudo-color−metallicity relationship and metallicity spread derived for the P1 stars.

Cluster	*N*star (1)	[M/H]err (2)	ΔC (3)	R_P (4)	*p*-value (5)	*a* (6)	*b* (7)	Δ[M/H] (8)
NGC 104	226	0.008	0.29	0.58	7.96e-22	0.21 ± 0.03	−0.787 ± 0.003	0.061 ± 0.008
NGC 1851	127	0.011	0.21	0.40	4.29e-06	0.25 ± 0.10	−1.169 ± 0.011	0.053 ± 0.016
NGC 2808	152	0.018	0.33	0.57	3.35e-14	0.41 ± 0.08	−1.072 ± 0.008	0.137 ± 0.021
NGC 3201	30	0.007	0.25	0.69	2.29e-05	0.38 ± 0.15	−1.449 ± 0.018	0.095 ± 0.029
NGC 362	113	0.017	0.12	0.21	2.54e-02	0.32 ± 0.20	−1.159 ± 0.009	0.039 ± 0.020
NGC 5286	115	0.023	0.29	0.67	3.27e-16	0.51 ± 0.10	−1.585 ± 0.016	0.148 ± 0.023
NGC 5904	113	0.015	0.20	0.63	7.36e-14	0.41 ± 0.09	−1.296 ± 0.006	0.083 ± 0.015
NGC 6093	234	0.025	0.20	0.24	2.02e-04	0.15 ± 0.09	−1.647 ± 0.008	0.030 ± 0.014
NGC 6218	64	0.015	0.13	0.22	7.55e-02	0.17 ± 0.18	−1.319 ± 0.011	0.022 ± 0.018
NGC 6254	83	0.016	0.23	0.68	1.51e-12	0.48 ± 0.13	−1.497 ± 0.009	0.110 ± 0.023
NGC 6362	33	0.025	0.21	0.38	2.76e-02	0.24 ± 0.17	−1.090 ± 0.013	0.050 ± 0.028
NGC 6388	97	0.018	0.80	0.81	1.45e-23	0.30 ± 0.04	−0.471 ± 0.012	0.240 ± 0.025
NGC 6441	148	0.018	0.48	0.35	1.34e-05	0.14 ± 0.07	−0.406 ± 0.010	0.067 ± 0.026
NGC 6541	255	0.017	0.13	0.16	1.12e-02	0.11 ± 0.08	−1.727 ± 0.006	0.014 ± 0.009
NGC 6624	68	0.017	0.36	0.48	3.33e-05	0.15 ± 0.09	−0.761 ± 0.013	0.054 ± 0.025
NGC 6656	72	0.014	0.23	0.55	5.14e-07	0.56 ± 0.17	−1.715 ± 0.035	0.132 ± 0.032
NGC 6681	38	0.024	0.17	0.02	8.99e-01	−0.05 ± 0.22	−1.546 ± 0.014	−0.008 ± 0.028
NGC 6752	92	0.012	0.16	0.72	5.50e-16	0.71 ± 0.11	−1.499 ± 0.006	0.117 ± 0.014
NGC 7078	220	0.033	0.17	0.47	2.86e-13	0.71 ± 0.15	−2.195 ± 0.010	0.122 ± 0.021
NGC 7089	171	0.019	0.24	0.70	1.23e-26	0.66 ± 0.09	−1.484 ± 0.008	0.156 ± 0.018
NGC 7099	67	0.021	0.10	0.05	6.63e-01	0.22 ± 0.45	−2.172 ± 0.020	0.021 ± 0.034

Notes. (1) Number of stars included in the P1 sample. (2) Median value of the errors on [M/H]. (3) $\Delta_{F275W,F814W}$ spread of the P1 stars. (4) Pearson correlation coefficient between $\Delta_{F275W,F814W}$ and [M/H] and its (5) *p*-value. (6) *a* and (7) *b* coefficients from the linear relationship derived as in $\Delta_{F275W,F814W} = a$[M/H] + *b* and their 95% confidence interval as uncertainties. (8) Metallicity spread.

metallicity[1] and not an iron abundance from individual iron lines. Thus, we always refer to it as metallicity ([M/H]) and we keep the use of iron abundance ([Fe/H]) for direct measurements from iron lines, unless specified otherwise.

The paper is organized as follows. In Sect. 2 we describe our spectroscopic data, and we explain the different aspects of our methodology in Sect. 3. Our results on the metallicity spread among the P1 stars and on the metallicity dispersions are presented in Sect. 4 and discussed in Sect. 5. We briefly conclude in Sect. 6.

2. Observational data

Our target stars were observed with MUSE as part of the guaranteed time observation (GTO) program dedicated to GCs (PI: S. Dreizler, S. Kamann). Of the 25 Galactic GCs observed for the survey, 20 (listed in Table 2) have the photometric data necessary to build their chromosome maps and a sufficient number of RGB spectra to perform our analysis. In addition to the GCs observed with the GTO program, we also used the MUSE observations of NGC 6362 taken as part of GO time (Prop ID: 0103.D-0545, PI: Dalessandro). MUSE is an integral field spectrograph mounted on UT4 of the Very Large Telescope and is in operation since 2014 (Bacon et al. 2010). It features a wide-field mode (WFM) with a field of view of 1′×1′ at a sampling of 0.2″ per pixel and a narrow-field mode (NFM) covering a smaller field of view (7.5″×7.5″) at a sampling of 0.025″ per pixel. Both modes result in a spectral coverage of 4750−9350 Å, with a spectral resolution $R \sim 3000$ although this varies slightly across the wavelength range (see Husser et al. 2016). The GC GTO program targeted the central region of the clusters, covering approximately up to their half-light radii. Depending on the cluster, a varying number of WFM pointings were used to cover the area of interest. Observations for the GTO program were taken between 2014 and 2022, most of the observations taken after mid-2017 made use of the adaptive optics (AO) system installed on UT4, whenever the observing conditions allowed it. In addition to the WFM pointings made for each GC, ten GCs have one or more additional NFM (with AO) observations located at their very center (see, e.g., Göttgens et al. 2021). However, because of the small field of view of the NFM, these observations contribute only a very small amount of RGB spectra per cluster.

The data reduction and spectral extraction processes have been described at length in previous papers using spectra from the Globular Cluster GTO program (e.g., Kamann et al. 2018, 2016) The basic data reduction of the MUSE datacubes is performed using the official MUSE pipeline (Weilbacher et al. 2020). The extraction of the individual stellar spectra is done with PAMPELMUSE[2] (Kamann et al. 2013, 2018) and relies on a reference source catalog to determine the position of each resolved star in the MUSE data. The photometric catalogs used for the GCs included in our study are from the ACS Survey of Globular Clusters (Sarajedini et al. 2007; Anderson et al. 2008).

3. Methods

3.1. Spectroscopic metallicities

The individual MUSE spectra of the RGB stars are fitted to derive atmospheric parameters using the Göttingen spectral

[1] Except for the α-enhancement factor.

[2] https://pampelmuse.readthedocs.io/en/latest/about.html

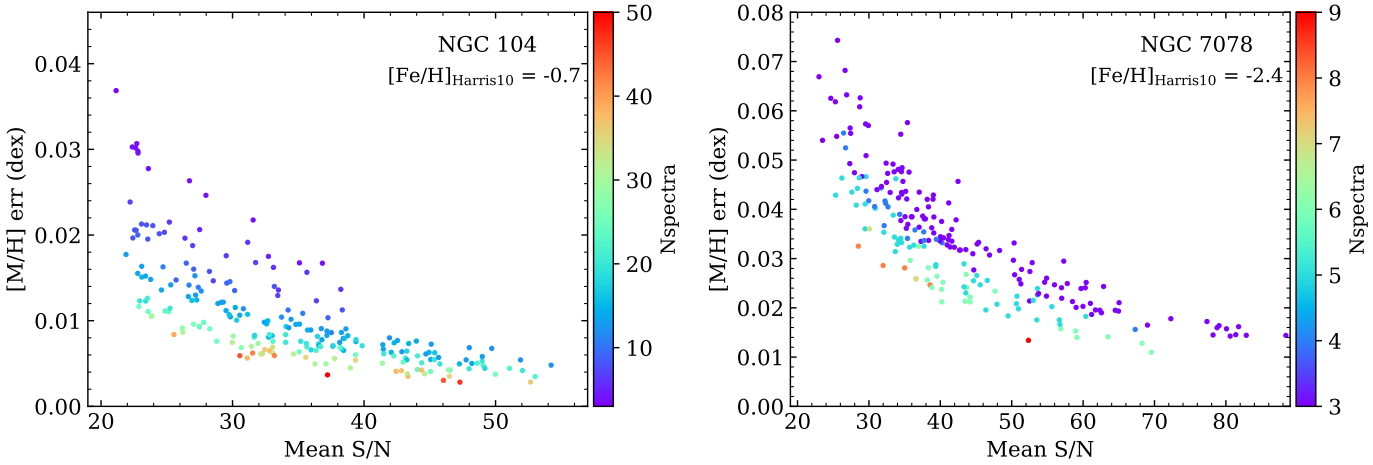

Fig. 1. Mean S/N versus the uncertainty on the metallicity ([M/H] err) for a subsample of RGB stars in NGC 104 and NGC 7078, color-coded by the number of spectra per star.

library of PHOENIX spectra (Husser et al. 2013) and the fitting framework *spexxy*[3] (Husser et al. 2016). Because the spectra of RGB stars are not very sensitive to changes in surface gravity (log g), this parameter is obtained from an isochrone. For each GC we find an isochrone (from Marigo et al. 2017) that best reproduces the HST photometry in the F606W-F814W versus F606W CMD, and we derived a value of log g and $T_{\rm eff}$ for each star by finding the nearest point on the isochrone. This was done with two sets of photometry: the ACS Survey of Globular Clusters (Sarajedini et al. 2007; Anderson et al. 2008) and the HST UV Globular Cluster Survey (HUGS; Piotto et al. 2015; Nardiello et al. 2018). The log g of the star is then fixed to the average value obtained from both sets of photometry during the spectral fit. The $T_{\rm eff}$ estimated from the isochrone is used as a starting value for the spectral fitting procedure and for selection criteria (see Sect. 3.3). The resulting best fit provides a value for the effective temperature ($T_{\rm eff}$), metallicity ([M/H]), radial velocity, a model for the telluric lines and a polynomial function describing the continuum. The metallicities [M/H] of the PHOENIX model grid are solar-scaled, except for the α elements that are enhanced to a value that is kept fixed during the fitting procedure. The α enhancement used varies from cluster to cluster and is between [α/Fe] = 0.1 and 0.4 (Dias et al. 2016). During the spectral fit, we did not target specific spectral regions but fit the whole spectral range, except the region containing the interstellar NaD lines and the AO gap (5780–5990 Å) in AO observations. Thus, what we measure from the spectral fit is the overall metallicity [M/H] and not a direct iron abundance. A more thorough description of the spectral fitting method is included in Husser et al. (2016) and Nitschai et al. (2023).

The next step is to combine, for a given star, the resulting parameters from the multiple observations to derive average values of $T_{\rm eff}$, radial velocity, and most important for this work, the metallicity ([M/H]). The general idea is to calculate the weighted average of the individual measurements (from the individual spectra), with the weight (w) being the inverse of the squared uncertainty (ϵ) returned by the Levenberg-Marquardt optimization routine used by *spexxy*. The resulting uncertainty on the weighted average is then

$$\sigma = \sqrt{\frac{1}{\sum w_i}}, \text{ where } w = \frac{1}{\epsilon^2}.$$

[3] https://github.com/thusser/spexxy

For a spectrum to be included in the calculation of the average parameters, two quality criteria are required: the spectrum must have a signal-to-noise ratio (S/N) > 20 and a magnitude accuracy above 0.8. To determine the magnitude accuracy parameter, PAMPELMUSE calculates the differences between the magnitudes recovered from the extracted spectra and the true magnitudes available in the photometric reference catalog. A value of 1 indicates that the magnitude difference determined for a star is fully consistent with the typical differences measured for similarly bright stars, whereas 0 indicates a strong outlier (see Sect. 4.4 in Kamann et al. 2018, for more details). The uncertainties on the [M/H] measurements are mainly defined by two factors. First, they scale with the mean S/N of the individual spectra because a higher S/N spectrum results in a smaller statistical uncertainty ϵ. Secondly, they depend on the number of measurements included in the average calculation. A final factor affecting the [M/H] uncertainties is the metallicity itself, meaning that the stars in low-metallicity GCs, such as NGC 7078 and NGC 7099, have larger uncertainties than stars at higher metallicity (like those of NGC 104) at a given S/N. The behavior of the uncertainties ([M/H] err) is illustrated in Fig. 1 where we show their relationship to the average S/N of the stars' spectra in two GCs: NGC 104 and NGC 7078. For illustration purposes, we include in these figures only the P1 stars included in the analysis described below. NGC 104 is relatively metal-rich ([Fe/H]=−0.7) and is one of the GCs with the largest number of observations, resulting in small typical uncertainties ([M/H] err < 0.015 dex). NGC 7078 is the most metal-poor GC in our sample ([Fe/H]=−2.4) and has fewer observations, typically below six measurements, resulting in larger uncertainties than in NGC 104. In the case of NGC 6362, we only have one measurement per star, so the uncertainties on [M/H] for that cluster are those returned by the Levenberg-Marquardt optimization routine of *spexxy*. The median values of [M/H] err for the P1 stars of each cluster are listed in Table 2.

3.2. Photometry and chromosome maps

The second essential ingredient for our analysis is the chromosome map of each GC. Our chromosome maps are constructed following the method described in Milone et al. (2017) and have been presented and used in previous work by our group (Husser et al. 2020; Martens et al. 2023; Latour et al. 2019). Before

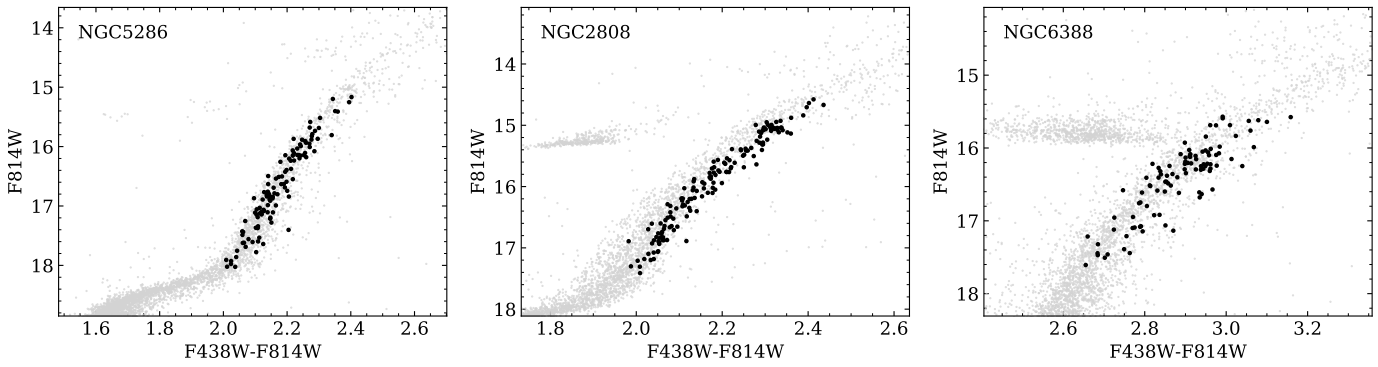

Fig. 2. CMDs of NGC 5286, NGC 2808, and NGC 6388. The P1 stars included in our final samples are shown with black dots. The small gray dots are the stars from the HUGS photometric catalog that have well-measured photometry (see Sect. 3.2).

making the chromosome maps, we cleaned the HST photometry following the procedure described in Sect. 3 of Nardiello et al. (2018, see also Milone et al. 2012). This procedure allowed us to select stars with well-measured photometry according to their photometric error and the shape and quality of their point spread function during the photometric extraction. We kept only the stars that pass the selection criteria for these three parameters in all four filters necessary to create the chromosome maps.

For our study, the chromosome maps are not necessary only to separate the stars into their respective population, but their exact pseudo-colors, especially $\Delta_{F275W,F814W}$ (the x-axis of the chromosome map), are important to look for metallicity trends. One improvement we made for this work is the inclusion of the differential reddening corrections provided by Legnardi et al. (2023) that is available for seven GCs in our sample (NGC 3201, NGC 5286, NGC 6254, NGC 6388, NGC 6441, NGC 6541, and NGC 6656). We also paid particular care to the fiducial lines defining the RGB envelope. Particularly, we removed the brightest RGB stars, down to two magnitudes (in F814W) below the tip of the RGB. This is to avoid the issue of poorly defined fiducial lines due to the small number of stars, which causes the pseudo-colors to be more uncertain for bright RGB stars. From the chromosome map, we separated the stars into their respective populations (P1 and P2), which is more or less straightforward depending on the cluster. In the Type II GCs, we also isolated the P3 stars (those from the red RGB), but these stars were not included in the analyses. The chromosome maps of the GCs in our study are shown in Appendix B along with the P1 and P2 stars included in our samples.

3.3. Sample selection

Our selection of stars from the chromosome maps is then matched with our sample of MUSE stars with [M/H] measurements. We applied several additional criteria to define our final sample of stars with the most reliable spectroscopic metallicity measurements and $\Delta_{F275W,F814W}$ position. We enumerate below the selection criteria.

- We kept stars that had at least three [M/H] measurements. The two exceptions are NGC 6681 and NGC 6362. For NGC 6681 we only have two observations of a single central pointing; in that case, we kept the stars that have two measurements. For NGC 6362 there is only one spectrum per star so we kept all the stars.
- We removed stars that have a radial velocity variability probability above 0.8, as defined in Sect. 5 of Giesers et al. (2019). This means that we removed from our sample the stars showing radial velocity variations, most likely due to binarity. Binaries were found to populate, among others, the blue extension of the $\Delta_{F275W,F814W}$ pseudo-color on the chromosome map (Kamann et al. 2020; Martins et al. 2020). While this is efficient at removing binaries in clusters with many observations, such as NGC 3201 and NGC 104, there are likely still binaries among the stars with a small number of observations. We note that this criteria was not applied to NGC 6362 because we have only one measurement per star.
- We kept only stars for which the difference between the $T_{\rm eff}$ estimated from the isochrone and the spectroscopic $T_{\rm eff}$ is within 3σ of the average $T_{\rm eff}$ difference between isochrone and spectroscopy. Effectively, this removed stars whose spectra returned a $T_{\rm eff}$ that is discrepant from their photometric properties (i.e., their position along the RGB). This can happen, for example, if the star is severely blended with a close neighbor causing contamination in the extracted spectra. This was done on a cluster-by-cluster basis.
- Similarly, we kept only stars for which the difference between the log g returned from both sets of photometry (ACS and HUGS) is within 3σ of the average difference. Effectively, this removed stars that have discrepant photometry between ACS and HUGS. If their position on the two CMDs is different, it results in different log g estimated from the isochrones. This was done on a cluster-by-cluster basis.
- We removed the brightest RGB stars, that is, those one to two magnitudes below the brightest star, depending on the cluster. The cut was made depending on how populated the RGB is; with fewer stars, the fiducial lines and the $\Delta_{F275W,F814W}$ positions are more uncertain. This cut also prevents contamination by AGB stars because they are found along the brightest part of the RGB. Finally, by limiting the range of magnitudes spanned by our stars, we also limited the extent of the metallicity-magnitude trend that affects the data (this is discussed in Appendix A).
- We crossmatched the stars with the HST proper-motion catalogs of Libralato et al. (2022) to remove any star that might not be a cluster member from its proper motion.
- We removed stars with a membership probability $p_{\rm member} < 0.5$, as defined in Kamann et al. (2016). Our membership probability is based on a combination of radial velocity and metallicity. While this is mostly redundant with the proper motion selection, it is still relevant for bright stars because some of them do not have proper motion measurements.

After applying this selection, we have between 30 and 250 P1 stars per cluster. The majority of them are fainter than the horizontal branch. In Fig. 2 we show the position of the P1 stars

in the $F438W$ versus $F814W$ CMD for three example clusters: NGC 5286, NGC 2808 and NGC 6388. For each star in our final samples, from P1 and P2, we provide their atmospheric parameters ([M/H], $T_{\rm eff}$, and, log g) and the pseudo-colors from our chromosome maps in Table 1, which is only available online from the CDS (see Sect. 6).

3.4. Metallicity variation and dispersion

To estimate the metallicity spread among the P1 stars of each cluster, we performed a linear regression in the $\Delta_{\rm F275W,F814W}$ − [M/H] plane. The individual measurements have varying precisions, depending on their errors, and we took that into account by performing a weighted least-square (WLS) regression, using the inverse of the squared error as weight. From the relationship obtained, we could estimate the metallicity variation for a given range of $\Delta_{\rm F275W,F814W}$ pseudo-color. To compute the pseudo-color extension (ΔC) of the P1 stars in each GC, we considered the range between the 4th and 96th percentile of the P1 stars in our final sample. The ΔC obtained from our spectroscopic sample is representative of that of the whole photometric sample on the chromosome maps because, as opposed to the previous studies using high-resolution spectroscopy (e.g., Lardo et al. 2018; Marino et al. 2023), we did not preselect the targets based on their position on the chromosome map. From the linear relation obtained with the WLS regression, we computed the metallicity predicted at the 4th and 96th pseudo-color percentiles and calculate the metallicity variation as Δ[M/H] = [M/H]$_{\rm 96th}$ − [M/H]$_{\rm 4th}$. We can consider here two different values as uncertainties, either from the standard errors returned on the parameters of the linear equation or from the 95% confidence interval of the regression. To be more conservative, in the following we compute and list uncertainties considering the 95% confidence interval.

We also computed the metallicity dispersion among our samples of stars. This was done using a Markov chain Monte Carlo ensemble sampler developed by Foreman-Mackey et al. (2013) to model our metallicity measurements with a Gaussian function. We then retrieved the most likely average metallicity and standard deviation (σ), with their uncertainties being the 16th and 84th percentile of the posterior probability distributions (equivalent to 1σ). This method takes the uncertainties on [M/H] into account and thus returns a standard deviation that should be representative, if the uncertainties are not significantly underestimated, of the intrinsic dispersion. We do not claim to measure an accurate intrinsic metallicity dispersion from our data, but we believe our values provide robust upper limits.

4. Results

4.1. Metallicity spread among the P1 stars

Our results are summarized in Table 2. For each GC we list the number of P1 stars included (Nstar), the median uncertainty on the stars' metallicity ([M/H]err) as an indicator of our measurements precision, the range in pseudo-color spanned by the P1 stars (ΔC, as described in Sect. 3.4), the Pearson correlation coefficient $R_{\rm P}$ and its associated p-value. We used the Pearson correlation coefficient because we expected a linear relationship. We also include the a and b coefficients of the relationship derived, as in $y = ax + b$, and their 95% confidence interval as uncertainties. Finally, we list the metallicity variations Δ[M/H] with their uncertainties. The first thing to examine is whether we find a significant correlation between metallicity and color. This is the case for 17 out of the 21 GCs. We measure a slope, a, that is consistent with zero, only for three clusters, namely NGC 7099, NGC 6681, and NGC 6218. These three clusters have a p-value larger than 7%. For all other clusters, the slope and $R_{\rm P}$ are positive, which demonstrates the sensitivity of the $\Delta_{\rm F275W,F814W}$ pseudo-color to changes in metallicity. We illustrate our results in Fig. 3 for four clusters: NGC 6752, NGC 104, NGC 3201, and NGC 6218. The results for the other clusters are shown in Appendix B.

The metallicity variations Δ[M/H] that we obtain are between ∼0.03 and 0.2 dex. In the hypothesis that the color extension of P1 is caused by metallicity variations, we expect a correlation between ΔC and Δ[M/H]. This is illustrated in Fig. 4. However, the relationship is not strictly linear but also shows a dependence on the average metallicity of the GC in the sense that for a given ΔC, we find the higher metallicity GCs to have smaller metallicity variations within their P1 stars. This behavior is illustrated with the results of Legnardi et al. (2022), shown in the right panel of Fig. 4. This clear dependence between the metallicity spread (Δ[M/H]) and the GC average metallicity is a natural consequence of the photometric method used by Legnardi et al. (2022) that relies on the color-metallicity relation of isochrones; a given metallicity increase, say by 0.05 dex, has a smaller impact on the F275W magnitude, where most iron and heavy-metal lines are found, at low metallicity than at larger ones. This dependence is also visible in our data for the higher metallicity clusters. We note here that even though the Legnardi et al. (2022) results are referred to as iron variations (Δ[Fe/H]), their isochrones were generated with different metallicities and not only with different iron abundances (Legnardi, priv. comm. 2024). We thus take the liberty to relabel their results as metallicity variations (Δ[M/H]).

It is worth mentioning here that NGC 6388, with its particularly large ΔC of 0.8, is not part of the Legnardi et al. (2022) sample because there is no estimate of a color spread for the P1 stars of that cluster. In any case, the RGB of NGC 6388 is especially wide in the $F275W$-$F814W$ color (see Milone et al. 2017, but also Carretta & Bragaglia 2022) and the pseudo-color − metallicity relation is conspicuous from Fig. B.3, this results in our largest Δ[M/H] of 0.24 dex.

4.2. Metallicity dispersion

Table 3 presents the average and dispersion values of the metallicities [M/H] obtained from our samples of P1 and P2 stars. The respective uncertainties correspond to 1σ. The number of stars included in each sample is also indicated. The P2 stars were first selected from their position on the chromosome maps (see Figs. B.6–B.8), and then filtered according to the criteria listed in Sect. 3.3. The metallicity dispersion of the P2 stars was computed in the same way as that of the P1 stars.

While we know that the chemical composition, at least in terms of light elements, is homogeneous in the P1 stars, this is not the case for the P2 stars. This could potentially affect our spectral fitting procedure because the PHOENIX models do not account for variations of individual atomic species like Na, O, Mg, and Al. These four elements have spectral lines, in the MUSE spectra, and their strength varies between populations (see, e.g., Latour et al. 2019 for line strength variations between the populations of NGC 2808). Admittedly, the wavelength ranges affected by these lines are small compared to the whole spectral coverage, but they could affect the metallicity of the best-fitting model. For example, if strong lines (essentially the magnesium triplet) are significantly stronger or weaker than the model predictions, it could influence the χ^2 minimum toward

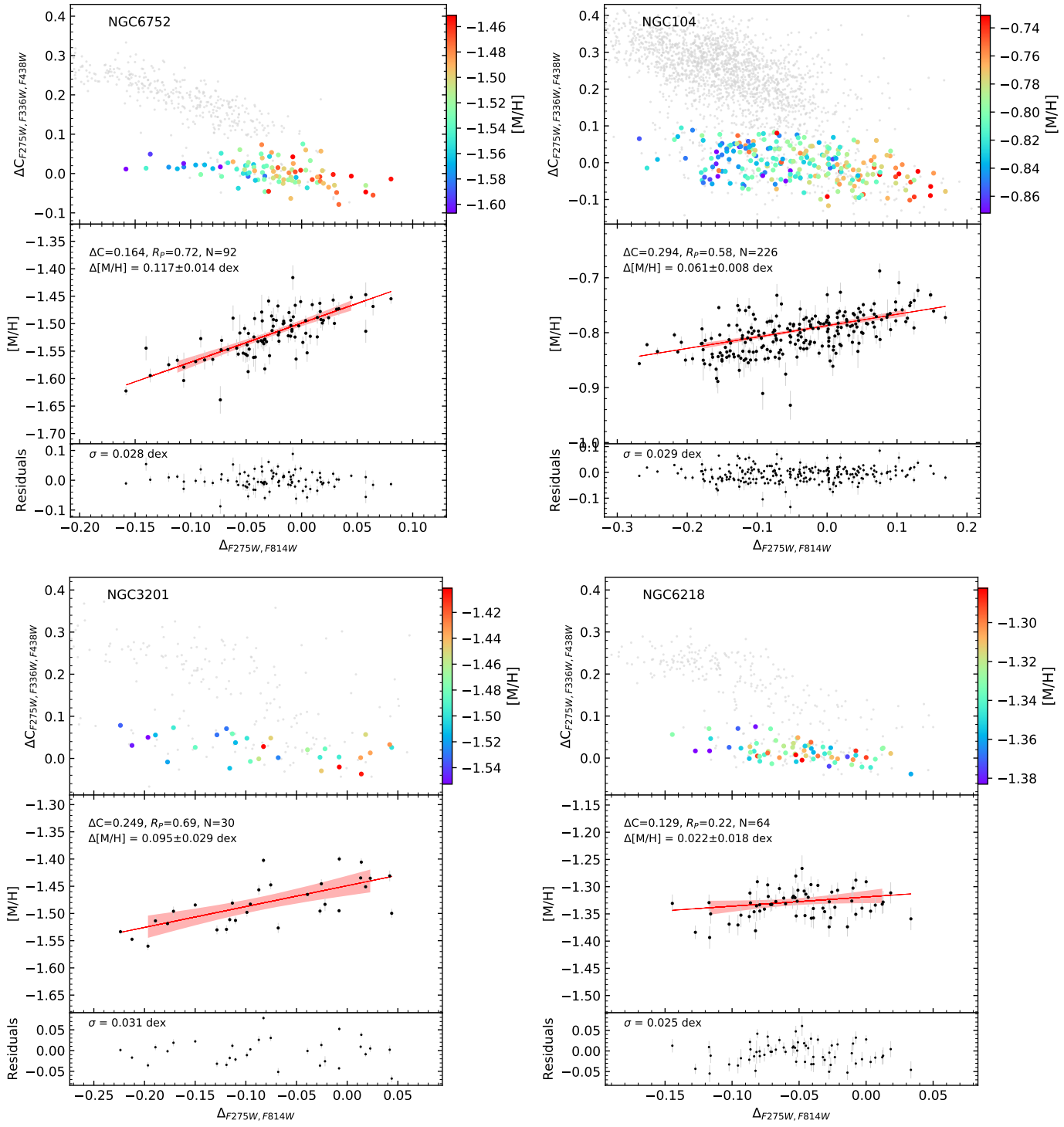

Fig. 3. Metallicity [M/H] versus $\Delta_{F275W,F814W}$ pseudo-color relationship for four GCs. Top panels: chromosome maps of the clusters with P1 stars within our sample color-coded by their metallicity. Middle panels: metallicity of each star, with the relationship derived from the WLS regression (red line) and the 95% confidence interval (red shaded area) over the color range ΔC. We also indicate the number of stars (N), the Pearson correlation coefficient (R_P), and the resulting metallicity variation Δ[M/H]. Bottom panels: residuals (i.e., observed minus predicted metallicity). We indicate the standard deviation (σ) of the residuals.

a higher or lower [M/H] value[4]. This could cause some small metallicity spread among the P2 stars.

We show the measured dispersion of the P2 stars versus that of the P1 stars in Fig. 5. In all but three clusters, the dispersion of the P2 stars is equal to or smaller than that of the P1 stars. If systematic effects were present in our MUSE sample, we would expect them to increase the metallicity dispersion of the P2 stars, as discussed above. The fact that we measure, in most clusters, a smaller metallicity dispersion among the P2 stars, is in line with

[4] We note here that the strong sodium lines from the NaD doublet are always masked during the spectral fit because they are blended with the lines from the interstellar medium. Thus, they do not influence the resulting metallicity.

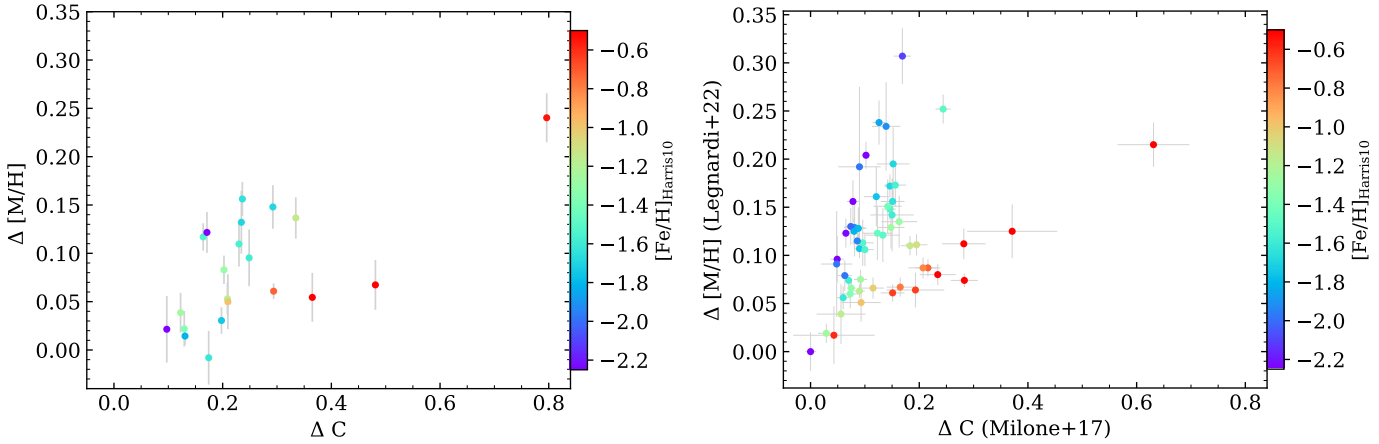

Fig. 4. Metallicity variations among the P1 stars versus their color extension for the GC included in our study (left panel) and in the study of Legnardi et al. (2022, right panel). Each point represents a GC and is color-coded by the average metallicity of the cluster (Harris, 2010 edition).

Table 3. Mean metallicity and dispersion measured in the P1 and P2 stars with their 1σ uncertainties.

Cluster	*N*stars	Mean [M/H]	σ [M/H]	*N*stars	Mean [M/H]	σ [M/H]	Mean [M/H]	σ [M/H]
		P1			P2		P1+P2	
NGC 104	226	–0.802 ± 0.002	0.033 ± 0.002	916	–0.802 ± 0.001	0.032 ± 0.001	–0.802 ± 0.001	0.033 ± 0.001
NGC 1851	127	–1.207 ± 0.003	0.036 ± 0.003	239	–1.210 ± 0.002	0.031 ± 0.002	–1.209 ± 0.002	0.033 ± 0.001
NGC 2808	152	–1.108 ± 0.005	0.056 ± 0.004	483	–1.096 ± 0.002	0.052 ± 0.002	–1.105 ± 0.003	0.055 ± 0.003
NGC 3201	30	–1.482 ± 0.008	0.045 ± 0.006	47	–1.474 ± 0.004	0.029 ± 0.003	–1.477 ± 0.004	0.036 ± 0.003
NGC 362	113	–1.178 ± 0.004	0.038 ± 0.003	281	–1.195 ± 0.002	0.037 ± 0.002	–1.191 ± 0.002	0.038 ± 0.002
NGC 5286	115	–1.663 ± 0.006	0.057 ± 0.005	174	–1.654 ± 0.004	0.046 ± 0.003	–1.658 ± 0.003	0.051 ± 0.003
NGC 5904	113	–1.307 ± 0.004	0.037 ± 0.003	355	–1.315 ± 0.001	0.022 ± 0.001	–1.313 ± 0.001	0.027 ± 0.001
NGC 6093	234	–1.661 ± 0.003	0.034 ± 0.003	324	–1.662 ± 0.002	0.027 ± 0.002	–1.662 ± 0.002	0.030 ± 0.002
NGC 6218	64	–1.330 ± 0.003	0.021 ± 0.003	97	–1.345 ± 0.003	0.022 ± 0.002	–1.339 ± 0.002	0.022 ± 0.002
NGC 6254	83	–1.519 ± 0.005	0.047 ± 0.004	171	–1.524 ± 0.002	0.026 ± 0.002	–1.523 ± 0.002	0.034 ± 0.002
NGC 6362	33	–1.101 ± 0.006	0.024 ± 0.007	18	–1.119 ± 0.007	0.012 ± 0.009	–1.108 ± 0.005	0.020 ± 0.006
NGC 6388	97	–0.541 ± 0.008	0.078 ± 0.006	312	–0.498 ± 0.004	0.064 ± 0.003	–0.507 ± 0.004	0.070 ± 0.003
NGC 6441	148	–0.433 ± 0.005	0.055 ± 0.004	374	–0.378 ± 0.004	0.083 ± 0.003	–0.394 ± 0.004	0.080 ± 0.003
NGC 6541	255	–1.736 ± 0.002	0.026 ± 0.002	242	–1.743 ± 0.002	0.027 ± 0.002	–1.740 ± 0.001	0.028 ± 0.001
NGC 6624	68	–0.786 ± 0.005	0.038 ± 0.004	175	–0.804 ± 0.003	0.029 ± 0.002	–0.799 ± 0.002	0.033 ± 0.002
NGC 6656	72	–1.831 ± 0.007	0.059 ± 0.005	107	–1.806 ± 0.006	0.065 ± 0.005	–1.816 ± 0.005	0.064 ± 0.004
NGC 6681	38	–1.547 ± 0.005	0.024 ± 0.005	180	–1.558 ± 0.003	0.032 ± 0.003	–1.556 ± 0.003	0.031 ± 0.002
NGC 6752	92	–1.518 ± 0.004	0.038 ± 0.003	220	–1.514 ± 0.002	0.027 ± 0.002	–1.517 ± 0.002	0.032 ± 0.002
NGC 7078	220	–2.242 ± 0.005	0.067 ± 0.004	329	–2.200 ± 0.003	0.053 ± 0.003	–2.237 ± 0.004	0.064 ± 0.003
NGC 7089	171	–1.515 ± 0.005	0.058 ± 0.004	670	–1.526 ± 0.002	0.038 ± 0.001	–1.524 ± 0.002	0.044 ± 0.001
NGC 7099	67	–2.184 ± 0.007	0.051 ± 0.005	150	–2.167 ± 0.003	0.028 ± 0.003	–2.172 ± 0.003	0.038 ± 0.002

the results of Legnardi et al. (2022) who found that the P1 main-sequence stars have a wider $F275W$-$F814W$ color spread than the P2 main-sequence stars in two particular GCs. The authors concluded that the metallicity spread among the P2 stars is lower than the metallicity spread of the P1 stars. Our results show that, at least for RGB stars, this conclusion can be extended to the massive GCs of the Milky Way. We note that this important result remains unchanged by the use of metallicities corrected for the magnitude (or $T_{\rm eff}$) trend discussed in Appendix A (see Fig. A.5). The three GCs for which we find a larger dispersion in P2 are NGC 6681, NGC 6656, and NGC 6441. In the last two, the differences between $\sigma_{\rm P2}$ and $\sigma_{\rm P1}$ is less than 2σ. In NGC 6656, there is one P2 star that is an outlier in terms of [M/H] (see Fig. A.1), removing this object lowers $\sigma_{\rm P2}$ very close to the value of $\sigma_{\rm P1}$. In NGC 6441, the difference is significant, but the separation of the populations is highly uncertain. The chromosome map of that cluster is peculiar and the RGB stars are essentially distributed continuously along a diagonal sequence so that it is not possible to see a distinction between P1 and P2 stars (see Fig. B.8), this is why Milone et al. (2017) did not estimate a $N_{\rm P1}/N_{\rm tot}$ ratio for NGC 6441. Our separation simply aimed at having a typical ratio of $N_{\rm P1}/N_{\rm tot} \approx 0.3$ (Milone et al. 2017).

5. Discussion

5.1. Comparison with literature results

As mentioned in the introduction, iron spreads among the P1 stars of GCs have been estimated from high-resolution spectroscopic measurements in only three GCs. We recall here that studies done with high-resolution spectroscopy are directly measuring iron abundances. Nonetheless, from the abundances of numerous atomic species in the P1 stars of NGC 104 and NGC 3201 it appeared that not only iron is varying but also the

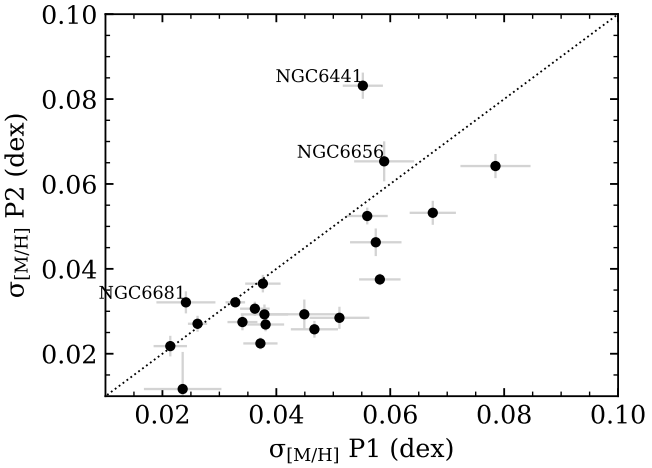

Fig. 5. Metallicity dispersion of the P2 stars versus that of the P1 stars. The error bars are 1σ uncertainties. The identity relation is drawn with a dashed line. The three GCs for which the P2 stars have a larger dispersion than the P1 stars are indicated, along with their names.

other elements studied (Marino et al. 2019b, 2023). A similar conclusion was made by Lardo et al. (2023) from their five P1 stars in NGC 2808. It is thus reasonable to assume that iron and metallicity variations go hand in hand. For NGC 2808, the linear relationship shown in Fig. 7 of Lardo et al. (2023), indicates an iron variation of about 0.15 dex among five P1 stars. For this cluster, we computed a Δ[M/H] of 0.14±0.02 dex for a similar ΔC as that of Lardo et al. (2023). In NGC 3201, Marino et al. (2019b) reported a difference of 0.1 dex among a dozen P1 stars spanning the same ΔC as in our sample. This is in perfect agreement with our Δ[M/H] of 0.095±0.03 dex. Finally, for NGC 104, we used our analysis method on the literature data (color and [Fe/H] for 21 P1 stars, Marino et al. 2023) and obtained a Δ[Fe/H] = 0.14±0.07 dex for a ΔC of 0.22. This is larger than our value of Δ[M/H] = 0.061 ± 0.01 dex, even though we have a larger color range (ΔC= 0.3). We note that because of its relatively high metallicity ([M/H] = −0.8), NGC 104 has a smaller Δ[M/H] than other more metal-poor clusters with a similar ΔC value in our sample (see also Fig. 4).

The MUSE spectra have also been used by Husser et al. (2020) to estimate the metallicity spread among the P1 stars on the RGB. In that work, the metallicities were estimated from an empirical calibration of the calcium triplet (CaT)−metallicity relationship. This means that the equivalent widths of the three Ca II lines ($\lambda\lambda$8498, 8542, 8662) and the stars' magnitude were used to derive the metallicity. Although the Husser et al. (2020) study is also based on the MUSE spectra, the method used is independent of ours because the metallicity derived from the CaT does not rely on any model atmospheres or atmospheric parameters such as $T_{\rm eff}$ and log g. The chromosome maps used in Husser et al. (2020) are slightly different than those presented here (see Sect. 3.2) and their analysis included stars almost up to the tip of the RGB. Besides a membership criteria, there was no further "cleaning" of the stellar sample. They calculated a metallicity spread in a similar way as done here, by computing the slope of the P1 stars in the [Fe/H]−$\Delta_{\rm F275W,F814W}$ plane (see their Fig. 21) but they used a smaller ΔC than in this work to compute their metallicity spread. In the left panel of Fig. 6 we show a comparison between our results and those of Husser et al. (2020). Here again, we relabel their [Fe/H] as [M/H] because the calibration of the CaT equivalent widths is made using the average metallicities of GCs. The agreement is very good; only for two clusters, namely NGC 6656 and NGC 7078, do we derive a significantly larger metallicity spread.

We also compared our results with those of Legnardi et al. (2022) for the 20 GCs in common (NGC 6388 is excluded) in the right panel of Fig. 6. As explained above, the authors used a different method to estimate metallicity variations, relying on photometry and isochrones instead of direct metallicity measurements and this results in a systematic shift between the two sets of results. In general, the metallicity spreads obtained from photometry are larger than our spectroscopic measurements, but they correlate well. At this point we do not know why the photometric method results in larger metallicity spread than our spectroscopic values.

Concerning the metallicity dispersions for the total P1+P2 populations (see Table 3 and Fig. 7), we find them to be generally very small, below 0.05 dex in the majority of our clusters. They are thus fully compatible with the conclusions reached by Carretta et al. (2009a), that $\sigma_{\rm [Fe/H]}$ is smaller than 0.05 dex in most of their 19 GCs, and by Bailin (2019), that $\sigma_{\rm [Fe/H]}$ is smaller than 0.1 dex with a median value of 0.045 dex for 55 GCs. However, we must keep in mind here that the P3 stars, which are possibly more metal-rich than those of P1 and P2, are excluded from our star samples in the Type II GCs (among our clusters those are NGC 362, NGC 1851, NGC 5286, NGC 6656, NGC 6388, and NGC 7089). So the total, in the sense of considering all RGB stars, metallicity dispersion in these six clusters might be larger. Few studies have measured intrinsic iron, or metallicity, dispersion below 0.05 dex, but Yong et al. (2013) conducted a high-precision differential abundance analysis of RGB stars in NGC 6752 and estimated an intrinsic iron and metallicity dispersion of ∼0.03 dex, which is fully compatible with our value of $\sigma_{\rm [M/H]}$ = 0.032 dex measured for this cluster. Our metallicity dispersion for NGC 362 of 0.038 dex is also in good agreement with the values obtained by Monty et al. (2023) of 0.035 and 0.041 dex (from Fe I and Fe II lines, respectively), using the same differential abundance analysis technique (see also Meléndez et al. 2009 for a description of the method).

NGC 6388 is a particular case, with its very large color spread on the chromosome map (ΔC=0.8) we measure a large Δ[M/H] of 0.24 dex and the color-metallicity relationship is clearly seen in Fig. B.3. It is also the cluster for which we derive the largest dispersion, 0.078 dex, among the P1 stars (0.07 dex for P1+P2). We note that the metallicity trends with magnitude and $T_{\rm eff}$ discussed in Appendix A are rather mild for this cluster, and by correcting the metallicities accordingly we only decrease the Δ[M/H] to 0.20 dex and the total dispersion (P1+P2) to 0.068 dex). This cluster has been classified as Type II by Milone et al. (2017) and is suspected to have some intrinsic iron spread (Wallerstein et al. 2007; Hughes et al. 2007). However, the recent studies by Carretta & Bragaglia (2022, 2023), based on a sample of about 150 bright RGB stars from all populations, concluded that there is no intrinsic iron (<0.04 dex) nor metallicity spread in the cluster, thus questioning its classification as a Type II GC. However, this would leave the significant color spread of the RGB stars in NGC 6388 unexplained.

5.2. Comparison with theoretical expectations

The studies of Bailin (2018) and McKenzie & Bekki (2021) simulated the formation of (proto)GCs and showed that self-enrichment leads to metallicity variations within the resulting GC. Although they use different types of simulations, their models predict a similar outcome in terms of intrinsic metallicity dispersion: an increase in dispersion with the cluster mass for

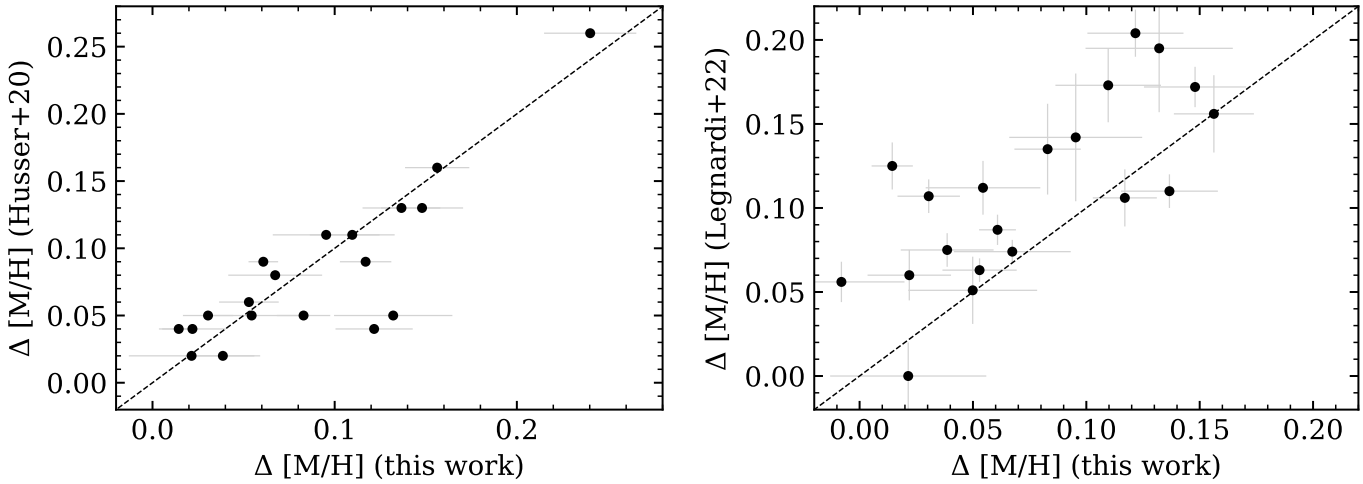

Fig. 6. Comparison between the Δ[M/H] found in this study and the Δ[Fe/H] obtained by Husser et al. (2020, left) and Legnardi et al. (2022, right). The identity relation is drawn with the dashed line.

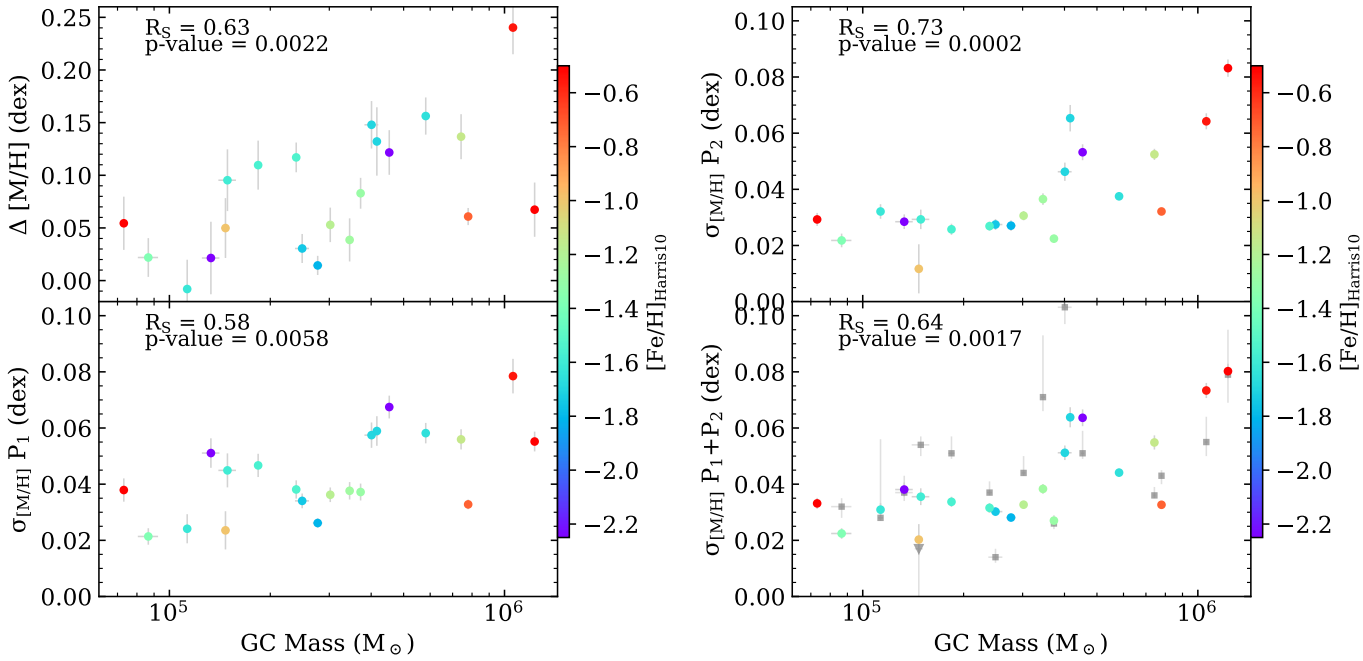

Fig. 7. Metallicity spread (Δ[M/H]) and dispersion ($\sigma_{[M/H]}$) versus the mass of the GCs (Baumgardt & Hilker 2018). The metallicity spreads (top left) are for the P1 stars as presented in Sect. 4.1. The other panels show the metallicity dispersions for the P1, P2, and P1+P2 stars. Our data points are color-coded by the average metallicity of the cluster (Harris, 2010 edition). The gray squares in the bottom-right panel represent the $\sigma_{[Fe/H]}$ taken from the literature compilation of Bailin (2019) for the GCs in common with our study. The gray downward triangle indicates the upper limit on $\sigma_{[Fe/H]}$ measured for NGC 6362.

$M > 10^5$ M$_\odot$. In the top-left panel of Fig. 7, we plot our resulting Δ[M/H] versus the mass of the GCs (from Baumgardt & Hilker 2018). We find a positive correlation between the two parameters with a Spearman's rank correlation coefficient $R_s = 0.63$. The majority of the GC in our sample have masses in the range 10^5–10^6 M$_\odot$ where the iron dispersion is expected to slightly increase with the cluster's mass and reach up to ∼0.1 dex (Bailin 2018; McKenzie & Bekki 2021). Our Δ[M/H] reach values that are larger than 0.1 dex, but they cannot be directly compared to the metallicity dispersions (σ) that are presented in the literature. Thus, in the bottom-left panel of Fig. 7 we plot the $\sigma_{[M/H]}$ of the P1 stars versus the GC masses. From this dataset, we retrieve a correlation coefficient $R_S = 0.58$, similar to that obtained from Δ[M/H], and our dispersion values are between 0.02 and 0.08 dex, in good agreement with the theoretical expectations.

We also show, in the right panels of Fig. 7, the metallicity dispersion of the P2 and P1+P2 stars versus the GC masses. Because the P2 stars dominate the total sample in most clusters, the behavior is very similar in the two cases: the dispersions are essentially constant and at their smallest ($\sigma_{[M/H]} < 0.04$ dex) until a mass of $\sim 3 \times 10^5$ M$_\odot$. At higher masses, the dispersion increases slightly up to ∼0.08 dex. This is qualitatively similar to what was measured by Carretta et al. (2009a) among their sample of 19 GCs using the absolute magnitudes of the clusters as a proxy for the GC mass. We note, however, that the correlation between iron dispersion and GC mass could not be reproduced

by Mészáros et al. (2020) with their spectra from the SDSS-IV APOGEE-2 survey, but their observed iron dispersions are systematically larger, with an average $\sigma_{[Fe/H]}$ close to 0.1 dex.

We also compared our dispersion values for P1+P2 with the $\sigma_{[Fe/H]}$ reported in the literature compilation of Bailin (2019) (updated in Bailin & von Klar 2022) for the GCs in our study (17 are in common, shown with gray symbols in Fig. 7). The two samples show a similar range of dispersion, although the correlation between the 17 Bailin (2019) data points and the GC mass is weaker (R_S = 0.44). It is worth specifying here that Bailin (2019) considered iron measurements from RGB stars regardless of their population, so it presumably includes a mix of P1 and P2 stars but also stars belonging to the red RGB in the case of type II clusters. The ratio of stars from each population included in their spectroscopic samples varies from one cluster to another.

In their photometric study of iron variations within P1 stars, Legnardi et al. (2022) observed a negative correlation between Δ[M/H] and the average GC metallicity. Within our sample, we do not have a correlation (R_S < 0.13) between the average metallicity of the GC and the metallicity spread (Δ[M/H]), nor any of the dispersions (P1, P2, or P1+P2). We note that there was also no such correlation among the observed $\sigma_{[Fe/H]}$ of Bailin (2019, see also Bailin & von Klar 2022) and no global correlation is expected from their theoretical models either. There is a relationship expected between metallicity and dispersion, but it is linked to the initial metallicity of the model (see Fig. 1 of Bailin & von Klar 2022). It essentially predicts a maximum possible dispersion that decreases with increasing metallicity.

6. Conclusion

We used the metallicities of RGB stars in 21 GCs to quantify the metallicity spread among the primordial stellar population of each cluster and characterize its intrinsic metallicity dispersion. The stellar metallicities were derived by fitting the MUSE spectra with model atmospheres, and, for each star, we averaged the results from at least three individual spectra with S/N > 20. The exceptions are NGC 6681 and NGC 6362, for which we had only two and one available measurement per star, respectively.

We characterized the relationship between the stellar metallicities, [M/H], and the pseudo-color, $\Delta_{F275W,F814W}$, which is the position of the star on the x-axis of the chromosome map. In almost all clusters, we observe a clear trend of increasing metallicity with increasing values of $\Delta_{F275W,F814W}$ among the P1 stars. From this relationship, we measured metallicity differences that mostly range from 0.03 to 0.15 dex. As expected, the metallicity variations are generally found to be larger in the GCs with a larger $\Delta_{F275W,F814W}$ color extension of their P1 stars. In this case, NGC 6388 stands out with, by far, the largest color spread among its P1 stars, resulting in a metallicity spread of 0.24 dex. Our metallicity spreads are in good agreement with those estimated in various literature works based on spectroscopy, but we find them to be, on average, systematically lower than those measured from photometry by Legnardi et al. (2022). We also find that the metallicity spread within the P1 stars correlates with the mass of the GCs.

In addition, we investigated the intrinsic metallicity dispersion ($\sigma_{[M/H]}$) of the P1 stars, the P2 stars, and the combination of the two (P1+P2). Our dispersion values are all relatively small (<0.08 dex) and comparable to, if not smaller than, the values presented in the literature based on high-resolution spectroscopy (Carretta et al. 2009a; Yong et al. 2013; Mészáros et al. 2020; Monty et al. 2023). For both the P1 and P2 stars we find a correlation between the intrinsic dispersion and the mass of the GC. This is particularly striking when looking at the metallicity dispersion of the P2 stars: the GCs less massive than $\sim 3 \times 10^5$ $M_\odot$ all have a similarly low dispersion (<0.04 dex), but for the more massive GCs, the dispersion increases with the cluster mass. Interestingly, we find that for all but one cluster, the metallicity dispersion of the P2 stars is lower than, or equal to, that of the P1 stars. This is in line with the results of Legnardi et al. (2022), who concluded, from photometry, that in NGC 6362 and NGC 6838, the metallicity spread among the P2 stars is lower than in the P1 stars. Our results show that this is a general tendency among the massive GCs of the Milky Way.

Even though the resolution of the MUSE spectra is too low to measure iron abundances from individual iron lines, we can achieve high precision in measuring small metallicity variations and dispersions thanks to the sheer amount of spectra accumulated over the last decade as part of the MUSE GTO Globular Cluster survey. In future work, we will explore in further detail the metallicities of the anomalous population (i.e., the red RGB) of Type-II GCs and some peculiarities among the P2 stars of selected GCs, such as NGC 104 and NGC 7078.

Data availability

Table 1 is available at the CDS via anonymous ftp to `cdsarc.cds.unistra.fr` (`130.79.128.5`) or via `https://cdsarc.cds.unistra.fr/viz-bin/cat/J/A+A/694/A248`.

The additional figures presented in Appendix B can be found online at `https://doi.org/10.5281/zenodo.14674172`.

Acknowledgements. We thank N. Bastian, M. Salaris, and S. Martens for useful comments and discussion. M.L. acknowledges funding from the Deutsche Forschungsgemeinschaft (grant LA 4383/4-1) and from the German Ministry for Education and Science (BMBF Verbundforschung) through grants 05A14MGA, 05A17MGA, 05A14BAC, 05A17BAA, and 05A20MGA. S.K. gratefully acknowledges funding from UKRI in the form of a Future Leaders Fellowship (grant no. MR/T022868/1). S.M. was supported by a Gliese Fellowship at the Zentrum für Astronomie, University of Heidelberg, Germany. This research has made use of the Astrophysics Data System, funded by NASA under Cooperative Agreement 80NSSC21M00561 and the PYTHON packages *pandas* (The pandas development team 2020; Wes McKinney 2010), MATPLOTLIB (Hunter 2007), and statsmodels (Seabold & Perktold 2010).

References

Albornoz, Á. L., Villanova, S., Cortés, C. C., Ahumada, J. A., & Parisi, C. 2021, AJ, 161, 76
Alfaro-Cuello, M., Kacharov, N., Neumayer, N., et al. 2019, ApJ, 886, 57
Anderson, J., Sarajedini, A., Bedin, L. R., et al. 2008, AJ, 135, 2055
Bacon, R., Accardo, M., Adjali, L., et al. 2010, SPIE Conf. Ser., 7735, 773508
Bailin, J. 2018, ApJ, 863, 99
Bailin, J. 2019, ApJS, 245, 5
Bailin, J., & Harris, W. E. 2009, ApJ, 695, 1082
Bailin, J., & von Klar, R. 2022, ApJ, 925, 36
Baratella, M., Prabhu, D. S., Lima, L., & Prugniel, P. 2022, A&A, 661, A138
Bastian, N., & Lardo, C. 2018, ARA&A, 56, 83
Baumgardt, H., & Hilker, M. 2018, MNRAS, 478, 1520
Cabrera-Ziri, I., Bastian, N., Longmore, S. N., et al. 2015, MNRAS, 448, 2224
Cabrera-Ziri, I., Lardo, C., & Mucciarelli, A. 2019, MNRAS, 485, 4128
Carretta, E., & Bragaglia, A. 2022, A&A, 659, A122
Carretta, E., & Bragaglia, A. 2023, A&A, 677, A73
Carretta, E., & Bragaglia, A. 2024, A&A, 690, A158
Carretta, E., Bragaglia, A., Gratton, R., D'Orazi, V., & Lucatello, S. 2009a, A&A, 508, 695
Carretta, E., Bragaglia, A., Gratton, R. G., et al. 2009b, A&A, 505, 117
Choi, J., Dotter, A., Conroy, C., et al. 2016, ApJ, 823, 102
Dias, B., Barbuy, B., Saviane, I., et al. 2016, A&A, 590, A9
Ferraro, F. R., Dalessandro, E., Mucciarelli, A., et al. 2009, Nature, 462, 483
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306

Giesers, B., Kamann, S., Dreizler, S., et al. 2019, A&A, 632, A3
Göttgens, F., Kamann, S., Baumgardt, H., et al. 2021, MNRAS, 507, 4788
Gratton, R., Bragaglia, A., Carretta, E., et al. 2019, A&A Rev., 27, 8
Gruyters, P., Nordlander, T., & Korn, A. J. 2014, A&A, 567, A72
Harris, W. E. 1996, AJ, 112, 1487
Hidalgo, S. L., Pietrinferni, A., Cassisi, S., et al. 2018, ApJ, 856, 125
Hughes, J., Wallerstein, G., Covarrubias, R., & Hays, N. 2007, AJ, 134, 229
Hunter, J. D. 2007, Comput. Sci. Eng., 9, 90
Husser, T.-O., Wende-von Berg, S., Dreizler, S., et al. 2013, A&A, 553, A6
Husser, T.-O., Kamann, S., Dreizler, S., et al. 2016, A&A, 588, A148
Husser, T.-O., Latour, M., Brinchmann, J., et al. 2020, A&A, 635, A114
Johnson, C. I., & Pilachowski, C. A. 2010, ApJ, 722, 1373
Johnson, C. I., Rich, R. M., Pilachowski, C. A., et al. 2015, AJ, 150, 63
Kamann, S., Wisotzki, L., & Roth, M. M. 2013, A&A, 549, A71
Kamann, S., Husser, T. O., Brinchmann, J., et al. 2016, A&A, 588, A149
Kamann, S., Husser, T. O., Dreizler, S., et al. 2018, MNRAS, 473, 5591
Kamann, S., Giesers, B., Bastian, N., et al. 2020, A&A, 635, A65
Kruijssen, J. M. D. 2014, Class. Quantum Gravity, 31, 244006
Lahén, N., Naab, T., & Szécsi, D. 2024, MNRAS, 530, 645
Lardo, C., Salaris, M., Bastian, N., et al. 2018, A&A, 616, A168
Lardo, C., Salaris, M., Cassisi, S., & Bastian, N. 2022, A&A, 662, A117
Lardo, C., Salaris, M., Cassisi, S., et al. 2023, A&A, 669, A19
Latour, M., Husser, T. O., Giesers, B., et al. 2019, A&A, 631, A14
Legnardi, M. V., Milone, A. P., Armillotta, L., et al. 2022, MNRAS, 513, 735
Legnardi, M. V., Milone, A. P., Cordoni, G., et al. 2023, MNRAS, 522, 367
Libralato, M., Bellini, A., Vesperini, E., et al. 2022, ApJ, 934, 150
Marigo, P., Girardi, L., Bressan, A., et al. 2017, ApJ, 835, 77
Marino, A. F., Milone, A. P., Karakas, A. I., et al. 2015, MNRAS, 450, 815
Marino, A. F., Milone, A. P., Renzini, A., et al. 2019a, MNRAS, 487, 3815
Marino, A. F., Milone, A. P., Sills, A., et al. 2019b, ApJ, 887, 91
Marino, A. F., Milone, A. P., Dondoglio, E., et al. 2023, ApJ, 958, 31
Martens, S., Kamann, S., Dreizler, S., et al. 2023, A&A, 671, A106
Martins, F., Morin, J., Charbonnel, C., Lardo, C., & Chantereau, W. 2020, A&A, 635, A52
McKenzie, M., & Bekki, K. 2021, MNRAS, 507, 834
McKenzie, M., Yong, D., Marino, A. F., et al. 2022, MNRAS, 516, 3515
Meléndez, J., Asplund, M., Gustafsson, B., & Yong, D. 2009, ApJ, 704, L66
Mészáros, S., Masseron, T., García-Hernández, D. A., et al. 2020, MNRAS, 492, 1641
Milone, A. P., Piotto, G., Bedin, L. R., et al. 2012, A&A, 540, A16
Milone, A. P., Marino, A. F., Piotto, G., et al. 2015, ApJ, 808, 51
Milone, A. P., Piotto, G., Renzini, A., et al. 2017, MNRAS, 464, 3636
Monty, S., Yong, D., Marino, A. F., et al. 2023, MNRAS, 518, 965
Morgan, S., & Lake, G. 1989, ApJ, 339, 171
Mucciarelli, A., Lapenna, E., Massari, D., Ferraro, F. R., & Lanzoni, B. 2015a, ApJ, 801, 69
Mucciarelli, A., Lapenna, E., Massari, D., et al. 2015b, ApJ, 809, 128
Nardiello, D., Libralato, M., Piotto, G., et al. 2018, MNRAS, 481, 3382
Nitschai, M. S., Neumayer, N., Clontz, C., et al. 2023, ApJ, 958, 8
Nordlander, T., Korn, A. J., Richard, O., & Lind, K. 2012, ApJ, 753, 48
Pfeffer, J., Lardo, C., Bastian, N., Saracino, S., & Kamann, S. 2021, MNRAS, 500, 2514
Piotto, G., Milone, A. P., Bedin, L. R., et al. 2015, AJ, 149, 91
Sarajedini, A., Bedin, L. R., Chaboyer, B., et al. 2007, AJ, 133, 1658
Seabold, S., & Perktold, J. 2010, in 9th Python in Science Conference
Smith, M. C., Bryan, G. L., Somerville, R. S., et al. 2021, MNRAS, 506, 3882
Strader, J., & Smith, G. H. 2008, AJ, 136, 1828
Tailo, M., D'Antona, F., Caloi, V., et al. 2019, MNRAS, 486, 5895
The pandas development team 2020, https://doi.org/10.5281/zenodo.13819579
Usher, C., Pfeffer, J., Bastian, N., et al. 2018, MNRAS, 480, 3279
VandenBerg, D. A., Richard, O., Michaud, G., & Richer, J. 2002, ApJ, 571, 487
Vargas, C., Villanova, S., Geisler, D., et al. 2022, MNRAS, 515, 1903
Wallerstein, G., Kovtyukh, V. V., & Andrievsky, S. M. 2007, AJ, 133, 1373
Weilbacher, P. M., Palsa, R., Streicher, O., et al. 2020, A&A, 641, A28
Wes McKinney 2010, in Proceedings of the 9th Python in Science Conference, eds. Stéfan van der Walt & Jarrod Millman, 56
Whitmore, B. C., Brogan, C., Chandar, R., et al. 2014, ApJ, 795, 156
Wirth, H., Jerabkova, T., Yan, Z., et al. 2021, MNRAS, 506, 4131
Yong, D., Meléndez, J., Grundahl, F., et al. 2013, MNRAS, 434, 3542

Appendix A: The metallicity−magnitude trend

Upon inspection of our data, we found that our metallicities ([M/H]) are not constant along the RGB: they increase with increasing luminosity. Given the relationship between the luminosity, $T_{\rm eff}$, and log g for RGB stars, this also means that the metallicities decrease with increasing log g and $T_{\rm eff}$. Similar issues were reported from the MUSE spectra of NGC 6397 obtained as part of the instrument commissioning (Husser et al. 2016; Baratella et al. 2022). From stellar evolution models including atomic diffusion, it is known that stars near the main-sequence turnoff display a lower metallicity due to the onset of diffusion processes as convection in the outer layers of the star ceases. Among other radiative transport processes, gravitational settling causes the heavier elements to sink at the bottom of the atmospheres, thus reducing the photospheric metallicity (VandenBerg et al. 2002; Nordlander et al. 2012). As the stars evolve on the SGB, convection reappears at the surface and the metallicity gradually returns to its original value as the heavy elements are brought back to the surface. According to stellar evolution calculations and isochrones that take atomic diffusion into account (e.g., Gruyters et al. 2014; Hidalgo et al. 2018; Choi et al. 2016), the original metallicity is still not fully recovered at the bottom of the RGB, so a difference in photospheric metallicity of ≈0.02-0.03 dex can be expected between the bottom and the middle part of the RGB. We note that this effect is expected to be larger at lower metallicities. The effect of atomic diffusion is nicely illustrated in Fig. 10 of Nitschai et al. (2023) where metallicity changes are seen in the measured metallicities from MUSE spectra in ω Centauri. Metallicity differences predicted from theoretical isochrones are also shown in that same figure.

The gradient seen in our data, in terms of [M/H] versus magnitude ($F814W$), is more or less important depending on the cluster. We show it for five GCs in Fig. A.1. In some clusters, like NGC 104, NGC 7089, and NGC 6752, the effect is mild (less than 0.03 dex difference between the faintest and brightest stars in our sample) and could be explained by atomic diffusion. But in some other cases, like in NGC 6441 and NGC 2808, the effect is much stronger than expected (more than 0.1 dex). We investigated our data and analysis method, but we could not find an explanation for this large [M/H]−magnitude trend. We do not see a correlation with the metallicity of the clusters, as is expected from diffusion. We show in Fig. A.2 the slope derived versus the average metallicity of the cluster. The three clusters with a slope larger than 0.04 are NGC 6656, NGC 2808, and NGC 6441.

To verify if this affects significantly our results, we computed "corrected" [M/H] values by removing the trend in the [M/H]−$F814W$ plane. For that, we fitted a linear relationship between [M/H] and $F814W$ for the P1 and P2 stars, individually. Then, we subtracted from the average metallicity the residual between the measured value of [M/H] and the predicted value from the linear relationship. Then we redid our analyses using the corrected metallicity for each star.

The metallicity spreads (Δ[M/H] corr) obtained with the corrected metallicities are consistent, within their uncertainties, with the results presented in the main text (see Fig. A.3). In the case of the metallicity dispersions, they decrease slightly but in most clusters, the decrease in dispersion is within the 3σ uncertainties (see Fig. A.4). The two exceptions are NGC 6441 and NGC 2808, those are the two GCs with the strongest [M/H]−magnitude gradient. The $\sigma_{\rm [M/H]}$ obtained with the corrected metallicities are significantly lower in these two clusters. For NGC 2808 $\sigma_{\rm [M/H]}$ decreases from 0.055 to 0.041 dex and for

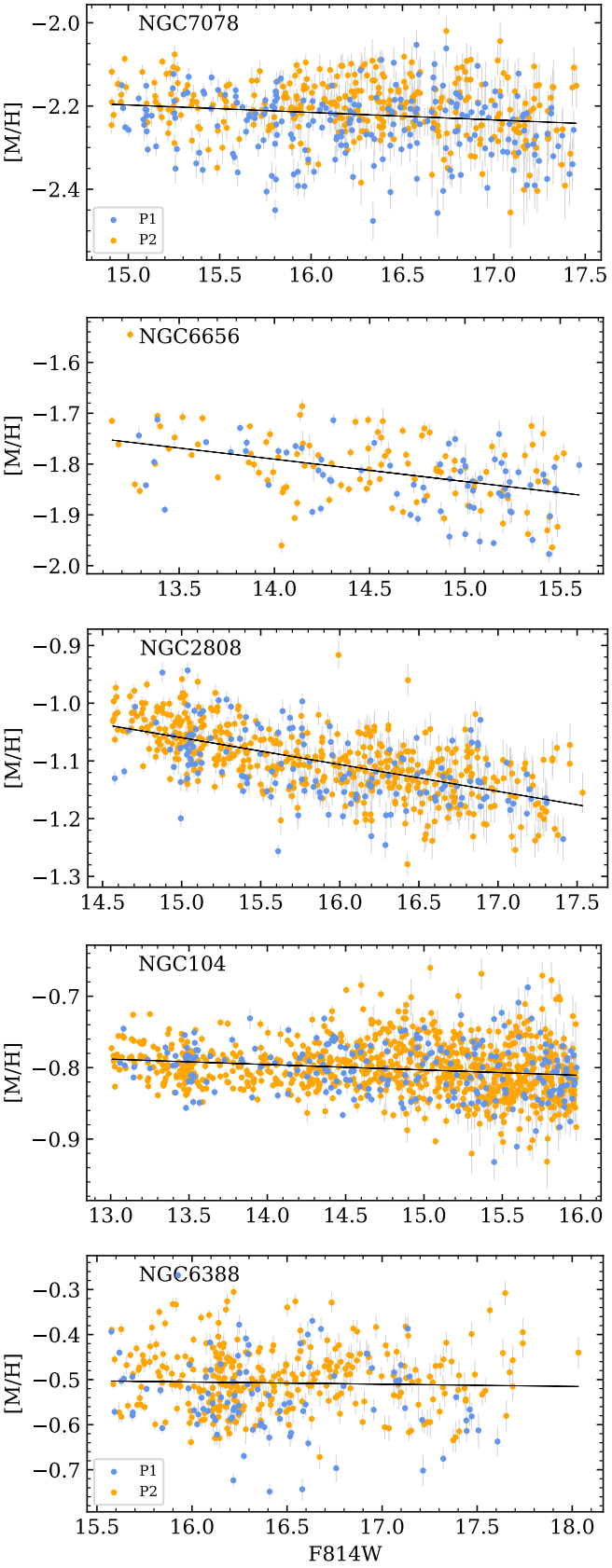

Fig. A.1. Metallicity−magnitude trend for five clusters spanning the metallicity range between [M/H]=−2.2 and −0.5 dex. The P1 and P2 stars are indicated by blue and orange dots, respectively. The black line indicates the best-fit linear relationship for all stars.

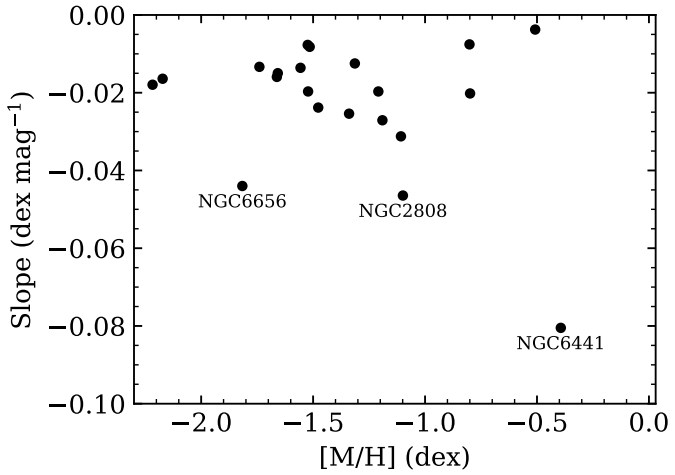

Fig. A.2. Slope of the [M/H]-$F814W$ relation versus the average [M/H] of the stars in our clusters.

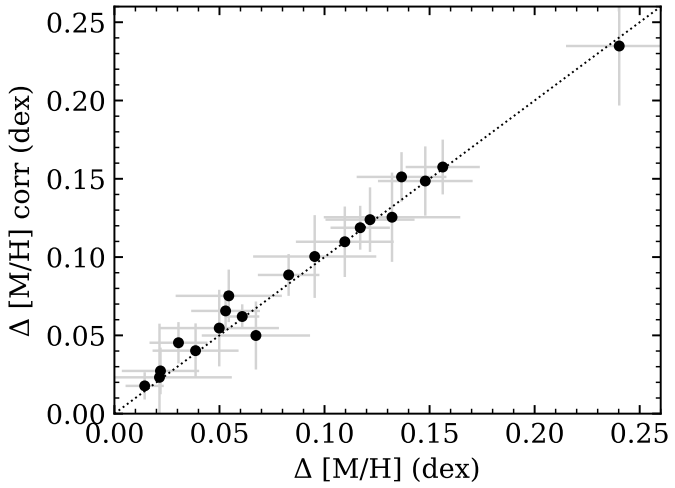

Fig. A.3. Comparison of the metallicity spreads, Δ[M/H] , obtained with the original and corrected metallicities.

NGC 6441 it decreases from 0.080 to 0.068 dex. When using the corrected metallicities, the P1 stars are still showing a larger dispersion than the P2 stars, except for the same three GCs (see Fig. A.5). Finally, the correlations between the GC masses and the metallicity spread (Δ[M/H]) and dispersion essentially remain the same whether we use the original or corrected metallicities. As a final remark, we mention that we also did this exercise of correcting individual metallicities to remove the [M/H]− $T_{\rm eff}$ gradient and the results are not significantly different than what is presented here except for the $\sigma_{\rm [M/H]}$ of NGC 6441 that decreases down to 0.062 dex.

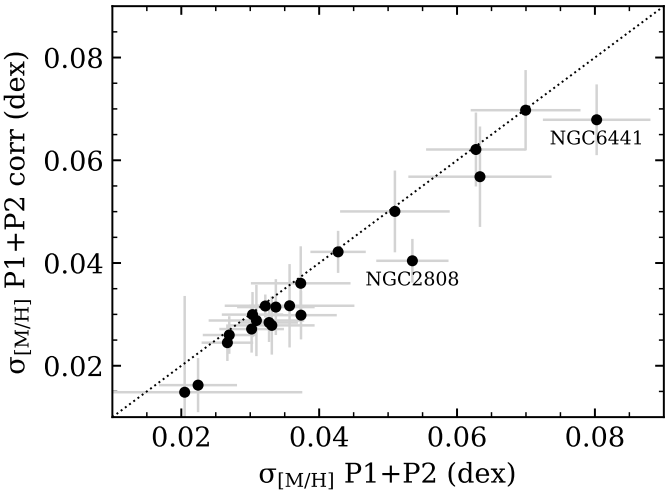

Fig. A.4. Comparison of the metallicity dispersions, $\sigma_{\rm [M/H]}$, obtained with the original and corrected metallicities.

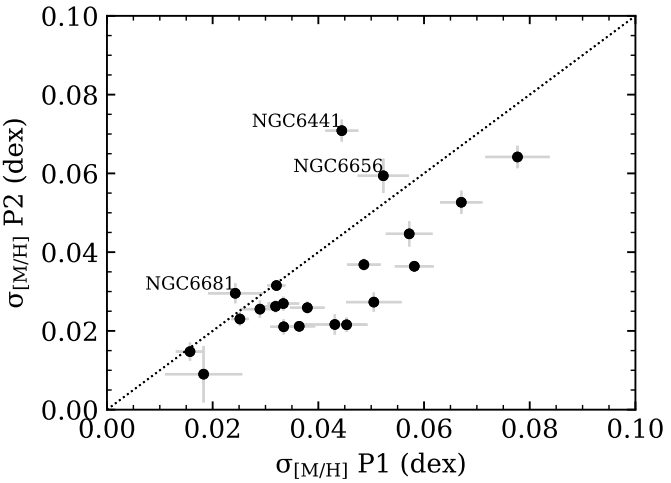

Fig. A.5. Same as Fig. 5 but for the dispersions computed with the corrected metallicities.