

How many components? Quantifying the complexity of the metallicity distribution in the Milky Way bulge with APOGEE

Alvaro Rojas-Arriagada¹,^{1,2}★ Gail Zasowski,³ Mathias Schultheis,⁴ Manuela Zoccali,^{1,2} Sten Hasselquist,³† Cristina Chiappini,⁵ Roger E. Cohen,⁶ Katia Cunha,^{7,8} José G. Fernández-Trincado,⁹ Francesca Fragkoudi,¹⁰ D. A. García-Hernández,^{11,12} Doug Geisler,^{13,14,15} Felipe Gran¹,^{1,2,16} Jianhui Lian,³ Steven Majewski,¹⁷ Dante Minniti,^{18,19} Antonela Monachesi¹,^{14,15} Christian Nitschelm²⁰ and Anna B. A. Queiroz⁵

¹Instituto de Astrofísica, Facultad de Física, Pontificia Universidad Católica de Chile, Avenida Vicuña Mackenna 4860, Santiago 8970117, Chile

²Millennium Institute of Astrophysics, Avenida Vicuña Mackenna 4860, Macul, Santiago 7820436, Chile

³Department of Physics & Astronomy, University of Utah, Salt Lake City, UT 84105, USA

⁴Observatoire de la Côte d’Azur, CNRS, Laboratoire Lagrange, Université Côte d’Azur, Blvd de l’Observatoire, F-06304 Nice, France

⁵Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany

⁶Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

⁷University of Arizona, Department of Astronomy and Steward Observatory, Tucson, AZ 85719, USA

⁸Observatório Nacional, Sao Cristóvão, Rio de Janeiro 20921400, Brazil

⁹Instituto de Astronomía y Ciencias Planetarias de Atacama, Universidad de Atacama, Copayapu 485, Copiapó 1531772, Chile

¹⁰Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str 1, D-85748 Garching, Germany

¹¹Instituto de Astrofísica de Canarias (IAC), E-38205 La Laguna, Tenerife, Spain

¹²Departamento de Astrofísica, Universidad de La Laguna (ULL), E-38206 La Laguna, Tenerife, Spain

¹³Departamento de Astronomía, Universidad de Concepción, Avenida Esteban Iturra s/n, Casilla 160-C, Chile

¹⁴Instituto de Investigación Multidisciplinario en Ciencia y Tecnología, Universidad de La Serena, Avenida Raúl Bitrán S/N, La Serena 1700000, Chile

¹⁵Departamento de Astronomía, Facultad de Ciencias, Universidad de La Serena, Avenida Juan Cisternas 1200, La Serena 1720236, Chile

¹⁶European Southern Observatory, Alonso de Cordova 3107, Vitacura, Casilla 19001, Santiago, Chile

¹⁷Department of Astronomy, University of Virginia, Charlottesville, VA 229044325, USA

¹⁸Departamento de Ciencias Físicas, Facultad de Ciencias Exactas, Universidad Andres Bello, Fernandez Concha 700, Las Condes, Santiago 7591538, Chile

¹⁹Vatican Observatory, I-V00120 Vatican City State, Italy

²⁰Centro de Astronomía (CITEVA), Universidad de Antofagasta, Avenida Angamos 601, Antofagasta 1270300, Chile

Accepted 2020 September 9. Received 2020 September 9; in original form 2020 July 27

ABSTRACT

We use data of $\sim 13\,000$ stars from the Sloan Digital Sky Survey/Apache Point Observatory Galactic Evolution Experiment survey to study the shape of the bulge metallicity distribution function (MDF) within the region $|\ell| \leq 11^\circ$ and $|b| \leq 13^\circ$, and spatially constrained to $R_{GC} \leq 3.5$ kpc. We apply Gaussian mixture modelling and non-negative matrix factorization decomposition techniques to identify the optimal number and the properties of MDF components. We find that the shape and spatial variations of the MDF (at $[\text{Fe}/\text{H}] \geq -1$ dex) are well represented as a smoothly varying contribution of three overlapping components located at $[\text{Fe}/\text{H}] = +0.32$, -0.17 , and -0.66 dex. The bimodal MDF found in previous studies is in agreement with our trimodal assessment once the limitations in sample size and individual measurement errors are taken into account. The shape of the MDF and its correlations with kinematics reveal different spatial distributions and kinematical structure for the three components co-existing in the bulge region. We confirm the consensus physical interpretation of metal-rich stars as associated with the secularly evolved disc into a boxy/peanut X-shape bar. On the other hand, metal-intermediate stars could be the product of *in-situ* formation at high redshift in a gas-rich environment characterized by violent and fast star formation. This interpretation would help us to link a present-day structure with those observed in formation in the centre of high-redshift galaxies. Finally, metal-poor stars may correspond to the metal-rich tail of the population sampled at lower metallicity from the study of RR Lyrae stars. Conversely, they could be associated with the metal-poor tail of the early thick disc.

Key words: infrared: stars – stars: fundamental parameters – stars: abundances – Galaxy: bulge – Galaxy: structure – Galaxy: stellar content.

★ E-mail: alvaro.rojas.astronomia@gmail.com, arojas@astro.puc.cl

† NSF Astronomy and Astrophysics Postdoctoral Fellow.

1 INTRODUCTION

Over the past decades, the bulge of the Milky Way (MW) has been an object of intense study and debate. A vast quantity of observations has contributed to an increasingly high-resolution, detailed characterization of its observed properties and their internal correlations. At the same time, considerable effort has been devoted to producing numerical simulations to provide theoretical prescriptions that link these observed properties with different mechanisms of bulge formation and evolution.

These efforts demonstrate recognition that the Galactic bulge is a key ingredient in our quest to understand galaxy formation and evolution. It represents the closest example of a mature bulge, which makes possible the detailed study of its resolved stellar content by means of photometric and spectroscopic measurements. The bulge provides a unique opportunity to understand the complex physics of baryons involved in galaxy formation, but at the same time represents a challenge since it requires dense coverage of at least 400 deg^2 of heavily extinguished sky.

The two broadest scenarios often invoked to explain bulge formation are motivated by observations of local external galaxies, and aim to account for their morphological dichotomy into the so-called *classical* and *pseudo-bulges* (Kormendy & Kennicutt 2004). Classical bulges are thought to be the product of mergers of primordial structures that yield spheroidal, massive (relative to the disc), kinematically hot, and isotropic structures, in a Lambda cold dark matter (Λ CDM) context (Baugh, Cole & Frenk 1996; Abadi et al. 2003). On the other hand, pseudo-bulges are the end product of the secular internal evolution of the disc, which rearranges angular momentum and stars to the centre and forms a bar, which subsequently undergoes vertical instabilities into an X-shaped boxy/peanut (B/P) bulge (Combes et al. 1990). In both cases, the mechanisms are dissipationless processes, driven by stellar dynamics (i.e. not gas dynamics) that assemble central structures from stars that have been previously formed elsewhere.

An additional, complementary scenario of bulge formation emerges from observational evidence at high redshift. Enhanced central star formation or even fully formed bulges have been observed in galaxies at $z \sim 2$, which otherwise are still vigorously forming stars in their massive, gas-rich discs (Tacchella et al. 2015; Nelson et al. 2016). This bulge assembly therefore takes place at an epoch preceding the formation of the bar, which for MW-mass galaxies happens at $z \sim 1$ (~ 8 Gyr ago; Sheth et al. 2008; Kraljic, Bournaud & Martig 2012; Fragkoudi et al. 2020). Characterized by a period of high star formation rate (SFR), the mechanism responsible here seems to produce bulges from the compaction of gas in a dissipative process, which is fast and forms stars *in situ*.

It is well established that the Galactic bulge hosts a bar and has a B/P morphology (e.g. Weiland et al. 1994; Dwek et al. 1995; Wegg & Gerhard 2013). *N*-body simulations indicate that such a structure is evidence of an origin in secularly evolved disc (Combes & Sanders 1981; Raha et al. 1991; Athanassoula 2005). This scenario has been reinforced and refined by the results of high-resolution hydrodynamical simulations, which, in addition to tracking the dynamical evolution of the particles, account for the detailed physics of a large number of baryonic processes (Gargiulo et al. 2019; Fragkoudi et al. 2020). Additional complexity has emerged as the outcome of a progressively larger amount of small programs and surveys mapping the stellar content of the inner Galaxy. In particular, it has become evident that this region hosts a variety of stellar populations or components, some of which have distinct kinematical or spatial distributions (e.g. Mira stars; Catchpole et al.

2016, RC stars; Zoccali et al. 2017, RR Lyrae; Kunder et al. 2020).

During the past 20 yr, enormous effort has gone into obtaining large samples of stellar spectra in the Galactic bulge with the aim of understanding the history of this massive component (for a review of these surveys, see Barbuy, Chiappini & Gerhard 2018). However, differences in target selection, stellar tracers, sightlines, and analysis methods long prevented a consistent picture of the bulge.

Starting with the pioneering low resolution spectroscopic study of Rich (1988) in Baade’s Window, the metallicity distribution function (MDF) and in general, the chemistry of K and M bulge giant stars started to be the object of dedicated observational efforts. Low resolution studies of the order of hundred stars (Minniti 1996b; Sadler, Rich & Terndrup 1996; Ramírez et al. 2000) were complemented with high resolution studies of the order of tens of stars, not only in Baade’s Window but also down to the high reddening regions close to the mid-plane, as near-infrared (NIR) spectroscopy was used (McWilliam & Rich 1994; Cunha & Smith 2006; Fulbright, McWilliam & Rich 2006; Cunha et al. 2007; Rich, Origlia & Valenti 2007).

The first homogeneous assessment of the MDF in several fields is that of Zoccali et al. (2008), who studied the MDF along the bulge’s minor axis (between $b = -4^\circ$ and -12°) by high-resolution observations of hundreds of stars per field, and confirmed the vertical metallicity gradient suggested earlier by Minniti et al. (1995); this was interpreted as a signature of classical bulge formation. Using a larger sample of red clump (RC) stars, Hill et al. (2011) revealed for the first time a bimodality in the MDF, the peaks of which are correlated with different kinematical signatures (Babusiaux et al. 2010). The metal-rich stars showed kinematics consistent with a bar-driven component, while the metal-poor ones had more isotropic kinematics that were associated with a classical spheroid. A similar metallicity bimodality was suggested from the analysis of the first available sample of dwarf, turn-off and subgiant bulge stars, observed during microlensing events (Bensby et al. 2011). In addition, this work suggested a different age–metallicity relation for metal-poor and metal-rich stars: While the former are generally old (~ 11 Gyr), the latter span a wide range of ages, with a significant fraction being younger than 9 Gyr. Together with the aforementioned different kinematic behaviour, these findings started to reveal an overall complex picture of the bulge stellar content.

Several subsequent studies, using different spectroscopic surveys of the bulge region, confirmed this bimodality and the relationship between metallicity and kinematics: e.g. *Gaia*-ESO Survey (GES; Rojas-Arriagada et al. 2014, 2017), GIRAFFE Inner Bulge Survey (GIBS; Gonzalez et al. 2015; Zoccali et al. 2017), and Apache Point Observatory Galactic Evolution Experiment (APOGEE; Ness et al. 2016; Zasowski et al. 2016, Schultheis et al. 2017; Fragkoudi et al. 2018; Queiroz et al. 2020b).

Using more than 10 000 stars from ARGOS, Ness et al. (2013) found that the MDF can be decomposed into up to 5 metallicity components, with 3 of them accounting for the majority of stars. Using the APOGEE Data Release (DR) 12 data, García Pérez et al. (2018) suggested the presence of four metallicity components that are of different strength. However, their sample did not include the coolest stars with temperatures below 3600 K. This complexity in the MDF structure, seen from the study of bulge giant stars, was reinforced by the results of the progressively larger sample of microlensed dwarf stars analysed in Bensby et al. (2013, 2017). Up to five peaks were identified in their MDF, with the metallicity locations being consistent with those detected in the MDF assembled from the larger ARGOS sample.

In this paper, we use the combined data of APOGEE-1 and APOGEE-2 (Section 2) to study the bulge’s MDF, its statistical properties, and correlations between stellar metallicity and kinematics over a large area of sky, including the still poorly explored inner degrees of the Galactic plane.

The structure of the paper is as follows: In Section 2, we describe the adopted data set and the computed set of spectro-photometric distances. We explore potential sources of bias in the MDF, as well as the appropriateness of the often adopted radial limit ($R_{GC} \leq 3.5$ kpc) used to select samples of bulge stars in Section 3. The shape of the bulge MDF and its spatial variations are explored in Section 4, where it is parametrized with a Gaussian mixture model analysis as well as a non-negative matrix factorization (NMF) decomposition. Finally, in Section 5 we discuss our results in the context of previous spectroscopic surveys of the bulge, as well as in the more general context of galactic formation.

2 DATA

2.1 APOGEE and *Gaia*

We use fundamental stellar parameters effective temperature (T_{eff}), surface gravity ($\log g$), and metallicity ($[M/H]$)¹ from the APOGEE (Majewski et al. 2017). APOGEE is a high-resolution, NIR spectroscopic survey designed to perform far-reaching chemical cartography of the MW stellar populations, using hundreds of thousands of stars. The main targets of the survey are giant stars (RGB, AGB, and RC), which are intrinsically luminous tracers present in nearly all stellar populations (Zasowski et al. 2013, 2017). By observing at NIR wavelengths, APOGEE overcomes much of the extinction imposed by the large amount of dust present in the Galactic plane, especially towards the Galactic bulge, which has limited past observational efforts in this important region of the Galaxy.

APOGEE, a component of both Sloan Digital Sky Survey-III (SDSS) and -IV (Eisenstein et al. 2011; Blanton et al. 2017), observes in the NIR H band (1.51–1.70 μm) using two custom-built, high-resolution ($R \sim 22\,500$) spectrographs at Apache Point Observatory’s 2.5 m Sloan Telescope and Las Campanas Observatory’s 2.5 m Irénée du Pont telescope (Bowen & Vaughan 1973; Gunn et al. 2006; Wilson et al. 2019). APOGEE spectra are extracted and wavelength calibrated, and radial velocities (RVs) are computed using the pipeline described in Nidever et al. (2015); stellar fundamental parameters and abundances of up to 26 elements (including alpha, iron-peak, odd-Z, and neutron-capture elements) are computed using the APOGEE Stellar Parameters and Chemical Abundances Pipeline (ASPCAP; García Pérez et al. 2016). An overview of the APOGEE parameter calibrations, data products, and elemental abundances can be found in Holtzman et al. (2018) and Jönsson et al. (2018).

The base catalogue for our sample selection has been reduced and analysed either with the APOGEE pipeline version used in SDSS DR16 (Ahumada et al. 2020; Jönsson et al. 2020) or a very similar one with a slightly updated data reduction version (r13). The full catalogue includes $\sim 134\,500$ additional stars observed after those released in DR16 (through 2019 November), and additional visits for $\sim 56\,000$ others.

¹Throughout this paper, we use $[M/H]$ as the estimate of metallicity, which results from the full-spectrum fitting performed by ASPCAP. These values are consistent with the $[\text{Fe}/\text{H}]$ measurements, so we compare our $[M/H]$ MDF components directly with the $[\text{Fe}/\text{H}]$ of the literature components.

We adopt proper motions for our sample from *Gaia* DR2 (Gaia Collaboration 2018). In addition, we adopt the renormalized unit weight error (RUWE) values for all the stars in our sample from the official release as available from the *Gaia* Archive. The RUWE is a recommended astrometric-quality diagnostic that can be used as a criterion to select good astrometric solutions. We do not use these values in selecting the bulge sample, but we include them in our study of the spatial variation of global kinematical properties of the sample (Section 3.2). On the other hand, we do not use *Gaia* parallaxes as they become error dominated for stars beyond 3–4 kpc, making them unsuitable to estimate distances for stars in the inner Galaxy. Instead, we compute spectro-photometric distances as described in the next section.

2.2 Distances and orbits

We calculated spectro-photometric distances for the whole set of stars available in the interim APOGEE catalogue (so, including data beyond the public DR16). To this end, we incorporated the stellar properties T_{eff} , $\log g$, $[\text{M}/\text{H}]$, and the 2MASS JHK_s photometry into the spectro-photometric method described in Rojas-Arriagada et al. (2017, 2019).

In summary, we compare a large set of theoretical isochrone points to the parameters T_{eff} , $\log g$, and $[\text{Fe}/\text{H}]$ of each star, and compute the distances from each star to each isochrone point in the theoretical space. These parameter-space distances serve as weights from which the most likely theoretical physical properties (e.g. the luminosity) of the star can be computed. A number of extra multiplicative weights are defined to account for the evolutionary speed of the points along the isochrones and for the initial mass function (IMF). Using these weights, the most likely absolute magnitudes (M_J , M_H , M_{K_s}) of the observed stars can be computed as the weighted mean or median of the theoretical values of the whole set of isochrone points. The computed absolute magnitudes are then compared to the observed photometry, allowing us to estimate the line-of-sight reddening and distance modulus. No prior on the Galactic stellar density is used. For these computations, we adopted a set of PARSEC² isochrones (version 3.1; Bressan et al. 2012; Marigo et al. 2017), spanning ages from 1 to 13 Gyr in steps of 1 Gyr and metallicities from -2.2 to $+0.5$ in steps of 0.1 dex.

In Appendix A, we validate our distances against other established metrics: the STARHORSE (Queiroz et al. 2018, 2020b) and ASTRONN (Leung & Bovy 2019) pipelines, *Gaia* Bayesian distances (Bailer-Jones et al. 2018), open clusters, and the distances to the Large Magellanic Cloud and Sagittarius dwarf galaxy. Overall, we find an approximate conservative uncertainty of ~ 25 per cent for the typical RGB stars that dominate our sample.

For the rest of the paper, we use Galactocentric cylindrical distances (R_{GC}) computed from our spectro-photometric heliocentric distances and the stellar (l , b) coordinates, assuming $R_{\odot} = 8.2$ kpc (Bland-Hawthorn & Gerhard 2016). We also adopt the Galactocentric velocity V_{GC} for our analysis, that is the heliocentric RV corrected for the solar reflex motion.

The combination of spectro-photometric distances, *Gaia* proper motions, and APOGEE RVs allows us to estimate orbital parameters by integrating orbits under a prescription of the Galactic potential. We use the GALPY³ code (Bovy 2015), adopting the MWPotentiaL2014 model for the MW gravitational potential. This model is

²Available at <http://stev.oapd.inaf.it/cgi-bin/cmd>.

³Available at <http://github.com/jobovy/galpy>.

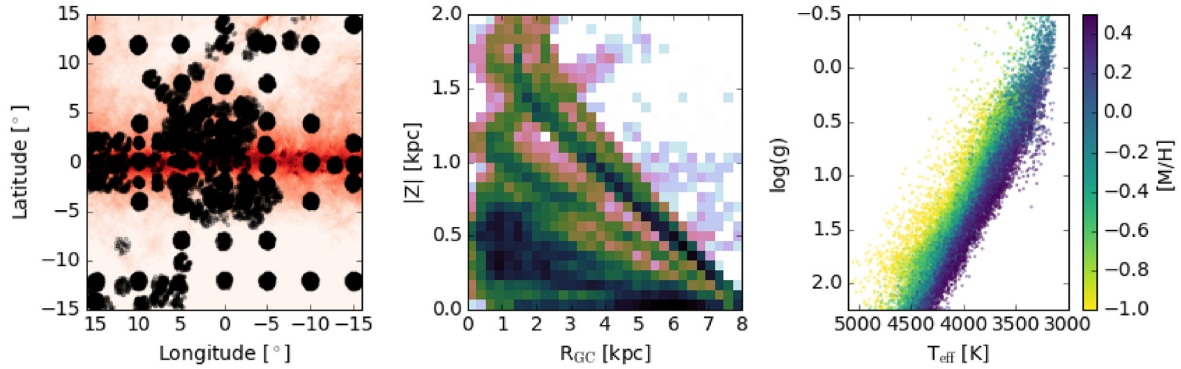


Figure 1. The clean general inner Galaxy sample as defined in Section 2.3. *Left:* Galactic (l , b) of the stars (black points) against the Schlegel, Finkbeiner & Davis (1998) $E(B - V)$ map (in red). *Centre:* Galactocentric radius (R_{GC}) and height from the mid-plane of the stars. *Right:* Kiel diagram, coloured by $[M/H]$.

a superposition of a Hernquist bulge (Hernquist 1990), a Miyamoto–Nagai disc (Miyamoto & Nagai 1975), and a Navarro–Frenk–White halo (Navarro, Frenk & White 1997), which contribute 5 per cent, 60 per cent, and 35 per cent of the rotational support at the solar circle, respectively.

From the full phase-space information available for each star [α , δ , $\mu_\alpha \cos(\delta)$, μ_δ , V_{GC} , d], the stellar orbits are integrated over 10 Gyr. In the calculations, the in-plane distance of the Sun from the Galactic Centre is adopted as $R_\odot = 8.2$ kpc, the velocity of the Local Standard of Rest (LSR) as $V_{LSR} = 220$ km s $^{-1}$ (see Bovy et al. 2012), and the peculiar velocity of the Sun with respect to the LSR as $(U, V, W)_\odot = (11.1, 12.24, 7.25)$ km s $^{-1}$ (Schönrich, Binney & Dehnen 2010). We estimate uncertainties for all computed orbital parameters by generating 600 random Gaussian realizations of the set of observed parameters from their respective individual uncertainties. From these variations of the initial conditions, we obtain distributions of the orbital parameters from which 1σ errors are estimated; these are used to restrict the sample solely to demonstrate the R_{GC} limit in Section 3.2.

2.3 Selection of a clean inner disc/bulge sample

Before examining the bulge MDF, we construct a clean sample of inner MW APOGEE stars with high-quality, reliable fundamental parameters but with no distance restrictions. We use this sample in Section 3 to determine the optimal R_{GC} limit for a bulge selection.

To select this clean inner sample, we start with the interim `allStar` file of all the APOGEE data observed through 2019 November (MJD 58814) and consider stars with $|\ell| \leq 16^\circ$ and $|b| \leq 15^\circ$. To ensure reliable ASPCAP values for the parameters of interest, we require a minimum signal-to-noise ratio (SNR) of 60 and that the following STARFLAG⁴ bits be set to zero: BAD_PIXELS, VERY_BRIGHT_NEIGHBOR, and PERSIST_HIGH. We require the NO_ASPCAP_RESULT and the STAR_BAD bits of the ASPCAPFLAG⁵ bitmask to be zero, and we remove stars with visit-to-visit RV variations (VSCATTER) greater than 0.6 km s $^{-1}$. The selected SNR limit includes stars below the APOGEE goal of SNR ≥ 80 per pixel, but it is set conservatively above the SNR = 50 limit over which ASPCAP provides reliable parametrizations (García Pérez et al. 2016). The VSCATTER limit is set to remove a minor

fraction of stars in the tail of the VSCATTER distribution, dominated by binary systems or stars with some other problem affecting their RV measurements. In addition to the previous cuts, we remove any stars that pass them but do not have valid ASPCAP values and/or spectro-photometric distances.

To eliminate biases due to non-standard target selection, we remove any star not flagged as a main survey target using the EXTRATARG⁶ bitmask. This leaves several inner Galaxy stars that were chosen as part of smaller APOGEE subprograms (in addition to the main survey); after evaluating those programs’ target selection and the resulting samples, we remove stars whose PROGRAM-NAME in the `allStar` file is set to `cluster_gc`, `clusters_gc1`, `geisler_18a`, `geisler_19a`, `sgr`, or `sgr_tidal`. The resulting sample after all these selections comprises 27 806 stars and is shown in Fig. 1.

3 DEFINITION AND BIAS CORRECTIONS OF OUR BULGE SAMPLE

The objective of this work is to study the shape of the MDF as an observational proxy of the complex mix of structures assembling the stellar content of the inner MW. In this context, there are two main issues we examine before defining a selection of bulge stars: potential sources of sampling biases in the APOGEE data (Section 3.1) and the most appropriate cylindrical distance cut to define the bulge spatial region (Section 3.2). We incorporate our findings into the definition of our bulge sample in Section 3.3.

3.1 Sample bias corrections

As we want to use the APOGEE spectroscopic data to study the underlying structure of the bulge MDF and its spatial variations, we must therefore account for potential sources of sampling biases. We calculated selection corrections for two different types of biases in our sample: (i) observational bias due to stars of different metallicities having different probabilities of being observed by APOGEE (P_{obs}), and (ii) analysis bias due to stars with some intrinsic metallicity– T_{eff} combinations not having reliable ASPCAP metallicities ($P_{[M/H]}$).

For the first of these – the observational probability – we simulated observing simple stellar populations (SSPs) as a function of $[Fe/H]$, distance, and extinction (see examples in Figs 2a and b). We generated SSPs based on 10 Gyr MIST isochrones (Paxton et al. 2011;

⁴https://www.sdss.org/dr16/algorithms/bitmasks/#collapseAPOGEE_STARFLAG

⁵https://www.sdss.org/dr16/algorithms/bitmasks/#collapseAPOGEE_ASPCAPFLAG

⁶https://www.sdss.org/dr16/algorithms/bitmasks/#collapseAPOGEE_EXT_RATARG

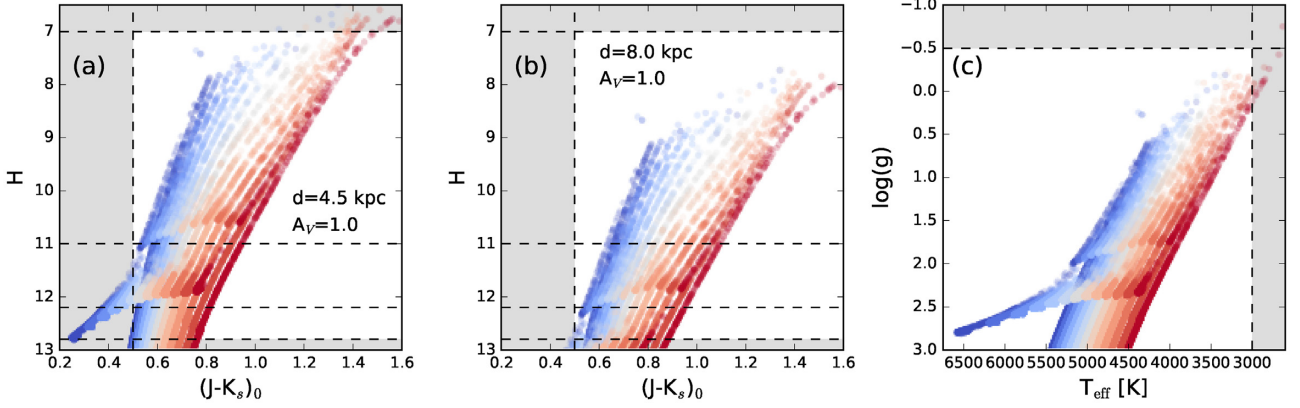


Figure 2. Example of sample selection limits, using 10 Gyr SSPs (based on MIST isochrones) with $[\text{Fe}/\text{H}]$ between -2 (blue) and $+0.4$ (red). Panel (a) simulates the photometric properties of populations with a heliocentric distance of 4.5 kpc ($R_{\text{GC}} \approx 3.5$ kpc) and an extinction of $A_V = 1$ mag. Dashed lines indicate the colour and magnitude limits used by APOGEE, with the shaded regions excluded from the main survey sample. Panel (b) is similar to (a), but for these same populations shifted to the approximate distance of the Galactic Centre. Panel (c) shows the $T_{\text{eff}}-\log g$ distribution of these stars, with the limits of the ASPCAP parameter grid shown as dashed lines.

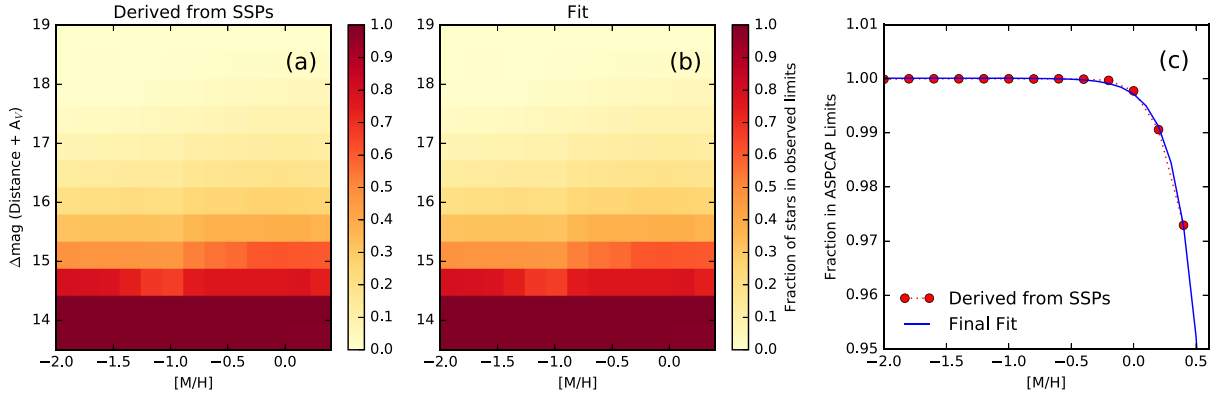


Figure 3. Panels (a) and (b): Example of fit to the ‘observed’ fraction of the RGB+AGB branch, as a function of stellar $[\text{M}/\text{H}]$ and μ for the selection bin defined by $(J - K_s)_0 \geq 0.5$ and $7 < H < 12.2$ (see Section 3.1). Panel (c): Fit to the fraction of simulated RGB+AGB stars within ASPCAP’s analysis range.

Dotter 2016) for $-2.0 \leq [\text{Fe}/\text{H}] \leq +0.4$, with $\Delta[\text{Fe}/\text{H}] = 0.2$ dex, using a Kroupa (2001) IMF. We computed H -band photometry for each SSP at a range of distances (d) and extinctions (A_H), together as $\mu = 5 \log d - 5 + A_H$; we explored $13.5 \leq \mu \leq 19.0$, which encompasses our closest bulge stars (assuming $A_H = 0$) out to distance/extinction combinations that would remove a population from APOGEE’s sample entirely. For each of the five APOGEE selection bins represented in our data set,⁷ we then counted the fraction of the $\log g \leq 2.2$ RGBs that fell within the bin. An example of these fractions is shown in Fig. 3(a), where the shading indicates P_{obs} as a function of $[\text{M}/\text{H}]$ and μ for one APOGEE selection bin. A two-dimensional linear spline interpolation is used to store the pattern for each selection bin (Fig. 3b). Given the bin in which each real APOGEE star was selected, along with its metallicity, distance, and extinction (Section 2.2), we can account for the probability of finding it in our sample, relative to other stars.

For the second selection effect, we again use SSPs to compute, as a function of metallicity, the fraction of the RGB that lies beyond the DR16 ASPCAP synthetic spectrum grid edge at $T_{\text{eff}} = 3000$ K

(Fig. 2c). This fraction turns out to be well represented by a simple function of metallicity (Fig. 3c):

$$P_{[\text{M}/\text{H}]} = -0.04 \times \exp\left(\frac{[\text{M}/\text{H}] - 0.46}{0.18}\right) + 1, \quad (1)$$

where $P_{[\text{M}/\text{H}]} = 1$ means that the entire upper RGB (with $\log g \leq 2.2$) lies within ASPCAP’s range.

The final sample weight (W_{tot}) for each star is taken as the inverse of the product of its observational and analysis probability, i.e. $W_{\text{tot}} = 1/(P_{\text{obs}} \times P_{[\text{M}/\text{H}]})$. $P_{[\text{M}/\text{H}]}$ is always large enough that it is not dominated by the $[\text{M}/\text{H}]$ uncertainties. However, at small P_{obs} , its value (and thus the value of W_{tot}) is dominated by the uncertainty in μ , which is roughly constant at ~ 0.55 , assuming 25 percent distance uncertainties and 0.1 mag extinction uncertainties. Pairs of stars with $\Delta\mu \approx 0.55$ mag have differences in P_{obs} clustered around 0.1, suggesting $P_{\text{obs}} \geq 0.1$ as a reasonable limit to ensure robust weighting. Stars below this value (~ 8 percent of the total bulge sample) are concentrated in the mid-plane, generally with $|Z| < 80$ pc; they cover the full R_{GC} span of our sample, but are much more heavily reddened than other stars at the same distance and latitude. Their metallicity distribution is identical to the other stars (with $P_{\text{obs}} > 0.1$) in the same latitude range, so our analysis and conclusions do not change at all if these stars with low probabilities

⁷ $(J - K_s)_0 \geq 0.5$ and one of: $7.0 \leq H < 11.0$, $7 \leq H < 12.2$, $11 \leq H < 12.2$, $11 \leq H < 12.8$, or $12.2 \leq H < 12.8$ (Zasowski et al. 2013, 2017).

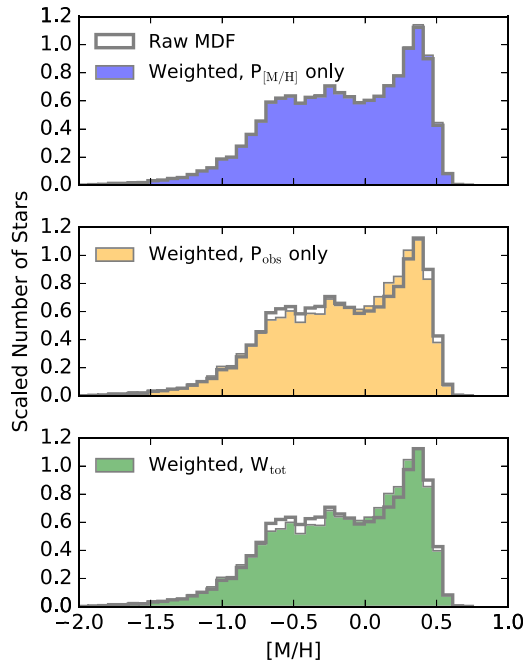


Figure 4. Selection- and metallicity-corrected MDFs for the total bulge sample (Section 3.1), compared to the raw sample counts in grey. *Top panel:* MDF weighted by $P_{[M/H]}$ alone. *Middle panel:* MDF weighted by P_{obs} alone. *Bottom panel:* MDF weighted by both $P_{[M/H]}$ and P_{obs} . In all cases, the weighting has a minimal impact on the global shape of the MDF.

are capped at $P_{\text{obs}} = 0.1$, or indeed if they are excluded altogether. For simplicity, we set $P_{\text{obs}} = 0.1$.

The impact of applying these two types of weights is summarized in the MDFs shown in Fig. 4. In short, the change in the MDF is minimal, suggesting that our bulge sample is not significantly affected by metallicity biases due to the survey selection function or to the effects of the current ASPCAP model grids.

3.2 Exploring the bulge R_{GC} limit

Seen from the Sun, the MW bulge has a B/P appearance, as initially established from the NIR light distribution and 2MASS star counts (e.g. Dwek et al. 1995; López-Corrodoira, Cabrera-Lavers & Gerhard 2005). Following early evidence from gas dynamics that indicates the presence of a non-axisymmetric rotating bar potential (de Vaucouleurs 1964; Rodríguez-Fernández & Combes 2008), studies using luminous stellar tracers (especially RC stars from massive photometric surveys) have characterized the boxy bar as a triaxial structure of length ~ 3.5 kpc, with axial ratios of 1:0.4:0.3 (Rattenbury et al. 2007; Robin et al. 2012; Wegg & Gerhard 2013). Its near side points towards positive Galactic longitudes, with a position angle of $\sim 25^\circ$ with respect to the Sun–Galactic Centre direction, and the asymmetric boxy isophotes are attributed to the effects of the near-end-on projection of a dynamically coherent bar.

Based on these considerations, many recent studies of bulge stellar populations have adopted a Galactocentric distance limit of $R_{\text{GC}} = 3.5\text{--}4$ kpc to define their bulge samples (Ness et al. 2013; Rojas-Arriagada et al. 2017; Schultheis et al. 2017). Our aim in this section is to use the kinematical properties of our clean inner Galaxy sample (Section 2.3) to assess the appropriateness of these limits and choose the best to use for this paper.

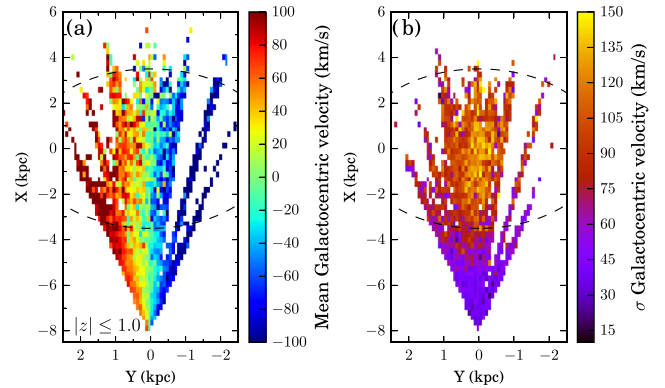


Figure 5. Galactocentric velocity (V_{GC}) maps of the clean inner MW sample in the XY plane. They are computed over a 60×60 bin grid for stars with $|Z| \leq 1.0$ kpc. *Left-hand panel:* mean Galactocentric velocity. Only pixels containing at least four samples are displayed. *Right-hand panel:* dispersion of the Galactocentric velocity. Only pixels containing at least six samples are displayed. The dashed black ellipses indicate in each panel the region defined by $R_{\text{GC}} = 3.5$ kpc.

In Figs 5 and 6, we show mean kinematical properties of our clean inner MW sample in the Galactocentric Cartesian (X, Y) plane. The Galactic Centre is at $(X, Y) = (0, 0)$ kpc, the Sun is at $(X, Y) = (-8.2, 0)$ kpc, and the close end of the bar lies where $X < 0$ and $Y > 0$. Each property is shown for stars limited to $|Z| \leq 1.0$ kpc. The dashed-line ellipses indicate $R_{\text{GC}} = 3.5$ kpc.

Fig. 5 shows the Galactocentric RV V_{GC} of the clean inner MW sample, with mean $\overline{V_{\text{GC}}}$ in the left-hand panel and dispersion $\sigma_{V_{\text{GC}}}$ in the right-hand panel. In the $\overline{V_{\text{GC}}}$ panels, we see an overall uniform rotation pattern for stars at nearly all R_{GC} , especially $R_{\text{GC}} > 3.5$ kpc, where the pattern is particularly smooth. In contrast, the velocity dispersion ($\sigma_{V_{\text{GC}}}$) shows marked spatial patterns, with low dispersion (indicating coherent rotation) at $R_{\text{GC}} \gtrsim 4$ kpc and higher dispersion (by a factor of 2 or more) at smaller radii, indicating more isotropic kinematics.

Fig. 6 shows additional kinematic and orbital (eccentricity) properties for stars from the clean inner MW sample with good proper motions and orbit measurements (Section 2.2). These are defined as stars with uncertainties in both $\mu_\alpha \cos(\delta)$ and μ_δ smaller than 0.5 mas yr^{-1} and uncertainties in tangential velocity (v_T , projected on to the plane), vertical velocity (v_Z), and eccentricity smaller than 40 km s^{-1} , 20 km s^{-1} , and 0.14 , respectively. The error distributions of these quantities are all sharply peaked, so these limits only remove stars in the long high-error tails (beyond the 95th percentile of the distribution). In addition, as recommended by the *Gaia* consortium, we select stars with $\text{RUWE} < 1.4$, to ensure that their astrometric solutions are reliable.

The top row of Fig. 6 contains the (X, Y) distribution of $\overline{v_T}$ (left) and $\sigma(v_T)$ (right), the middle row contains similar maps for $\overline{v_Z}$ and $\sigma(v_Z)$, and the bottom row contains a map for mean eccentricity \bar{e} . Together, these distributions tell a complementary story of a kinematical transition between the disc- and bulge-dominated regions. The disc is dominated by stars that are in roughly circular, coplanar rotation, with $v_T \approx 150\text{--}250 \text{ km s}^{-1}$, low dispersion in both v_T and v_Z , and roughly uniform $e \lesssim 0.25$. Inside $R_{\text{GC}} \sim 3.5$ kpc, however, rather sharply the kinematics become significantly less coherent, with $\sigma(v_T)$ and $\sigma(v_Z)$ greater than 100 km s^{-1} and mean eccentricity exceeding $\bar{e} = 0.6$.

These trends indicate a relative dearth of circular, coplanar orbits and higher orbital isotropy in the inner Galaxy, and a transi-

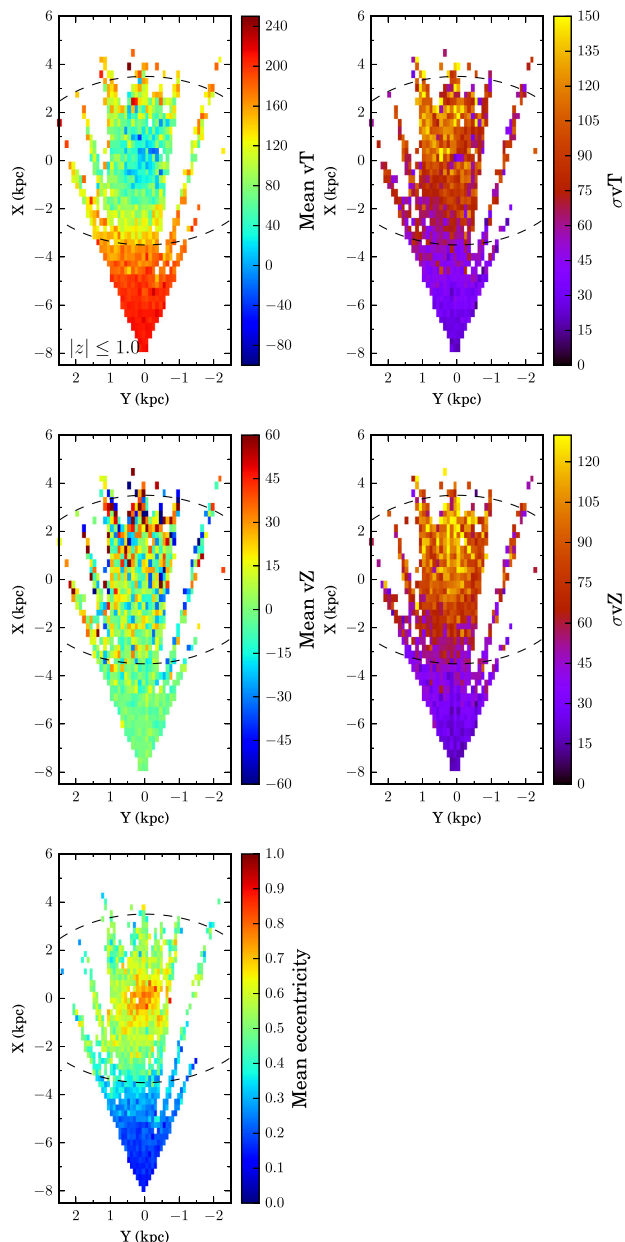


Figure 6. The binned statistics of tangential velocity (v_T , upper row), vertical velocity (v_z , middle row), and eccentricity (e , lower row) are computed in the (X, Y) Galactic Cartesian plane. In each case, a 50×50 bin grid has been adopted, except for eccentricity in which case we adopt 60×60 bins. Only stars with $|Z| \leq 1.0$ kpc are considered. In each panel, the red dashed-line ellipse depicts the $R_{GC} = 3.5$ kpc limit.

tion between domination by the rotation-supported disc and the bar+pressure-supported bulge in a narrow region at $R_{GC} \approx 3.5$ kpc. Based on this behaviour, we adopt this as our spatial limit to select a sample of likeliest bulge stars (Section 3.3).

3.3 Selection of a bulge sample

We use the results of Sections 3.1 and 3.2 to define our sample of bulge stars. Starting with the clean inner MW sample defined in Section 2.3, we remove stars with $\log g > 2.2$ dex, to ensure our sample is adequately free of the selection and analysis effects of

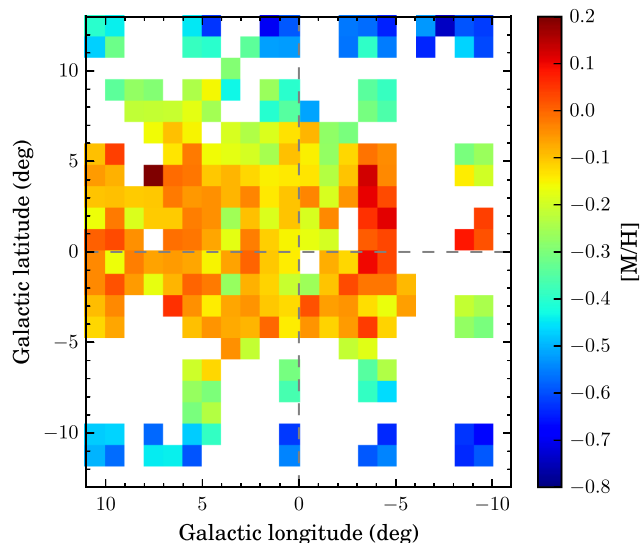


Figure 7. Mean metallicity map of our bulge sample. The mean metallicity is computed over a 22×22 grid within $|\ell| \leq 11^\circ$ and $|b| \leq 13^\circ$. Only pixels with more than seven stars are shown.

Section 3.1. Then, we select stars with $R_{GC} \leq 3.5$ kpc as those in the kinematically distinguished bulge/bar region explored in Section 3.2. Finally, we focus for the rest of the analysis on the spatial variation of the MDF within the region bounded by $|\ell| \leq 11^\circ$ and $|b| \leq 13^\circ$, so we keep only stars inside those limits. The final bulge sample comprises 13 031 stars.

Fig. 7 shows the mean metallicity map (in ℓ, b) of the bulge sampled by the APOGEE bulge stars. The mean metallicity was computed over a 22×22 grid in the $|\ell| \leq 11^\circ, |b| \leq 13^\circ$ region, displaying only pixels containing more than seven stars. The dominant gradient is in latitude, with average metallicity increasing towards the mid-plane. In addition, although the sampling of the APOGEE data is less dense beyond 5° from the Galactic plane, the general pattern appears symmetric around $b = 0^\circ$. Beyond $|b| > 10^\circ$, however, a weak longitudinal gradient is visible in the map, with the mean population becoming more metal-rich towards positive longitudes. As we shall see, this region is dominated by stellar populations other than bulge stars, and it is at the border of the region we are sampling. Consequently, we do not quantify the significance of this effect. Qualitatively, this mean metallicity map compares quite well with the photometric map of Gonzalez et al. (2013) derived from VISTA Variables in the Vía Láctea survey (Minniti et al. 2010) data, and shows the vertical metallicity gradient (see also Fig. 12).

4 THE MDF

Fig. 8 shows the (l, b) distribution of our bulge sample, emphasizing (along with Fig. 1) that APOGEE preferentially samples the region within 5° of the Galactic plane, even with stars at $R_{GC} \leq 3.5$ kpc. This region has been largely avoided by large-scale optical spectroscopic surveys because of the severe limitations imposed by the high dust extinction. Thanks to its NIR H -band sensitivity, APOGEE is less affected by this extinction and offers the opportunity to systematically explore the nature of the MDF with good statistics across dozens of square degrees in the innermost parts of the MW.

Numerous bulge studies have shown that the proportion of metal-rich to metal-poor stars decreases as a function of the distance ($|b|$ or $|Z|$) from the mid-plane (Zoccali et al. 2008; Ness et al. 2013;

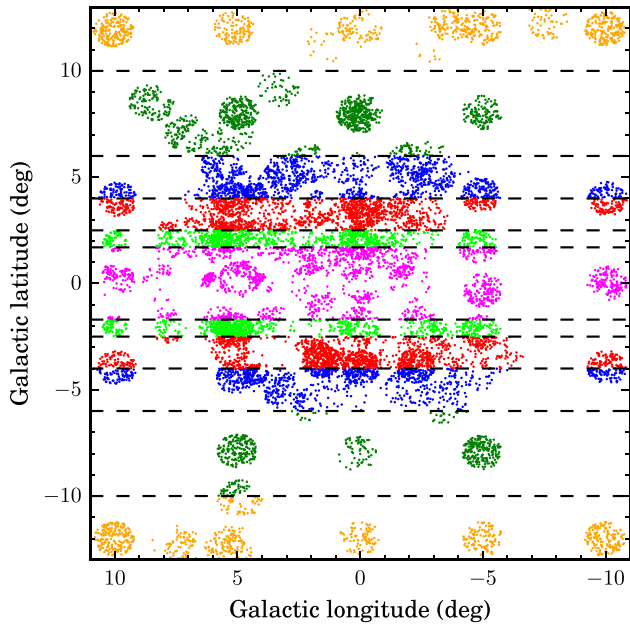


Figure 8. Distribution of our bulge sample of 13 031 stars in (l, b) . The horizontal black dashed lines define our latitude strips, with a binning pattern that is symmetric about the mid-plane (Section 4). Points are colour coded by latitude strip.

Rojas-Arriagada et al. 2017). As we also focus here on the latitude dependence of the MDF, we split our bulge sample into latitude strips that extend across the full longitude range and are symmetric around $b = 0^\circ$ (Fig. 8); we combine stars into strips of $|b|$ after confirming that the MDFs of the matched strips at $\pm b$ are statistically identical (see also the symmetry of Fig. 7). The varying widths of the strips are chosen to maximize the resolution in $|b|$ while still ensuring more than 1600 stars per strip, each of which is shown in a different colour in Fig. 8. The three outermost strips include stars spanning areas previously studied by optical spectroscopic surveys (e.g. ARGOS, GIBS, and GES at $b < -4^\circ$).

The goal of this paper is to study in greater detail the *shape* – especially the ‘peakiness’ – of the MDF and how that shape changes with latitude in the relatively poorly explored inner few degrees of the bulge, all enabled by the large size and wide angular distribution of the APOGEE sample. In Section 4.1, we describe the qualitative properties of the MDFs in our six bins of $|b|$. In Sections 4.2 and 4.3, we explore statistical decompositions of these MDFs, the stability of these decompositions with $|b|$, and the kinematical properties of the resulting components. In Section 4.4, we use an alternate decomposition method to highlight how different methods may affect our interpretation of bulge ‘components’.

4.1 MDF properties

Fig. 9 shows the MDFs of the spatial bins defined above, along with Gaussian mixture components that are discussed in Section 4.2.

The MDF of the outermost spatial (folded) bin at $|b| \geq 10^\circ$ is dominated by a single peak at $[M/H] \sim -0.5$ dex (top panel of Fig. 9). Its overall shape is highly symmetric, but not Gaussian, with pronounced tails extending towards the metal-rich and metal-poor ends. The location of the dominant peak is too metal rich to be attributed to halo stars (although halo stars might be present in the metal-poor tail of the distribution), but it is consistent with the

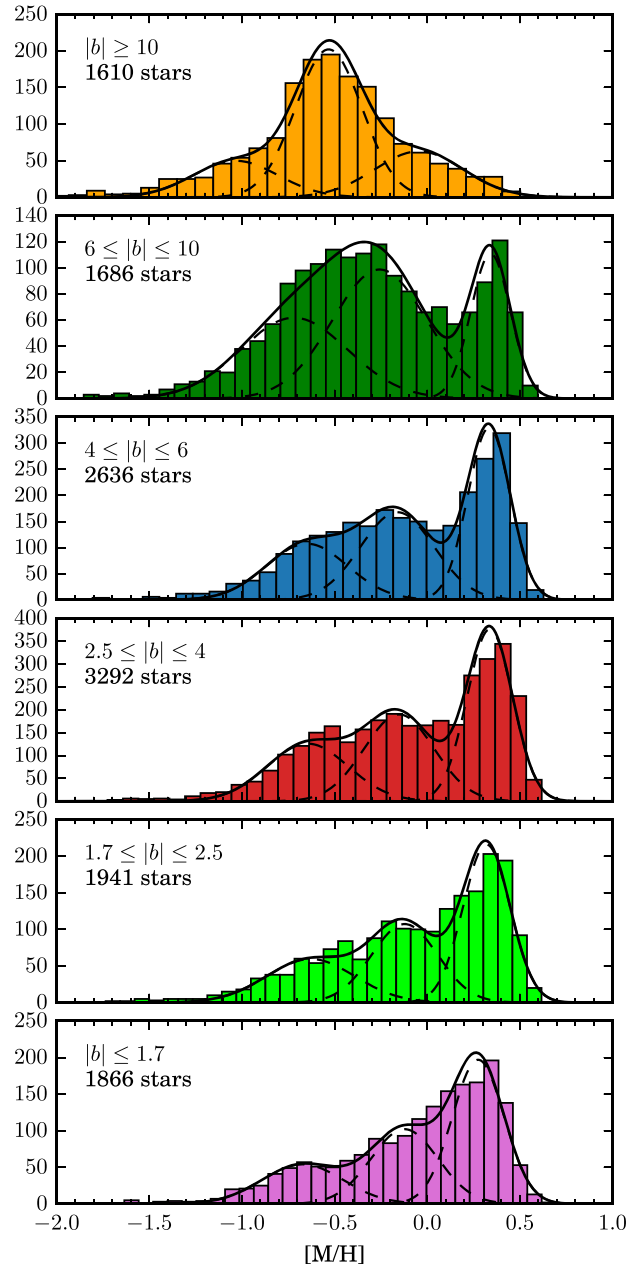


Figure 9. MDFs of the latitude strips shown in Fig. 8, with the same colour scheme. The black dashed and solid lines indicate the individual and summed components, respectively, of the optimal GMM modelling of the data’s density distribution (Section 4.2).

presence of thick disc stars dominating the line-of-sight mix of stellar populations at high Galactic latitude (for reference, stars located at $|b| = 10^\circ$ are at $|Z| \sim 1.45$ kpc from the mid-plane at a $d = 8.2$ kpc distance to the bulge).

At $6^\circ \leq |b| \leq 10^\circ$ the MDF appears bimodal, dominated by a broad metal-poor distribution with a peak at $[M/H] \sim -0.4$ dex. The second prominent peak is narrower and more metal rich, centred at $+0.4$ dex. A similar, but less peaky, bimodal distribution is seen in the stars at $4^\circ \leq |b| \leq 6^\circ$. In this case, the metal-rich peak is dominant and well defined, while metal-poor stars are present in a flatter distribution. The spatial region covered by these two strips has been previously sampled by optical spectroscopic surveys such as

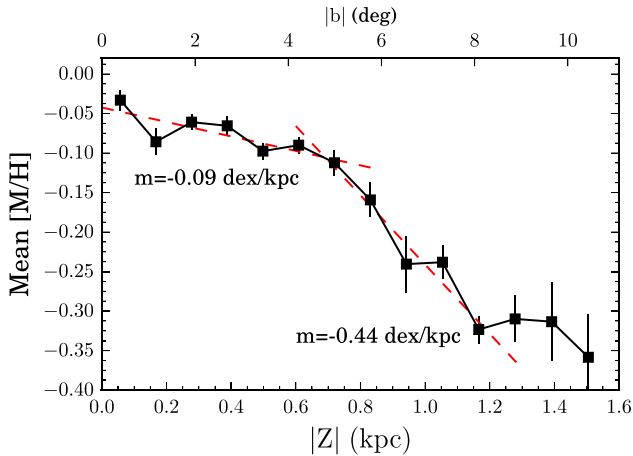


Figure 10. Vertical metallicity gradient computed from the mean metallicity of stars in a number of narrow latitude intervals. Angular distances (in the additional top axis) have been converted to Z by assuming a projection on a plane at 8 kpc (to be consistent with other studies and allow comparison). The gradient has been estimated in two independent regions, as indicated by the red dashed lines.

GIBS and GES (at negative latitudes). The overall picture proposed from these surveys is that the bulge MDF is intrinsically bimodal, with the relative proportion of metal-rich stars increasing as closer to the mid-plane. The visual inspection of the MDFs from our data over the same spatial region seems to qualitatively confirm this picture.

Inside $|b| < 4^\circ$, the MDF becomes increasingly dominated by the metal-rich peak (lower three panels of Fig. 9). As $|b|$ gets smaller, the metal-poor distribution becomes increasingly flat, resembling more of a heavy tail to the metal-rich peak; upon close examination, this tail appears to comprise two wide peaks, separating around $[M/H] \sim -0.4$ dex. This separation is weaker in the innermost strip ($|b| \leq 1.7^\circ$).

In Fig. 10, we divide the bulge sample in several narrow latitude strips in order to compute the vertical mean metallicity gradient. The angular distances were converted to spatial vertical distances by assuming all stars in each bin projected in a plane at 8 kpc. This is done in order to be consistent with other studies and allow comparison. The vertical gradient has been measured in previous studies as a summary quantity of the vertical variation of the composition of lines of sight towards the bulge. The resulting gradient appears to have two slopes: a flatter one of -0.09 dex kpc^{-1} within 0.7 kpc ($|b| \sim 5^\circ$) from the mid-plane, and a steeper one of -0.44 dex kpc^{-1} from there out to 1.2 kpc ($|b| \sim 8^\circ$). Beyond that point, the slope becomes flat but noisy. The slope we measure beyond $|b| \sim 5^\circ$ is lower than the value reported over a similar region ($-12^\circ \leq b \leq -4^\circ$) by Zoccali et al. (2008) and Rojas-Arriagada et al. (2017) (-0.24 dex kpc^{-1} , both). The shallower value reported in these works may be driven by the inclusion of outer fields, which as we see in our data are flatter but noisy. On the other hand, the inner flattening of the metallicity gradient is in agreement with previous suggestions from the analysis of smaller samples of high-resolution NIR spectra (Rich et al. 2007; Rich, Origlia & Valenti 2012; Schultheis et al. 2019).

This spatial variation of the shape of the MDF reflects a complex mix of stellar populations present in the bulge. The identification and characterization of this complexity has been a key outcome from the optical surveys exploring the southern bulge region at $b \gtrsim -4^\circ$. The presence of at least two separate ‘populations’ (one metal rich and one metal poor) with different spatial distributions,

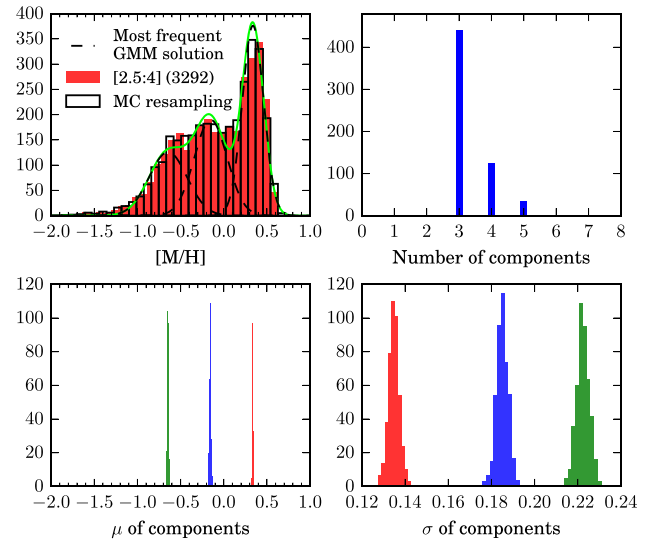


Figure 11. Example of the Monte Carlo analysis implemented to assess the stability of the GMM solution of the MDF. *Top left panel:* the red histogram stands for the observed MDF in a given latitude strip indicated in the legend, as well as the number of stars in that strip. The black step histogram depicts one random sampling of the observed MDF given metallicity measurement uncertainties. The black dashed lines show the individual components, while the solid green line shows total profile of the optimal trimodal Gaussian mixture. *Top right panel:* frequency distribution of the optimal number of Gaussian components found from the GMM runs on the 600 Monte Carlo resamplings of the observed MDF. *Bottom panels:* frequency distributions of the centroids (left) and widths (right) of each Gaussian component in the preferred trimodal parametrization.

different kinematics, and possibly different origins is argued to drive the variable behaviour of stellar distributions, including the MDF.

4.2 MDF decomposition with Gaussian mixture modelling (GMM)

Identifying and separating by chemistry these ‘populations’ described above has largely been done by assuming that the intrinsic shapes of their underlying distributions can be described by simple functions (most commonly, Gaussians) that do not have a large overlap. Expanding on this tradition, we use GMM to decompose the $|b|$ -dependent MDFs and study not only the structure of the distribution, but also how robust this structure is to uncertainties in the data, how robust it is to the stochasticity of GMM decomposition, and how the peakiness itself depends on $|b|$.

To this end, we adopted a Monte Carlo resampling approach, which is illustrated in Fig. 11. Given an observed MDF (red histogram, top left panel), we draw 600 resamplings of the individual stellar metallicities by assuming a Gaussian variability of 0.05 dex (approximately APOGEE’s nominal metallicity measurement uncertainty; e.g. Holtzman et al. 2018; Jönsson et al. 2018). An example of the MDF resulting from one such resampling is shown as a black empty histogram in the top left panel of Fig. 11.

A GMM parametrization was computed for each of the 600 resamplings, considering mixtures with $N = 1-6$ components, and using the Bayesian Information Criterion (BIC) to identify the optimal model. The frequency of the optimal number of components over the whole set of resamplings was examined to see if a preferred model complexity emerged (top right panel). A trimodal solution is found to be the optimal one in the majority of the cases (~ 75 per cent

Table 1. Relative frequency of the preferred number of GMM components used to model the MDF over the 600 Monte Carlo resamplings in each latitude strip.

Galactic latitude	$N = 1$	$N = 2$	$N = 3$	$N = 4$	$N = 5$	$N = 6$
$10^\circ \leq b $	0.00	0.00	0.99	0.01	0.00	0.00
$6^\circ \leq b < 10^\circ$	0.00	0.37	0.56	0.05	0.02	0.00
$4^\circ \leq b < 6^\circ$	0.00	0.11	0.47	0.34	0.09	0.00
$2.5^\circ \leq b < 4^\circ$	0.00	0.00	0.75	0.19	0.06	0.00
$1.7^\circ \leq b < 2.5^\circ$	0.00	0.00	0.56	0.38	0.07	0.00
$ b < 1.7^\circ$	0.00	0.00	0.93	0.07	0.00	0.00

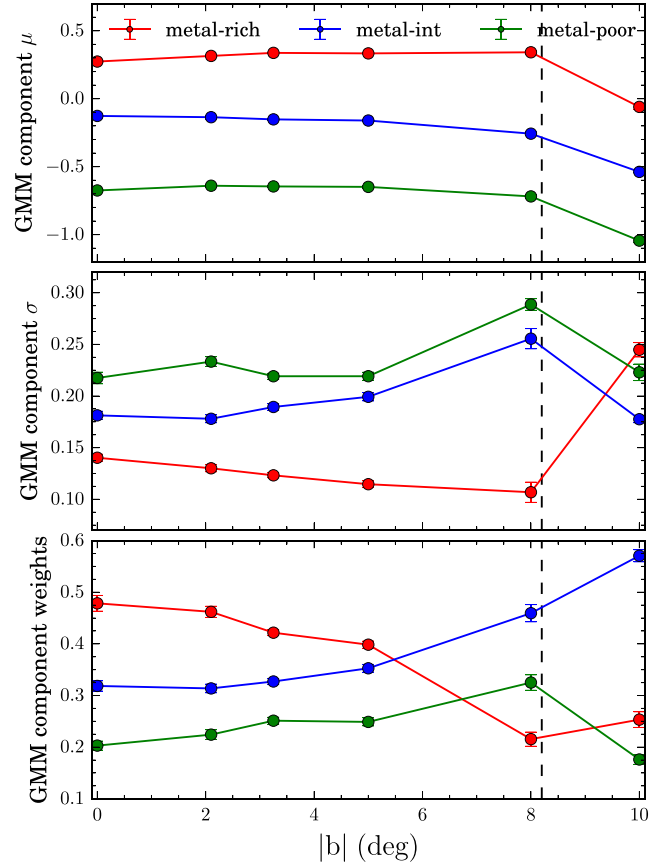
in this example). Thus, we choose the trimodal mixture as the stable optimal parametrization of this MDF density structure.

The lower panels of Fig. 11 show the distributions of the centroids and widths of the individual Gaussian components in the trimodal solutions. The narrowness of these distributions indicates that the component parameters are robust over the set of Monte Carlo resamplings. The optimal trimodal Gaussian mixture is shown on top of the observed MDF in the top left panel. Dashed black lines denote the individual components, while the solid green line shows the total density distribution. This distribution appears to be a fairly good description of the density distribution of the observed data at this latitude.

We applied the procedure described above to the observed MDFs in each of our latitude strips (Fig. 9). The components of the optimal model are shown on top of each MDF in dashed lines, with their sum as the solid line. Despite the changing shape of the MDF with latitude, *in all cases the preferred GMM solution is trimodal*. This can be seen in Table 1, which contains the relative frequency of solutions with $N = 1$ –6 components over the whole set of Monte Carlo resamplings of each latitude strip. Table 2 contains the parameters of the optimal Gaussian mixture for each of the latitude strips (Fig. 9).

The metal-rich peak, which is visually apparent in all the MDFs, corresponds closely to the consistently narrowest of the GMM components (except in the outermost latitude bin at $|b| \geq 10$, which as mentioned before is likely dominated by contamination). Interestingly, the broad metal-poor distribution identified in the qualitative assessment of the MDF shape (Section 4.1) appears to be best described by two Gaussian components. This is the case even in the strips where visual inspection may suggest a single metal-poor peak. In the rest of this analysis, we will refer to these components as metal rich, metal intermediate, and metal poor. As we shall argue, this nomenclature does not necessarily imply physically different structures/populations, but rather distinct *data* components needed to parametrize the density structure of the observed MDF.

The narrowness of the probability distribution functions for the centroid μ of each component (e.g. bottom left panel of Fig. 11), and the constancy of μ with $|b|$ (e.g. top panel of Fig. 12), suggests

**Figure 12.** Summary plots of the GMM analysis on the latitude-separated MDFs. The curves display the variation as a function of the distance from the mid-plane of the individual Gaussian parameters of the preferred trimodal mixture: centroids (μ , top panel), widths (σ , middle panel), and relative weights (w , bottom panel). In all panels, the vertical black dashed line separates the region seemingly dominated by three similar bulge components (i.e. $|b| < 8^\circ$) from that dominated by different components, possibly describing the halo+thick disc (Section 4.2).

that the metallicities of the components do not themselves depend on distance from the Galactic plane. In this sense, the shape of the MDF is determined by the combination of components whose relative weights change systematically, with metal-rich stars becoming progressively important closer to the mid-plane. This observation is consistent with the qualitative picture drawn from optical surveys.

We note that although the MDF of the sample at $|b| \geq 10^\circ$ is also best represented by a trimodal mixture, the individual components look different from those at other latitudes. The most prominent component peaks at $[M/H] \sim -0.5$ dex, which is located between

Table 2. The mean $[M/H]$ (μ), width (σ), and relative weights (w) of the three individual components found to be the preferred and stable GMM solution in the MDF of our latitude strips.

Strip	Metal-poor			Metal-intermediate			Metal-rich		
	μ	σ	w	μ	σ	w	μ	σ	w
$10 \leq b $	-1.04	0.23	0.18	-0.53	0.18	0.58	-0.04	0.24	0.24
$6 \leq b \leq 10$	-0.72	0.29	0.33	-0.26	0.26	0.46	0.34	0.11	0.22
$4 \leq b \leq 6$	-0.65	0.22	0.25	-0.16	0.20	0.35	0.33	0.11	0.40
$2.5 \leq b \leq 4$	-0.64	0.22	0.25	-0.15	0.19	0.33	0.34	0.12	0.42
$1.7 \leq b \leq 2.5$	-0.63	0.24	0.23	-0.12	0.18	0.32	0.32	0.13	0.45
$ b \leq 1.7$	-0.68	0.22	0.20	-0.13	0.18	0.32	0.27	0.14	0.48

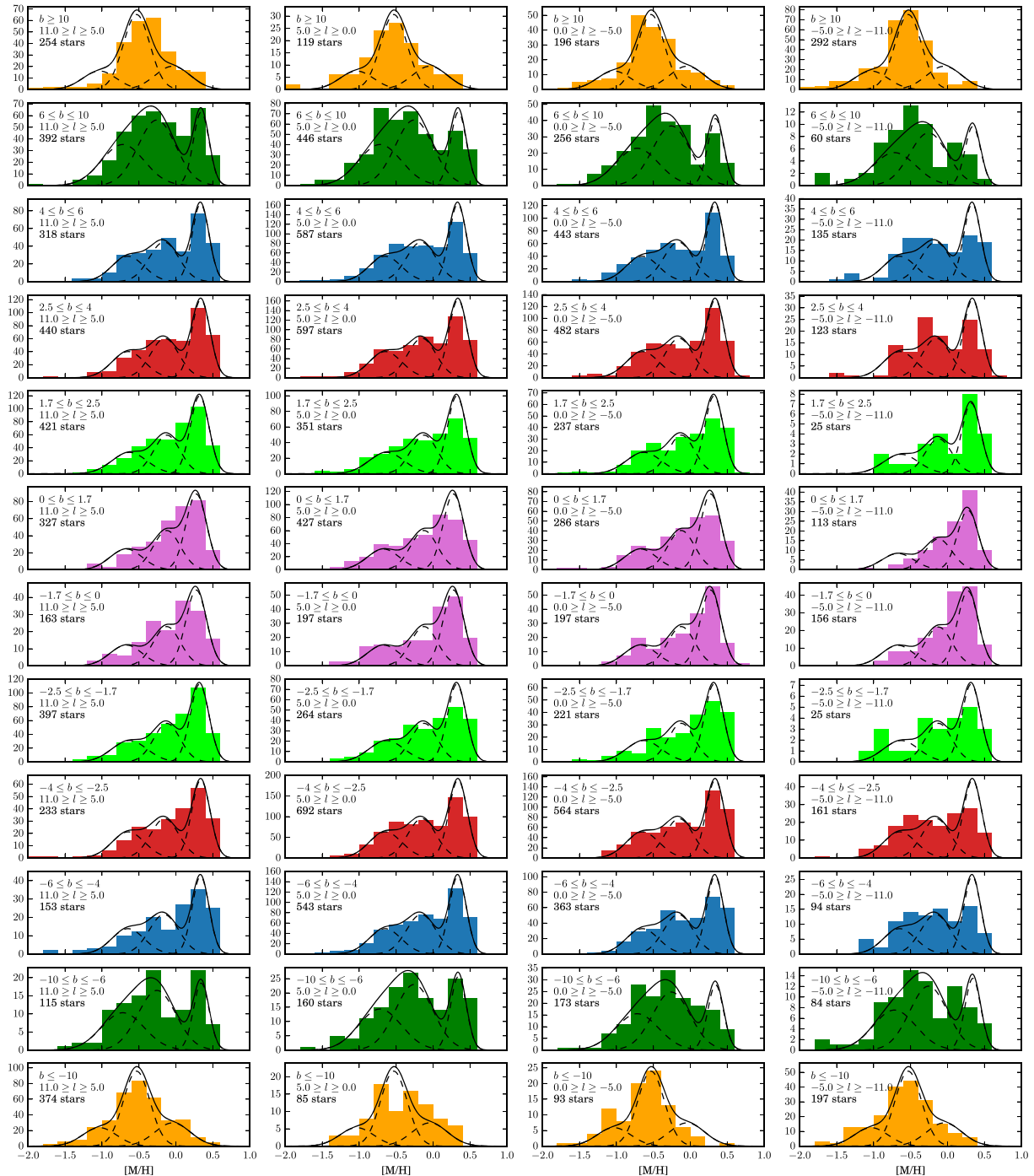


Figure 13. Variations of the bulge MDF in binned regions in (l, b) . Stars in each latitude strip (see Fig. 8) are split into four longitudinal ranges: $5^\circ \leq l \leq 11^\circ$, $0^\circ \leq l \leq 5^\circ$, $-5^\circ \leq l \leq 0^\circ$, and $-11^\circ \leq l \leq -5^\circ$. The MDFs of the resulting samples are distributed in the figure to preserve their relative position in the (l, b) plane, as shown in Fig. 8. The Galactic longitude/latitude range and number of stars are quoted in each panel. The best-fitting GMM to the full sample in each latitude strip is overlotted with dashed and solid black lines (individual components and full mixture, respectively).

the metal-intermediate and metal-poor peaks of the other latitude samples. A similar offset is observed for the two other components of this mixture. This reinforces our previous hypothesis that our highest latitudes are dominated by stars from the thick disc and halo.

Fig. 12 contains a summary of the individual component means (μ), widths (σ), and relative weights as a function of angular distance from the mid-plane. The top panel highlights the constancy of the metallicity centroids of the individual components over the whole

$0 \leq |b| \lesssim 8^\circ$ region, only deviating in the outermost bin at $|b| \geq 10$. A similar trend is seen for the widths of the components (second panel from top), although with larger deviations than the centroids.

The trend of the relative weight of components with $|b|$ (third panel) suggests that the global shape of the MDF is largely determined by the variation of the relative weights of the metal-rich and metal-intermediate components. In this sense, the metal-poor component accounts for a smaller (and relatively constant) proportion

of the data density, with only a hint of a modest enhancement at higher latitudes. The inversion in the relative importance of the metal-rich and metal-intermediate components happens in the range $|b| \sim 5^\circ - 7^\circ$, with the ratio becoming stable at the innermost 2–3 deg.

In Fig. 13, we compare the $|b|$ -dependent best-fitting GMM solutions to smaller subsamples separated around the mid-plane (i.e. using b instead of $|b|$) and further divided by Galactic longitude: $-11^\circ \leq l \leq -5^\circ$, $-5^\circ \leq l \leq 0^\circ$, $0^\circ \leq l \leq 5^\circ$, and $5^\circ \leq l \leq 11^\circ$. The arrangement of the MDFs of these angular pixels in Fig. 13 is consistent with the layout in Fig. 8. The distributions are noisier, as expected, but the number of stars per bin is still typically >100 . As a visual reference, in each panel, we show the best-fitting GMM to the full sample in each latitude strip with dashed (individual components) and solid black (summed mixture) lines. Stars from low to supersolar metallicity are observed in all individual MDFs across the whole (l, b) region. At fixed l , the MDFs of fields symmetric about the mid-plane (i.e. at $\pm b$) appear consistent with each other.

On the other hand, if one compares the MDF of individual fields along a latitude strip, one sees that the strength of the metal-rich peak finds better agreement at positive longitudes to the respective whole-latitude-bin GMM model. This asymmetry is clearer in the strips beyond $|b| = 2.5^\circ$. The trend may be explained by the asymmetry introduced by the bar, whose major axis lies at an angle to the Sun–Galactic Centre line and whose near-end is located at positive longitudes (the sightlines at negative longitudes must extend to larger distances to reach the metal-rich-dominated bar). This is consistent with the association of metal-rich stars with the bar (also see Wegg et al. 2019 and Hasselquist et al. in preparation). We emphasize that these asymmetries manifest themselves only as changes in the relative component weights of the observed (noisy) MDF with respect to the latitude best-fitting GMM, not the centres or widths of the components.

In summary, our GMM analysis shows that, when based on statistically significant samples, the density structure of the bulge MDF is optimally and robustly parametrized by a trimodal Gaussian mixture. The metallicity of the metal-rich, metal-intermediate, and metal-poor components remains nearly constant with latitude at approximately -0.66 , -0.17 , and $+0.32$ dex, respectively. The strongest variation of the MDF shape is in the vertical direction, and appears mostly driven by the variation of the relative weight of the metal-rich and metal-intermediate components. The metal-poor component seems to account for a relatively constant and low fraction of the total MDF density, regardless of the distance to the mid-plane.

4.3 Kinematic properties of the MDF GMM components

We examine in this section the kinematic patterns of stars most closely associated with each of the three GMM components. Fig. 14 contains maps of the mean velocity and velocity dispersion in the (l, b) plane, computed from the bulge sample (folded about the mid-plane) in a grid of $(\Delta l, \Delta b) \sim (1.25^\circ, 1.5^\circ)$. As a complementary view, we display in Fig. 15 the mean and dispersion Galactocentric velocity curves for the three metallicity components separated into the latitude strips adopted elsewhere in this work. In this general qualitative assessment, we attempt to construct subsamples of stars most likely to belong to each of the MDF components. To this end, we separate the bulge sample into three metallicity ranges, centred on the three components inferred in Section 4.2 but avoiding the regions of largest overlap (and so, where the association of a star with a given component is more uncertain). The sample is thus separated into the following ranges: $[M/H] \geq +0.1$ dex (metal rich), $-0.4 \leq [M/H] \leq 0.0$ dex (metal intermediate), and $-1.2 \leq [M/H] \leq -0.5$ dex (metal poor).

From the upper set of panels of Figs 14 and 15, one can see that all three components show a cylindrical rotation pattern about the minor axis. The latitudinal variation of the rotation pattern (i.e. the departures from perfect cylindrical rotation) seems to be somehow larger for the metal-rich component compared to the other two. This component has in addition the most pronounced rotation in the mid-plane (compare, for example, curves for $|b| \leq 1.7^\circ$). On the other hand, a more uniform rotation pattern can be seen in stars belonging to the metal-poor component, which also show an overall slower rotation. Although the cylindrical rotation pattern in the bulge is known from previous optical spectroscopic surveys (Zoccali et al. 2014), we characterize it here from stars distributed in a previously ill-explored region close to the mid-plane. Our results here update those previously found from the APOGEE DR12 data (Ness et al. 2016; Zasowski et al. 2016), although with higher spatial resolution due to the larger sample available now.

The patterns of velocity dispersion (lower row of Figs 14 and 15) show some differences between the metallicity components. In the metal-rich stars, the velocity dispersion has overall a large vertical variation, increasing sharply with decreasing latitude, with a pronounced enhancement in the central region within $|\ell| < 4^\circ$ and $|b| < 4^\circ$.

This is also consistent with the results of Ness et al. (2016) and Zasowski et al. (2016). In contrast, Figs 14 and 15 show that the metal-intermediate component is overall kinematically hotter (less variation with latitude) and displays a less pronounced increment of velocity dispersion towards the centre. An even larger isotropy is seen for stars in the metal-poor group, which appears kinematically hot over nearly the entire sampled spatial region. This is clearly seen from both the map and the relative flat velocity dispersion curves. As we discuss in Section 5, these varying projected kinematical patterns can be used to speculate on the orbital properties of the stars associated with each of the metallicity components.

4.4 MDF decomposition with NMF

The primary aim of this paper is to explore the complexity of the bulge's MDF using GMM (Sections 4.2 and 4.3), but in this section we describe an alternative decomposition using NMF (Lee & Seung 1999, with some examples of astrophysical applications in Igual & Llinares 2008 and Hurley et al. 2014). NMF is a dimensionality reduction technique similar to principal component analysis in spirit; one key difference is that the eigenvectors are constrained to be non-negative. This constraint makes NMF an appealing option for decomposing a spatially variant MDF, since in principle, the MDF at any location is a superposition of every MDF component scaled by a non-negative value (including 0). An additional benefit of NMF is that the components themselves are not limited to Gaussians or any pre-defined functional form.

Using the same absolute latitude bins described above (Fig. 8), and motivated by the robustness of the three-component GMM solution, we perform a three-component NMF decomposition on the same MDFs over the range $-2 \leq [M/H] \leq +0.7$ dex ($\Delta[M/H] = 0.1$ dex). The NMF components and the reconstructed MDFs are shown in the left-hand and middle panels of Fig. 16, respectively. When rerunning the decomposition while removing one of the input MDFs, the missing MDF is able to be reconstructed using the output components to the same apparent quality as when all MDFs are included. We performed this same procedure on the subsampled MDFs described in Section 5.1 and found no difference in the shape of the components, simply a reduced amplitude.

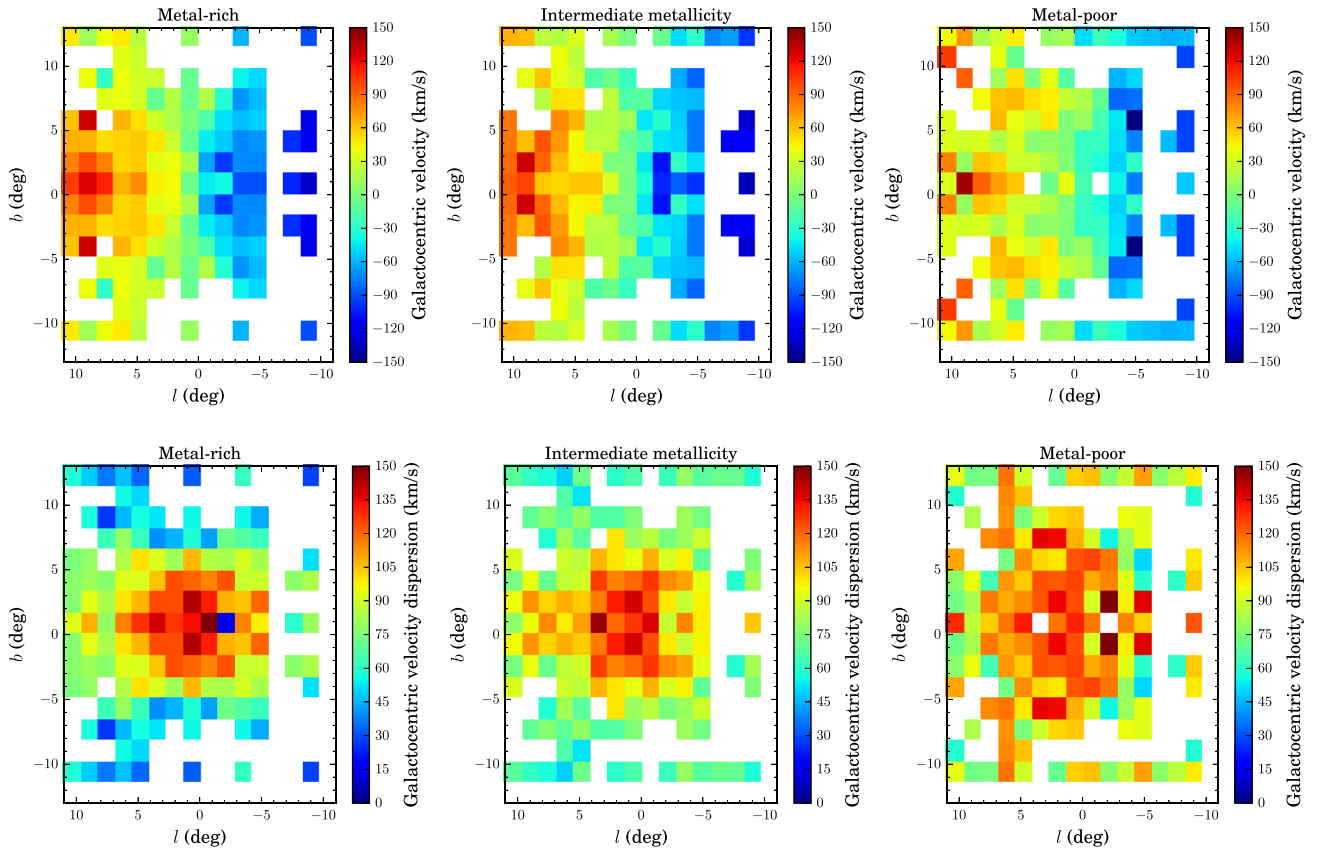


Figure 14. (l, b) kinematic maps for the three metallicity components found in the bulge MDF. In each case, the sample has been two-folded with respect to the Galactic plane, and the map has been constructed by binning into a 16×16 grid. Only pixels containing more than five stars are displayed. Top panels: mean Galactocentric velocity maps. Bottom panels: Galactocentric velocity dispersion maps.

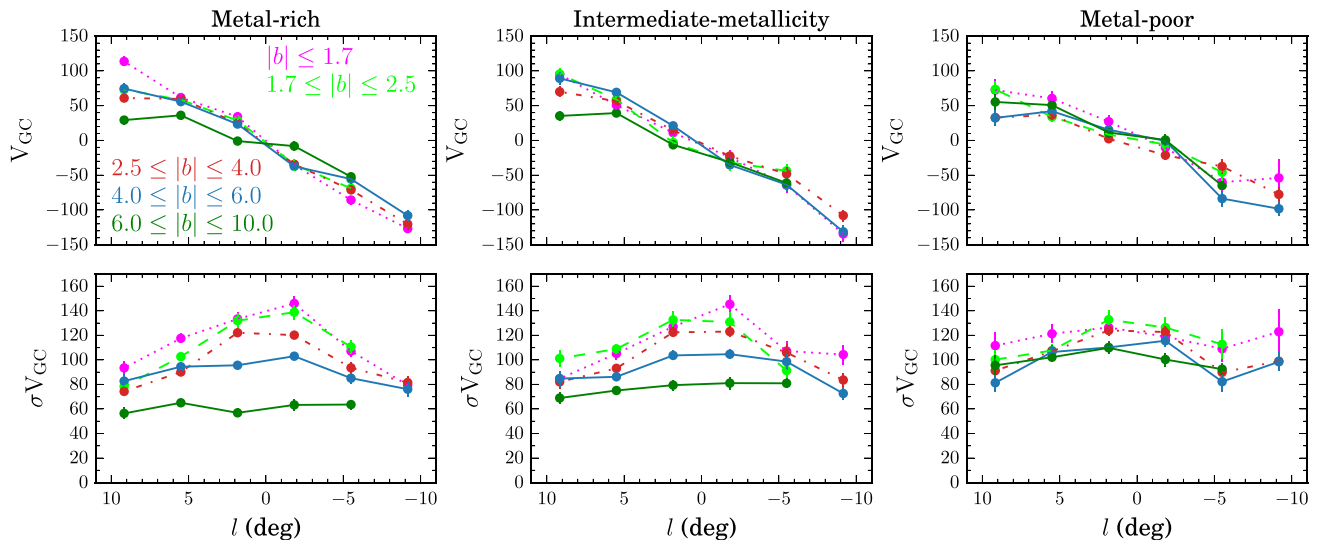


Figure 15. Mean Galactocentric velocity and velocity dispersion curves for stars in the three GMM-inferred metallicity components. The colour of the points and lines indicates the latitude range, using the same colour scheme adopted elsewhere in this paper (e.g. Figs 8 and 9).

In comparing the NMF components to the range of GMM components (vertical lines in Fig. 16a), we find broad consistency along with some marked differences, some of which are simply due to the different methodologies and some of which may contain meaningful information. Regardless of the weights given to the three components, a trimodal distribution results in agreement with the

GMM findings.⁸ The metal-poorest NMF component (NMF 2, green

⁸The exception is if a zero weight is given to the broadest, most metal-rich eigenvector and not to the other two, which does not match any of the patterns in our data set.

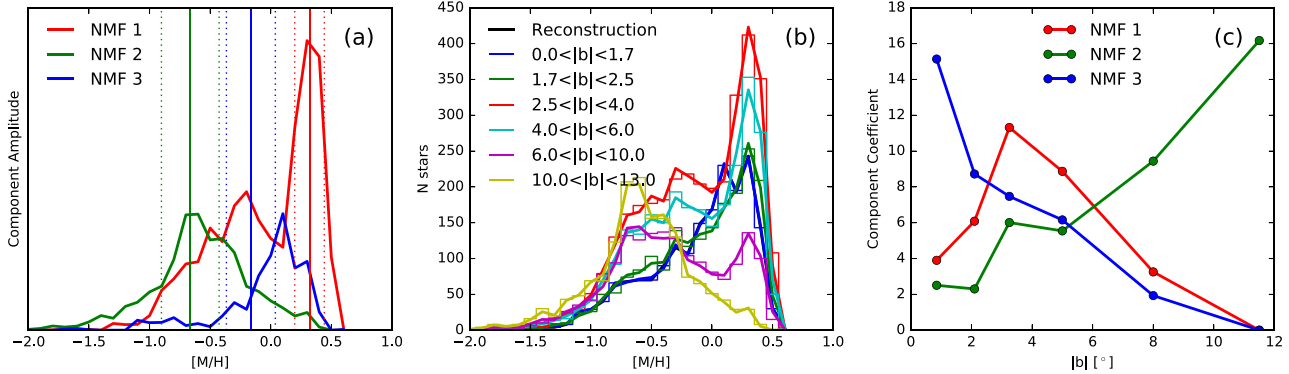


Figure 16. NMF as applied to our bulge MDFs. Panel (a): The components inferred from a three-component decomposition applied to the $|b|$ -dependent MDFs discussed in Section 4.1 and shown in Fig. 9. The vertical solid and dotted lines indicate the mean and $\pm\sigma$ of the three GMM components extracted in Section 4.2. Panel (b): Reconstruction of each MDF using the components in Panel (a). Panel (c): Relative weight of the NMF components in each latitude strip.

line) peaks at the same metallicity as the metal-poor GMM component, with a large width. The metal-richest NMF component (NMF 1, red line) peaks at the same metallicity as the metal-rich GMM component, with a small width in that peak at $[M/H] \sim +0.3$ dex; however, the metal-richest NMF component is the widest of the three, and significantly, shows a secondary peak that coincides with the metallicity of the metal-intermediate GMM component.

These two peaks in the eigenvector (or ‘eigenMDF’) imply that the numbers of stars at these metallicities are somewhat correlated with each other across the range of $|b|$ spanned by the MDFs. An alternative way of interpreting this is as follows: In contrast to representing the MDF by summed Gaussians with smoothly varying relative heights, the MDF can be modelled equally well with a skewed, multi-peaked distribution that is modulated by a second (narrower, symmetric, single-peaked) component (NMF 3, blue line in Fig. 16a, centred at $[M/H] \sim 0$ dex). The near-total dominance of NMF 2 at high $|b|$ suggests that it may represent the combined distributions of inner halo and thick disc stars, which contribute relatively little at low $|Z|$; the inner bulge MDF variations are driven by the relative strength of the skewed, multi-peaked distribution and the symmetric, solar-metallicity distribution.

The details of this comparison depend only weakly on our choice of a three-component decomposition. If a different number of components is chosen, each one’s shape changes slightly, but the general pattern is robust. For example, in a two-component decomposition, the dominant NMF 1 has a shoulder instead of a dip at $[M/H] \sim 0$ dex, while NMF 2 is nearly identical. In a four- or five-component decomposition, the three eigenMDFs shown in Fig. 16(a) are again nearly identical, with additional components providing only slight modulations with very low weights.

Our goal in exploring this technique is to demonstrate an alternative approach to MDF analysis, as a contrast to adopting Gaussian (or similar) bases and identifying or classifying ‘populations’ to associate with known or assumed bulge constituents. These types of non-parametric decompositions may be useful when comparing to, e.g. stellar abundance distributions produced by chemical enrichment models with an extended, complex star formation history, which do not predict distributions easily modelled by a small number of symmetric Gaussians. A deeper exploration of the full information contained in a NMF decomposition (including the optimal number of components) and in other techniques is deferred to future work.

5 DISCUSSION

5.1 MDF complexity from the perspective of other surveys

In Section 4.2, we studied the varying structure of the bulge MDF in several bins sampling different distances from the Galactic mid-plane. In this section, we compare this picture with the general conclusions drawn from the results of optical spectroscopic surveys in the bulge region, some of which found a larger optimal number of components and others which found a smaller number.

A complex picture, with a higher number of components, emerged as a main outcome of the ARGOS survey. The primary targets were bulge RC stars that were observed in the calcium triplet (CaT) region at $R = \lambda/\Delta\lambda \sim 11\,000$. A full-spectrum-fitting approach was used to estimate fundamental parameters, including metallicity (but not T_{eff} , which was fixed from photometric calibrations), with an uncertainty of ~ 0.1 dex. The structure of the MDF was studied from three latitude bins at $b = -5^\circ, -7.5^\circ$, and -10° (within $l \pm 15^\circ$), each containing between 2000 and 4000 stars. From these data, Ness et al. (2013) found the MDF shape to be optimally decomposed by up to five Gaussian components, lettered A through E in order of decreasing metallicity, with the three most metal-rich components (at $[\text{Fe}/\text{H}] = +0.1, -0.28$, and -0.68 dex) dominating the density structure of the MDF. An overall comparable picture was reached from the analysis of the whole sample of microlensed bulge dwarf stars of Bensby et al. (2017). Their MDF, based on a small sample of 90 stars but observed at high resolution ($R = 42\,000\text{--}48\,000$), suggested the presence of up to five peaks, at similar positions to those of ARGOS.

Our results (Fig. 9) are comparable with those of ARGOS, at least with regards to the number of Gaussian components needed to reproduce the observed MDF at $[\text{Fe}/\text{H}] \geq -1$ dex. Our sample has too few stars with $[\text{Fe}/\text{H}] \leq -1$ dex to be represented by extra components in the Gaussian mixture. Thus, the larger size of the ARGOS sample may explain the presence of two additional minor components at $[\text{Fe}/\text{H}] \sim -1.2$ and -1.7 dex. Although the number of components at $[\text{Fe}/\text{H}] \geq -1$ dex is the same, the centres of our metallicity components ($-0.66, -0.17$, and $+0.32$ dex) are not consistent with those of ARGOS, nor are they simply shifted due to a constant offset between the survey metallicity scales. For example, the separation between our metal-rich and metal-intermediate components is 0.10 dex larger than the separation between components A and B of ARGOS (which is statistically

significant, given the errorbars of the respective centroid estimates in both data sets).

A similar multimodal MDF decomposition was obtained from a sample of ~ 7500 stars from APOGEE DR12 (García Pérez et al. 2018). A three Gaussian decomposition was evaluated in a number of mid-size samples (~ 200 stars in average) separating the bulge area in broad spatial (R_{GC} , Z) pixels (see fig. 9 of García Pérez et al. 2018). As a result, four different metallicities (+0.32, +0.0, -0.46 , and -0.83 dex) were found as the centroid of the components considering the results of all fields together (i.e. not every component was detected in all fields). As in the case of ARGOS, the position of the components seems inconsistent with the results obtained here. Note that the results presented in this work, as based on a more recent and improved APOGEE data release, and on a larger number of stars, supersede those of García Pérez et al. (2018).

In contrast, Hill et al. (2011) used a GMM analysis on 219 RC stars to investigate the shape of the MDF in Baade’s Window and found evidence for only two populations, of roughly equal size, located at $[\text{Fe}/\text{H}] = -0.3$ and $+0.3$ dex. A bimodal MDF was also posited to describe a large area below the mid-plane, based on results from the GIBS and GES surveys (Rojas-Arriagada et al. 2014, 2017; Gonzalez et al. 2015; Zoccali et al. 2017). As optical surveys – both of them observing with FLAMES@VLT – the footprints of these observations were concentrated in the area below $b \lesssim -3.5^\circ$ (although GIBS also had a strip of fields at $b = -2^\circ$). As in ARGOS, the targets were RC giants observed in the region around the CaT. In the case of GIBS, spectra were obtained at $R \sim 6500$, and a CaT calibration was used to obtain metallicities with a nominal uncertainty of 0.2 dex (Vásquez et al. 2015). GES data were observed at higher spectral resolution ($R \sim 17000$) and analysed through a full-spectrum-fitting analysis, reaching a metallicity uncertainty of ~ 0.1 – 0.15 dex. In both surveys, the typical sample size per observed field is around a couple hundred stars. The overall picture, consistent between Hill et al. (2011), GES, and GIBS, is of a bimodal bulge MDF. The metallicity positions of the two components vary slightly between the different fields in each study, but are found to be located at about $-0.4 / +0.3$ dex (GIBS) and $-0.35 / +0.40$ dex (GES), compared to ± 0.3 in Hill et al. (2011).

In a different approach, Fragkoudi et al. (2018) compared the shape of the MDF obtained from APOGEE DR13 with a N -body simulation in a number of (l , b) pixels. The simulation tracked the secular evolution of a thin disc and two thick discs into a B/P bulge. From the model, simulated MDFs were ‘observed’ by reproducing the line-of-sight distance sampling of APOGEE. These simulated MDFs were found to qualitatively agree with the data in the sense of being visually bimodal; they also reproduced the spatial variation of the mean metallicity in the bulge region (vertical and longitudinal gradients) and the variation in the relative contribution of particles from the different discs in each observed field.

In spite of the simpler ($N = 2$) decomposition found by these studies to explain the shape of the MDF, the optically derived distributions are in good qualitative agreement with those in our latitude strips spanning similar areas. In fact, the position of our metal-rich component is between the metal-rich components found by GIBS and GES. The shape of the MDF in our latitude bins spanning areas similar to that sampled by GIBS/GES can be visually described as a bimodal distribution (Section 4.1).

Quantitatively, however, this lower metallicity half of the bimodal distribution is better described in our data set by two Gaussian components, which we posit are unresolved when the sample sizes are small and/or the individual metallicity uncertainties are large. Thanks to the large samples we have in each of our latitude strips,

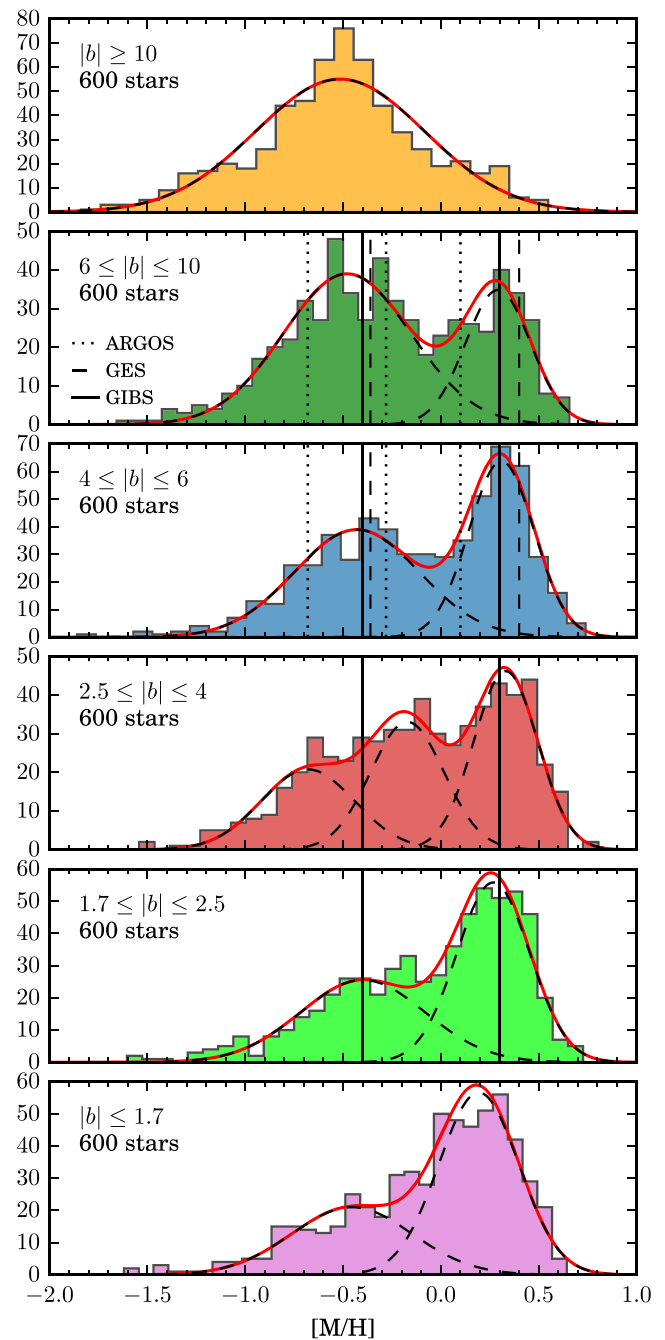


Figure 17. Gaussian mixture decomposition of MDFs designed to reproduce the sampling of previous spectroscopic surveys (analogous to Fig. 9). From each latitude sample, a subsample of 600 stars was taken and their individual measurement errors inflated to 0.1 dex. The best-fitting GMM mixture to each of these resulting samples is overplotted with dashed black and red solid lines, depicting individual components and the total mixture, respectively. The vertical black lines indicate the position of the peaks as found from ARGOS (dotted), GES (dashed), and GIBS (solid) MDF decompositions. They are only displayed in the panels approximately corresponding to the latitude ranges they cover.

and the small individual measurement errors, we can perform an experiment simulating previous optical surveys’ sampling of the bulge MDF, which is illustrated in Fig. 17. We downsampled our six latitude strips by randomly choosing 600 stars from each of them. Although this sample size is a considerable reduction from the current

sample size, it is still larger than that in the typical GIBS/GES fields (~ 200). Gaussian noise was added to the individual metallicities to simulate inflating their uncertainties to an ~ 0.12 dex level.

The optimal GMM solution was computed for each distribution allowing for one to six components and adopting the BIC as model selection criterion. The optimal mixtures fitted to each MDF are displayed on top of them in Fig. 17. As we can see, the MDF of the outermost bin at $|b| \geq 10^\circ$ is now found to be well represented by a single Gaussian component centred at ~ -0.5 dex. On the other hand, the MDFs of the bins inside $|b| = 10^\circ$ are now well represented by bimodal distributions that in every case look as a fair representation of the observed density distribution. In one case, the $2.5^\circ \leq |b| \leq 4^\circ$ strip, the solution is still found to be trimodal and consistent with the result obtained before from the full original sample. In fact, the whole exercise was run several times, and in some cases this distribution was well represented by a bimodal mixture. We choose to leave this instance of the experiment to show that with smaller samples with larger individual errors the results of a GMM run become simpler, or in any case, less robust, since the observed GMM does not provide enough information to be explained by a more complex Gaussian mixture.

Overall, Fig. 17 shows that under similar sampling conditions, the qualitative shape of the MDF in fields over the same spatial area is quite consistent between APOGEE and previous optical surveys. In all cases, the metal-rich component, apparent by simple visual inspection, is appropriately detected and described by a component in the Gaussian mixture. On the other hand, only a larger and more precise sampling allows us to find extra complexity at the subsolar metallicity domain of the MDF.

The metallicity distribution in a given bulge line of sight results from the varying contributions of what could eventually be several structures coexisting in the inner Galaxy. If any pair of them is characterized by intrinsically similar distributions with a significant overlap, they might become difficult to disentangle from the observed MDF. The situation can be even more complex if the intrinsic shape of the distributions departs from the simple assumed Gaussianity (e.g. Section 4.4).

To summarize, the need for three Gaussian components to explain the shape of the bulge MDF qualitatively agrees with the results from the ARGOS survey. It is also in agreement with the density structure reported from other large optical spectroscopic surveys once considerations of sample size and individual error measurements are taken into account. Whether these components represent underlying physical components or not can be assessed from the examination of the distributions of other observed properties.

5.2 The bulge's MDF in the context of bulge formation mechanisms and extragalactic evidence

Our trimodal characterization of the bulge MDF reveals different spatial variations and kinematical correlations for stars in the three metallicity groups. These observations are evidence for understanding the metallicity components as distinct physical populations (see also Queiroz et al. 2020a). It is interesting to review these findings in the context of extragalactic observations and theoretical predictions from bulge formation scenarios. We will start with the highest metallicity stars.

Many lines of observational evidence have been used to argue that the bulge's metal-rich stars are entrained in the B/P X-shaped bar (e.g. Ness et al. 2012; Bensby et al. 2017; Rojas-Arriagada et al. 2017; Zoccali et al. 2017). In the canonical prescription for bar formation, secular evolution proceeds from spontaneous gravitational instabil-

ities that rearrange angular momentum in the pre-existing disc and result in mass transfer, building a stellar concentration in a central bar. Subsequent bending instabilities produce a B/P bulge from the stars being kinematically heated in the vertical direction. The point here is that secular evolution is fundamentally a dissipationless process taking place in a gas-poor phase of galactic evolution, driven by stellar dynamics. This process has implications for the ages and present-day kinematics of stars entrained in the bar.

With respect to age, one must be careful to distinguish between ages of the bar stars and the age of the bar structure itself. The high metallicity of the stars argued to be in the bar may imply younger ages, relative to the rest of the bulge, which may suggest that the bar itself is a relatively recent addition to the MW (Di Matteo 2016; Fragkoudi et al. 2018; Wegg et al. 2019, Lian et al. in preparation, but for a different view, Bovy et al. 2019). It has been established from the study of the redshift-dependent fraction of galactic bars, in observations (Sheth et al. 2008) and cosmological simulations (Kraljic et al. 2012; Fragkoudi et al. 2020), that bars appear at lookback times of ~ 8 Gyr in MW-mass galaxies – that is, at a later time compared with the main epoch of mass assembly and star formation. However, it has also been shown that the presence of a rotating non-axisymmetric bar potential can induce a loss of angular momentum in gas clouds, which fall to the centre, trigger star formation (Ellison et al. 2011), and produce stars the same age or younger than the initial bar itself. Thus, the ages of individual stars constrained, by other means, to be part of the bar population are essential for disentangling the formation history of not only the bar's stars but also the bar's structure. Detections of a significant fraction of relatively young stars associated with the bar are important lines of evidence (e.g. Bensby et al. 2011, 2013, 2017; Schultheis et al. 2017; Hasselquist et al. 2020).

The kinematics of our metal-richest stars support the interpretation of bar membership. The cylindrical rotation pattern exhibited by metal-rich stars in Figs 14 and 15 is an observational outcome predicted by N -body simulations of B/P bulges formed from the secular evolution of the disc through bar formation and buckling (Zhao 1996). It is also a phenomenology observed in external local galaxies with central B/P bar-dominated morphology (Molaeinezhad et al. 2016). The slightly enhanced rotation pattern towards the mid-plane observed in our data represents an additional signature of bar-dominated dynamics (Gómez et al. 2018). The $\sigma(V_{GC})$ maps (Fig. 14) show that this component has a steep vertical kinematical gradient, consistent with the progressive dominance of bar streaming orbits towards the mid-plane. At the same time, such an orbital structure may explain the central $\sigma(V_{GC})$ enhancement (at $|\ell| \leq 4^\circ$, $|b| \leq 4^\circ$) as an effect of the lines of sight crossing inner regions where the several families of orbits supporting the bar structure intersect each other.

Let us now turn our attention to the group of metal-intermediate bulge stars. This component, in our analysis, comprises most of the stars identified as belonging to the metal-poor component in optical studies such as GES/GIBS (Section 5.1), so that comparisons between them are pertinent in what follows. As shown in Figs 14 and 15, this component is overall kinematically hotter than the metal-rich one but has a less pronounced enhancement in dispersion towards the centre. However, it displays a more coherent cylindrical rotation than the metal-rich stars. Can these stars be associated with the buckling instability of stars formed in the disc or with a central *in-situ* formation scenario instead?

Stars in this metallicity range [(-0.5)–(0.0) dex] have been interpreted as the bulk of the early thin disc rearranged into the B/P bulge by the bar instability processes (Ness et al. 2013). In this sense,

the splitting of the RC luminosity distribution (as a proxy of distance) was taken by Ness et al. (2012) as evidence of them participating in the X-shape. However, in the same metallicity range, GES data suggest a less structured, more homogeneous spatial distribution (Rojas-Arriagada et al. 2017). In addition, such a scenario conflicts with the relatively high levels of α -element enhancement observed for these stars in the APOGEE data (in the same $R_{GC} \leq 3.5$ kpc sampled in this work; Rojas-Arriagada et al. 2019; Queiroz et al. 2020b); this enhancement is instead indicative of a rapid star forming episode (i.e. the chemical evolution models of Rojas-Arriagada et al. 2017; Lian et al. 2020; Matteucci et al. 2020).

Using simulations, Debattista et al. (2017) argue that the bar could kinematically separate co-spatial populations according to their initial in-plane radial random motions. This *kinematical fractionation* determines the height stars can reach during the buckling bar event, with those being initially kinematically cooler forming a stronger X-shaped spatial distribution, while hotter populations ending instead in a weaker bar vertically thicker box. As older stars have hotter initial kinematics and different chemistry, this scenario implies a continuum of current bulge stellar content properties as a function of metallicity. In particular, this scenario can reconcile the weaker (or absent) RC luminosity distribution splitting, and the particular alpha-enhancement of stars over the metallicity range of our metal-intermediate component, given an appropriate chemo-dynamical structure of the initial disc(s).

In a different proposal, Di Matteo (2016) and Fragkoudi et al. (2018) associate these metal-intermediate stars with the larger thick disc in order to reproduce their observed chemical and kinematical distributions. Although a qualitatively similar chemistry is found for the sequences of bulge and thick disc stars in the $[\alpha/\text{Fe}]$ versus $[\text{Fe}/\text{H}]$ plane, a more metal-rich position has been found for the bulge ‘knee’ (Cunha & Smith 2006; Rojas-Arriagada et al. 2017, although this difference might be smaller if comparing with the inner thick disc; Zasowski et al. 2019), indicating a higher early SFR in the bulge than that in the thick disc inside the solar circle. This has been also suggested to be the case from the comparison of microlensed bulge dwarf stars with local thick disc dwarfs (Bensby et al. 2013, 2017). In addition, these works suggest that bulge stars might be overall slightly more alpha-enhanced than thick disc stars. This latter result has been confirmed in Queiroz et al. (2020a) by comparing the bulge and inner thick disc ($|X| < 5$ kpc, $|Y| < 3.5$ kpc, and $|Z| < 1$ kpc) chemical distributions.

Alternatively, these stars have been proposed as a morphologically classical/spheroidal bulge component (Babusiaux et al. 2010; Hill et al. 2011; Rojas-Arriagada et al. 2017). In Λ CDM, this component corresponds to mass assembled from the dissipationless merging of stellar substructures. There are a couple of observational constraints limiting the plausibility of this scenario: (i) a bulge assembled from many different building blocks might lead to a rather chemically inhomogeneous structure, unless the independent building blocks were already uniform to begin with. What is observed, however, is a tight, seemingly singular chemical evolutionary sequence, at least in the alpha elements (Rojas-Arriagada et al. 2019; Zasowski et al. 2019). (ii) This hypothesis would require a rather high mass for the individual building blocks. Indeed, if their individual potential wells were shallow, the consequent slow SFR would be unable to produce a substantial amount of stars of a high enough metallicity to match our component distribution.

Moreover, the analysis of 30 high-resolution hydrodynamical cosmological simulations has shown that the amount of accreted stars in the inner/bulge region of MW-mass galaxies is very low, even negligible in some cases. Indeed, most of the (already low)

accreted components of these simulated bulges come from a very small number of building blocks (Gargiulo et al. 2019). This result is confirmed by the analysis of zoom-in simulations of Fragkoudi et al. (2020), which has revealed that the modelled galaxies that most closely resemble the chemodynamical properties of the MW are characterized by a quiescent merger history: the last major merger happening long before bar formation, and overall, with a negligible fraction of *ex-situ* stars in the bulge region.

Finally, an *in-situ* formation scenario may provide an appropriate explanation for the present-day observed properties of metal-intermediate stars. Observations at high redshifts ($z = 1.5\text{--}2.2$) have revealed the presence of central stellar overdensities in galaxies that otherwise are still actively forming stars in their gas-rich discs in an inside-out way (Tacchella et al. 2015; Nelson et al. 2016). This implies that mature bulges are already in place early during the main epoch of star formation. This mechanism rapidly forms rotating bulges out of the gas-rich disc in an internal dissipative process, without the need for major mergers (Tadaki et al. 2017). Theoretical prescriptions of such a scenario include the formation of giant disc gas clumps, their migration and coalescence (Immeli et al. 2004; Elmegreen, Bournaud & Elmegreen 2008) at the centre, or global violent disc instabilities leading to a massive central accumulation of star-forming gas (Dekel & Burkert 2014). In addition, in the aforementioned study of cosmological simulations by Gargiulo et al. (2019), they found that in 75 per cent of the cases most of the stars found in the bulge at present day formed *in situ* in the inner regions of the simulated galaxies. A fast and early vigorous star formation in gas-rich high-redshift environments can produce a bulge with the properties observed in our metal-intermediate stars: a spatially extended, rotating, kinematically hot, relatively metal-rich and alpha-enhanced stellar structure composed predominantly by old stars. In this sense, our metal-intermediate stars may correspond to the present-day structure assembled centrally *in situ* and early in the formation of the MW.

Our metal-poor component amounts for a relatively constant and minor fraction of the stars below solar metallicity. They display a tight but slow cylindrical rotation pattern and are kinematically hot and isotropic over most of the studied area. In Ness et al. (2013), these stars were associated with the thick disc present in the inner Galaxy before the instability event. This interpretation conflicts with the peak of this component being located at $[\text{M}/\text{H}] = -0.66$ dex, which is around 0.3/0.4 dex lower than the peak of the thick disc at $3 \leq R_{GC} \leq 5$ kpc (Hayden et al. 2015; Queiroz et al. 2020b). On the other hand, they are too metal rich to be associated with the inner halo. None the less, our sample contains only a minor fraction of stars with $[\text{M}/\text{H}] \leq -1$ dex, and therefore we cannot rule out that some of these stars belong to the stellar halo. In fact, from the early evidence of Minniti (1996a) it is known that giant stars more metal poor than $[\text{M}/\text{H}] = -1$ dex display halo-like kinematics, and should be naturally found in the inner Galactic regions but as an overall very small fraction.

In that metal-poor metallicity range, Kunder et al. (2020) have proposed the existence of two different components traced by the intrinsically metal-poor old RR Lyrae stars, with the most centrally concentrated one proposed to be a classical bulge in the sense of being produced by an accretion event at high redshift. Whether our metal-poor stars can be associated with the stellar populations being traced by RR Lyrae at $[\text{Fe}/\text{H}] \sim -1$ dex remains to be proven. None the less, a merger scenario for this component has to account, as in the case of metal-intermediate stars, for their low dispersion of α -elements. The latter also challenges a potential interpretation of this component as the product of accreted globular cluster stars; not

only the chemical homogeneity of α -elements is not guaranteed, but also the peak in metallicity of this component is in between the two peaks of the metallicity distribution of bulge globular clusters (Bica, Ortolani & Barbuy 2016, their fig. 4). Field stars with chemical anomalies have been reported in this metallicity regime (Schiavon et al. 2017; Fernández-Trincado et al. 2019), but they are estimated to amount for ~ 2 per cent of the total budget mass. The more recent work of Horta et al. (2020) estimates that up to 1/3 of all bulge metal-poor stars (with $[\text{Fe}/\text{H}] \leq -0.8$ dex) might be accreted, as identified from their chemical composition resembling those of low-mass satellites of the MW. Further data will be needed to resolve the formation scenario for these stars.

6 SUMMARY

We have investigated the shape of the bulge MDF from a sample of $\sim 13\,000$ giant stars spatially located in the bulge region coming from the APOGEE DR16 and further internal incremental data releases. This sample allowed us to study the inner-Galaxy stellar populations over the whole $|\ell| \leq 11^\circ$, $|b| \leq 13^\circ$ area, and down to the mid-plane where the high levels of extinction have hindered previous attempts from optical surveys.

Our main findings are as follows:

(i) The effect of the APOGEE selection function on the shape of the MDF is rather small and does not impact the complex density structure revealed by our analysis.

(ii) The $R_{\text{GC}} = 3.5$ kpc limit often adopted to spatially separate bulge/disc stellar samples finds qualitative justification from the changes in kinematical patterns characterized from our data around that limit.

(iii) The shape and spatial variations of the bulge MDF can be explained as the varying contribution of three main components at $[\text{Fe}/\text{H}] \geq -1$ dex.

(iv) Our GMM analysis and robustness assessment show that the trimodal Gaussian mixture correctly captures the density structure of the MDF over the whole set of latitude strips we adopted to separate the bulge sample. We estimate that bulge populations dominate the line-of-sight samples inside $|b| = 8^\circ$. Over that region, the mean metallicity of the components remains rather constant, and the widths present only mild variations. The relative weights of the metal-rich and metal-intermediate components determine the vertical variations of the density shape of the MDF, and so, naturally explain the vertical variations of summary quantities such as the metallicity gradient.

(v) In the same vein, we show that the bimodal Gaussian mixture found to well represent the MDF from optical surveys data could be an effect of the combination of relatively small sampling and larger individual measurement errors.

(vi) The kinematics of our three metallicity components show some differences that we relate to their different orbital structure.

(vii) Finally, we cast our results in the context of previous works on the bulge MDF, as well as of pieces of evidence from the study of high-redshift galaxies and theoretical formation models.

With the present data, we are not in a position to settle down a definitive physical interpretation of the nature and origin of the bulge components. None the less, our data allowed us to trace the big empirical picture, conciliating previous observational evidence and contributing to assemble a state-of-the-art observational picture of the bulge stellar populations down to the mid-plane. The complexities revealed here and in other studies highlight the need of considering both local (resolved stellar populations) and high redshift (integrated properties) evidence to piece together the sequence of events behind

galaxy formation and evolution. This observational approach may lead us to assemble, with the help of theoretical models/prescriptions, a scenario providing a satisfactory and harmonious account of galaxy phenomenology from $z = 2$ to the present-day picture offered by the MW.

ACKNOWLEDGEMENTS

ARA and GZ are grateful to the Observatoire de la Côte d’Azur for hospitality during visits that greatly benefited the analysis in this paper. ARA acknowledges support from FONDECYT through grant 3180203. The authors thank R. M. Rich for his compelling analogy between bulge metallicity components and martinis during the 2018 European Southern Observatory (ESO) bulge meeting in Pucón, which served as an inspiration for this paper. DG gratefully acknowledges support from the Chilean Centro de Excelencia en Astrofísica y Tecnologías Afines (CATA) BASAL grant AFB-170002. DG and DM gratefully acknowledge support from the Chilean Centro de Excelencia en Astrofísica y Tecnologías Afines (CATA) BASAL grant AFB-170002. DM also thanks support from proyecto FONDECYT Regular No. 1170121. DG acknowledges financial support from the Dirección de Investigación y Desarrollo de la Universidad de La Serena through the Programa de Incentivo a la Investigación de Académicos (PIA-DIDULS). DAGH acknowledges support from the State Research Agency (AEI) of the Spanish Ministry of Science, Innovation and Universities (MCIU) and the European Regional Development Fund (FEDER) under grant AYA2017-88254-P. JGF-T is supported by FONDECYT No. 3180210 and Becas Iberoamérica Investigador 2019, Banco Santander Chile. AM acknowledges financial support from FONDECYT Regular 1181797 and funding from the Max Planck Society through a Partner Group grant. CC acknowledges partial support from DFG Grant CHI 188/2-1 and from the ChETEC COST Action (CA16117), supported by COST (European Cooperation in Science and Technology).

This work has made use of data from the European Space Agency (ESA) mission *Gaia* (<https://www.cosmos.esa.int/gaia>), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC; <https://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.

Funding for the SDSS-IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss.org.

SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU)/University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de

México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

DATA AVAILABILITY

The data used in this article are from an internal incremental release of the SDSS-IV/APOGEE survey, following the SDSS-IV public DR16. This means that a minor fraction of the sample studied here is not publicly available now but will be included in the final public data release of SDSS-IV.

REFERENCES

- Abadi M. G., Navarro J. F., Steinmetz M., Eke V. R., 2003, *ApJ*, 597, 21
- Ahumada R. et al., 2020, *ApJS*, 249, 3
- Athanassoula E., 2005, *MNRAS*, 358, 1477
- Babusiaux C. et al., 2010, *A&A*, 519, A77
- Bailer-Jones C. A. L., Rybizki J., Fouesneau M., Mantelet G., Andrae R., 2018, *AJ*, 156, 58
- Barbuy B., Chiappini C., Gerhard O., 2018, *ARA&A*, 56, 223
- Baugh C. M., Cole S., Frenk C. S., 1996, *MNRAS*, 283, 1361
- Bensby T. et al., 2011, *A&A*, 533, A134
- Bensby T. et al., 2013, *A&A*, 549, A147
- Bensby T. et al., 2017, *A&A*, 605, A89
- Bica E., Ortolani S., Barbuy B., 2016, *Publ. Astron. Soc. Aust.*, 33, e028
- Bland-Hawthorn J., Gerhard O., 2016, *ARA&A*, 54, 529
- Blanton M. R. et al., 2017, *AJ*, 154, 28
- Bovy J., 2015, *ApJS*, 216, 29
- Bovy J. et al., 2012, *ApJ*, 759, 131
- Bovy J., Leung H. W., Hunt J. A. S., Mackereth J. T., García-Hernández D. A., Roman-Lopes A., 2019, *MNRAS*, 490, 4740
- Bowen I. S., Vaughan A. H., Jr., 1973, *Appl. Opt.*, 12, 1430
- Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, *MNRAS*, 427, 127
- Catchpole R. M., Whitlock P. A., Feast M. W., Hughes S. M. G., Irwin M., Alard C., 2016, *MNRAS*, 462, 3765
- Combes F., Sanders R. H., 1981, *A&A*, 96, 164
- Combes F., Debbasch F., Friedli D., Pfenniger D., 1990, *A&A*, 233, 82
- Cunha K., Smith V. V., 2006, *ApJ*, 651, 491
- Cunha K., Sellgren K., Smith V. V., Ramirez S. V., Blum R. D., Terndrup D. M., 2007, *ApJ*, 669, 1011
- de Vaucouleurs G., 1964, in Kerr F. J., ed., *Proc. IAU Symp. 20, The Galaxy and the Magellanic Clouds*. Kluwer, Dordrecht, p. 195
- Debatista V. P., Ness M., Gonzalez O. A., Freeman K., Zoccali M., Minniti D., 2017, *MNRAS*, 469, 1587
- Dekel A., Burkert A., 2014, *MNRAS*, 438, 1870
- Di Matteo P., 2016, *Publ. Astron. Soc. Aust.*, 33, e027
- Dotter A., 2016, *ApJS*, 222, 8
- Dwek E. et al., 1995, *ApJ*, 445, 716
- Eisenstein D. J. et al., 2011, *AJ*, 142, 72
- Ellison S. L., Nair P., Patton D. R., Scudder J. M., Mendel J. T., Simard L., 2011, *MNRAS*, 416, 2182
- Elmegreen B. G., Bournaud F., Elmegreen D. M., 2008, *ApJ*, 688, 67
- Fernández-Trincado J. G., Beers T. C., Tang B., Moreno E., Pérez-Villegas A., Ortigoza-Urdaneta M., 2019, *MNRAS*, 488, 2864
- Fragkoudi F., Di Matteo P., Haywood M., Schultheis M., Khoperskov S., Gómez A., Combes F., 2018, *A&A*, 616, A180
- Fragkoudi F. et al., 2020, *MNRAS*, 494, 5936
- Fulbright J. P., McWilliam A., Rich R. M., 2006, *ApJ*, 636, 821
- Gaia Collaboration, 2018, *A&A*, 616, A1
- García Pérez A. E. et al., 2016, *AJ*, 151, 144
- García Pérez A. E. et al., 2018, *ApJ*, 852, 91
- Gargiulo I. D. et al., 2019, *MNRAS*, 489, 5742
- Gómez A., Di Matteo P., Schultheis M., Fragkoudi F., Haywood M., Combes F., 2018, *A&A*, 615, A100
- Gonzalez O. A., Rejkuba M., Zoccali M., Valent E., Minniti D., Tobar R., 2013, *A&A*, 552, A110
- Gonzalez O. A. et al., 2015, *A&A*, 584, A46
- Gunn J. E. et al., 2006, *AJ*, 131, 2332
- Hasselquist S. et al., 2020, preprint ([arXiv:2008.03603](https://arxiv.org/abs/2008.03603))
- Hayden M. R. et al., 2015, *ApJ*, 808, 132
- Hernitschek N. et al., 2019, *ApJ*, 871, 49
- Hernquist L., 1990, *ApJ*, 356, 359
- Hill V. et al., 2011, *A&A*, 534, A80
- Holtzman J. A. et al., 2018, *AJ*, 156, 125
- Horta D. et al., 2020, preprint ([arXiv:2007.10374](https://arxiv.org/abs/2007.10374))
- Hurley P. D., Oliver S., Farrar D., Leboutteiller V., Spoon H. W. W., 2014, *MNRAS*, 437, 241
- Igual J., Llinares R., 2008, *IEEE J. Sel. Top. Signal Process.*, 2, 697
- Immeli A., Samland M., Gerhard O., Westera P., 2004, *A&A*, 413, 547
- Jönsson H. et al., 2018, *AJ*, 156, 126
- Jönsson H. et al., 2020, *AJ*, 160, 120
- Kormendy J., Kennicutt R. C., Jr., 2004, *ARA&A*, 42, 603
- Kraljic K., Bournaud F., Martig M., 2012, *ApJ*, 757, 60
- Kroupa P., 2001, *MNRAS*, 322, 231
- Kunder A. et al., 2020, *AJ*, 159, 270
- Lee D. D., Seung H. S., 1999, *Nature*, 401, 788
- Leung H. W., Bovy J., 2019, *MNRAS*, 483, 3255
- Lian J. et al., 2020, *MNRAS*, 497, 3557
- López-Corredoira M., Cabrera-Lavers A., Gerhard O. E., 2005, *A&A*, 439, 107
- McWilliam A., Rich R. M., 1994, *ApJS*, 91, 749
- Majewski S. R. et al., 2017, *AJ*, 154, 94
- Marigo P. et al., 2017, *ApJ*, 835, 77
- Matteucci F., Vassini A., Grisoni V., Schultheis M., 2020, *MNRAS*, 494, 5534
- Minniti D., 1996a, *ApJ*, 459, 175
- Minniti D., 1996b, *ApJ*, 459, 579
- Minniti D., Olszewski E. W., Liebert J., White S. D. M., Hill J. M., Irwin M. J., 1995, *MNRAS*, 277, 1293
- Minniti D. et al., 2010, *New Astron.*, 15, 433
- Miyamoto M., Nagai R., 1975, *PASJ*, 27, 533
- Molaeinezhad A., Falcón-Barroso J., Martínez-Valpuesta I., Khosroshahi H. G., Balcells M., Peletier R. F., 2016, *MNRAS*, 456, 692
- Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493
- Nelson E. J. et al., 2016, *ApJ*, 828, 27
- Ness M. et al., 2012, *ApJ*, 756, 22
- Ness M. et al., 2013, *MNRAS*, 430, 836
- Ness M. et al., 2016, *ApJ*, 819, 2
- Nidever D. L. et al., 2015, *AJ*, 150, 173
- Paxton B., Bildsten L., Dotter A., Herwig F., Lesaffre P., Timmes F., 2011, *ApJS*, 192, 3
- Pietrzyński G. et al., 2013, *Nature*, 495, 76
- Queiroz A. B. A. et al., 2018, *MNRAS*, 476, 2556
- Queiroz A. B. A. et al., 2020a, preprint ([arXiv:2007.12915](https://arxiv.org/abs/2007.12915))
- Queiroz A. B. A. et al., 2020b, *A&A*, 638, A76
- Raha N., Sellwood J. A., James R. A., Kahn F. D., 1991, *Nature*, 352, 411
- Ramírez S. V., Stephens A. W., Frogel J. A., DePoy D. L., 2000, *AJ*, 120, 833
- Rattenbury N. J., Mao S., Sumi T., Smith M. C., 2007, *MNRAS*, 378, 1064
- Rich R. M., 1988, *AJ*, 95, 828
- Rich R. M., Origlia L., Valenti E., 2007, *ApJ*, 665, L119
- Rich R. M., Origlia L., Valenti E., 2012, *ApJ*, 746, 59
- Robin A. C., Marshall D. J., Schultheis M., Reylé C., 2012, *A&A*, 538, A106
- Rodríguez-Fernández N. J., Combes F., 2008, *A&A*, 489, 115
- Rojas-Arriagada A. et al., 2014, *A&A*, 569, A103
- Rojas-Arriagada A. et al., 2017, *A&A*, 601, A140
- Rojas-Arriagada A., Zoccali M., Schultheis M., Recio-Blanco A., Zasowski G., Minniti D., Jönsson H., Cohen R. E., 2019, *A&A*, 626, A16
- Sadler E. M., Rich R. M., Terndrup D. M., 1996, *AJ*, 112, 171
- Schiavon R. P. et al., 2017, *MNRAS*, 465, 501
- Schlegel D. J., Finkbeiner D. P., Davis M., 1998, *ApJ*, 500, 525
- Schönrich R., Binney J., Dehnen W., 2010, *MNRAS*, 403, 1829

- Schultheis M. et al., 2017, *A&A*, 600, A14
 Schultheis M., Rich R. M., Origlia L., Ryde N., Nandakumar G., Thorsbro B., Neumayer N., 2019, *A&A*, 627, A152
 Sheth K. et al., 2008, *ApJ*, 675, 1141
 Tacchella S. et al., 2015, *Science*, 348, 314
 Tadaki K.-i. et al., 2017, *ApJ*, 834, 135
 Vázquez S., Zoccali M., Hill V., Gonzalez O. A., Saviane I., Rejkuba M., Battaglia G., 2015, *A&A*, 580, A121
 Wegg C., Gerhard O., 2013, *MNRAS*, 435, 1874
 Wegg C., Rojas-Arriagada A., Schultheis M., Gerhard O., 2019, *A&A*, 632, A121
 Weiland J. L. et al., 1994, *ApJ*, 425, L81
 Wilson J. C. et al., 2019, *PASP*, 131, 055001
 Zasowski G. et al., 2013, *AJ*, 146, 81
 Zasowski G., Ness M. K., García Pérez A. E., Martínez-Valpuesta I., Johnson J. A., Majewski S. R., 2016, *ApJ*, 832, 132
 Zasowski G. et al., 2017, *AJ*, 154, 198
 Zasowski G. et al., 2019, *ApJ*, 870, 138
 Zhao H., 1996, *MNRAS*, 283, 149
 Zoccali M., Hill V., Lecureur A., Barbuy B., Renzini A., Minniti D., Gómez A., Ortolani S., 2008, *A&A*, 486, 177
 Zoccali M. et al., 2014, *A&A*, 562, A66
 Zoccali M. et al., 2017, *A&A*, 599, A12

APPENDIX A: DISTANCE VALIDATION

We perform a number of tests and comparisons to verify the quality of our derived spectrophotometric distances, which are summarized in Figs A1 and A2.

In the left-hand panel of Fig. A1, we show the comparison of our spectrophotometric distances with those from the catalogue of *Gaia*-based Bayesian distances of Bailer-Jones et al. (2018, hereafter CBJ). We cross-match our catalogue for the APOGEE giant sample ($\log g \leq 3.5$) with that of CBJ and select stars with distances smaller than 2.5 kpc and relative error in parallaxes $\sigma_{\pi}/\pi \leq 0.1$. We find a

median offset of $(d - d_{\text{CBJ}})/d = 3.4$ per cent (i.e. 85 pc at 2.5 kpc), with a standard deviation of 21.1 per cent.

The middle panel displays the comparison of our distances with those from the neural network ASTRONN (aNN; Leung & Bovy 2019), for giant stars within 10 kpc. We find a median offset of $(d - d_{\text{aNN}})/d = 4.8$ per cent, with a standard deviation of 21.1 per cent.

In the right-hand panel, our distances are compared with those computed with the STARHORSE code (Queiroz et al. 2018; SH) for giant stars within 10 kpc of the Sun. Here, we find even better agreement than that with other two methods, with a median offset of $(d - d_{\text{SH}})/d = 2.1$ per cent and a standard deviation of 13.4 per cent.

Fig. A2 shows distance properties for some different systems using our spectrophotometric distances. In the left-hand panel, we show the fractional distance *spread* of several open clusters whose members are selected using dynamics only, with no assumption of distance (e.g. from isochrone-fitting; Poovelil et al. in preparation). The black histogram is the standard deviation of each of the cluster members' distances, and the red histogram shows the median absolute deviation. We adopt the more conservative typical uncertainty of 25 per cent (vertical dashed line) for our sample in this paper.

In the right-hand panel of Fig. A2, we show the distance distribution of the entire APOGEE sample (grey histogram, predominantly within $d < 10$ kpc), compared to two dwarf galaxies. Stars targeted in the Sagittarius (Sgr) dwarf galaxy are shown in purple, and stars targeted in the Large Magellanic Cloud (LMC) are shown in blue. We see the expected disc and halo contamination, but the samples are dominated by stars at distances entirely consistent with previously determined literature values – here, shown as $d_{\text{Sgr}} = 28.2$ kpc based on RR Lyrae (Hernitschek et al. 2019) and $d_{\text{LMC}} = 49.88$ kpc based on eclipsing binaries (Pietrzyński et al. 2013). The width of these peaks is entirely consistent with the 25 per cent uncertainty established from the left-hand panel.

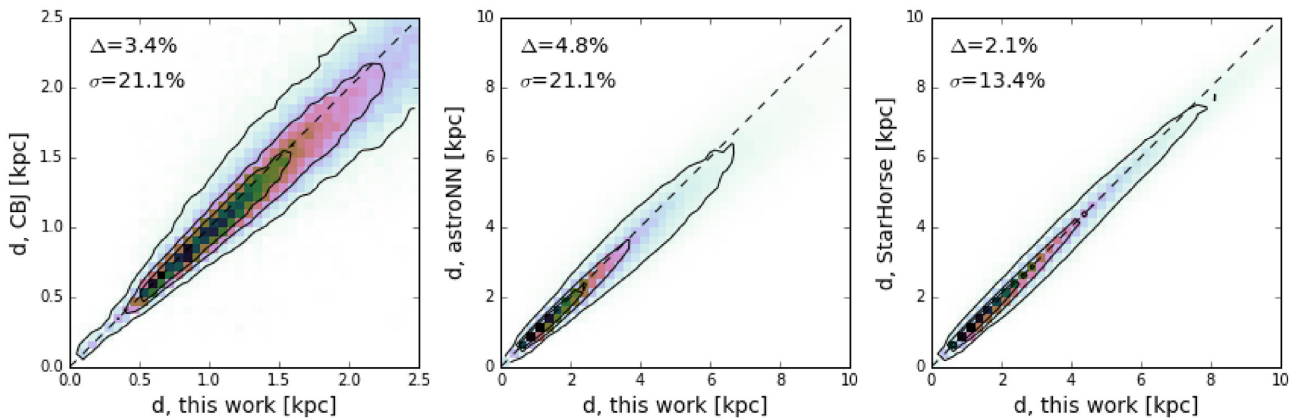


Figure A1. Comparison of the spectro-photometric distances computed in this work (Section 2.2) with three others: *Gaia* Bayesian distances from Bailer-Jones et al. (2018), the neural network ASTRONN distances from Leung & Bovy (2019), and the spectrophotometric distances computed with the STARHORSE code (Queiroz et al. 2018). In each case, we display the direct comparison of our spectrophotometric distances with those of the different catalogues for the APOGEE DR16 sources in common. The dashed black line indicates the 1:1 relation, and the contours trace the 50 per cent, 25 per cent, and 5 per cent isodensity lines. We find generally good agreement with all three methods.

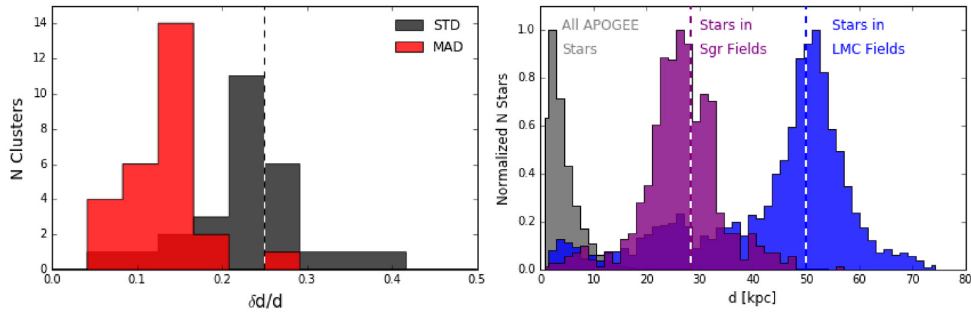


Figure A2. *Left-hand panel:* distance dispersion for stars in kinematically identified open clusters (Poovelil et al. in preparation), measured using the standard deviation (STD, black histogram) and the median absolute deviation (MAD, red histogram) of the cluster members' distance moduli. We adopt the more conservative typical uncertainty of 25 per cent (vertical dashed line) for our sample in this paper. *Right-hand panel:* distance distribution of stars targeted in fields towards the Large Magellanic Cloud (LMC) in blue and towards the Sagittarius (Sgr) dwarf galaxy in purple, compared with the distance distribution of the general APOGEE sample in grey. All three histograms have been scaled to peak at 1. The vertical dashed lines indicate the $d = 28.2$ kpc distance to the Sgr core based on RRL (Hernitschek et al. 2019) and the $d = 49.88$ kpc distance to the LMC based on eclipsing binaries (Pietrzyński et al. 2013).

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.