

**Department of Physics and Astronomy  
University of Heidelberg**

**Verification and chemo-dynamic analysis  
of the 4MOST Milky Way Halo  
High-Resolution Survey**

Master's Thesis in Physics  
submitted by

**Jens Richter**

born in Saarbrücken (Germany)

**07.11.2023**

This Master Thesis has been carried out by Jens Richter at the  
Zentrum für Astronomie der Universität Heidelberg, Landessternwarte  
under the supervision of  
Prof. Norbert Christlieb

## Abstract

The stellar halo contains a lot of the Galaxy's most metal-poor stars and is also made up to a significant extent by accreted material. As such, it is an appropriate site to study the formation, early evolution, and merger history of the Galaxy. The 4MOST S2 survey will target the Milky Way halo in high resolution and provide the spectra of about 2 million stars over the course of 5 years. The subject of this thesis is the target catalogue of 4MOST S2, with the aim of verifying its selection and providing a kinematic and chemical analysis of the catalogue's content using Gaia DR3 and GALAH. To achieve these goals, the Pristine-survey-based metallicity estimates used in the selection of the catalogue are verified via several reference samples including GALAH, SAGA, and APOGEE. The comparisons indicate that the Pristine calibration underestimates metallicities by about 0.1 dex, but otherwise agrees very well. In addition, the Besançon galaxy model is used as a reference for the expected spatial distribution of the target catalogue. The agreement is good overall, but towards the inner Galaxy and close to the Sun, there are significant deviations, which are explained by extinction effects and metallicity uncertainties, respectively. A stepwise application of the selection criteria to reference catalogues gives insight into the effects of the individual criteria and provides further verification. For the kinematic analysis, the target catalogue is restricted to the subset with available radial velocities. After calculating the necessary kinematic parameters, a kinematic halo selection is carried out and then verified via its stellar number distribution, with a single power-law slope slightly higher than common literature values at  $\alpha \approx 4.4$  being found. The kinematic space is explored for any substructures, leading to Gaia-Enceladus being clearly identified visually, whereas the hierarchical clustering algorithm HDBSCAN identifies further substructures, which are found to correspond to Sequoia, Thamnos, and the Helmi streams. Finally, the neutron-capture element abundance patterns are investigated using data from GALAH. In particular, by estimating their similarity to reference patterns, about 319, 253, and 283 candidates are identified for enrichment via the r-process, s-process, and weak r-process, respectively.

## Zusammenfassung

Der stellare Halo beinhaltet viele der metallärmsten Sterne der Galaxie und ist zu einem bedeutenden Anteil aus akkretiertem Material zusammengesetzt. Das macht ihn zu einem geeigneten Ort zur Untersuchung der Entstehung, frühen Evolution, sowie der Verschmelzungsgeschichte der Galaxie. Die 4MOST S2 survey zielt darauf ab, den Halo der Milchstraße in hoher Auflösung zu beobachten und über einen Zeitraum von 5 Jahren die Spektren von circa 2 Millionen Sternen zur Verfügung zu stellen. Gegenstand dieser Arbeit ist der Zielkatalog von 4MOST S2, mit der Zielsetzung, dessen Selektion zu verifizieren und eine kinematische und chemische Analyse seiner Inhalte mithilfe von Gaia DR3 und GALAH durchzuführen. Zum Erreichen dieser Zwecke werden die in der Selektion verwendeten Pristine-Metallizitätsabschätzungen mithilfe mehrerer Referenzstichproben wie GALAH, SAGA und APOGEE verifiziert. Die Vergleiche zeigen, dass die Pristine-Kalibration die Metallizitäten um circa 0.1 dex unterschätzt, aber die Übereinstimmung ansonsten sehr gut ist. Zusätzlich wird das Besançon Galaxy Model als Referenz für die erwartete räumliche Verteilung des Zielkatalogs verwendet. Die Übereinstimmung ist insgesamt gut, aber in Richtung der inneren Galaxie und nahe der Sonne gibt es signifikante Abweichungen, wobei Ersteres durch Extinktionseffekte und Letzteres durch Metallizitätsunsicherheiten erklärt wird. Eine schrittweise Anwendung der Selektionskriterien zu Referenzkatalogen gibt Aufschluss über die Effekte einzelner Kriterien und bietet weitere Verifikation. Für die kinematische Analyse ist der Zielkatalog auf die Teilmenge beschränkt, für die Radialgeschwindigkeiten zur Verfügung stehen. Nach der Berechnung der nötigen kinematischen Parameter wird eine kinematische Halo-Selektion durchgeführt und anschließend basierend auf der resultierenden Sternanzahlverteilung verifiziert, wobei ein im Vergleich zu typischen Literaturwerten leicht erhöhtes Gefälle von  $\alpha \approx 4.4$  für ein einfaches Potenzgesetz gefunden wird. Der kinematische Raum wird nach Substrukturen abgesucht, wobei Gaia-Enceladus rein visuell identifizierbar ist, während der hierarchische Clustering-Algorithmus HDBSCAN weitere Substrukturen identifiziert, die Sequoia, Thamnos sowie den Helmi-Strömen entsprechen. Schließlich werden Elementhäufigkeitsmuster von Neutroneneinfangelementen mithilfe von GALAH-Daten untersucht. Indem deren Ähnlichkeit zu Referenzmustern abgeschätzt wird, werden insbesondere 319, 253, und 283 Kandidaten für Anreicherung durch den r-Prozess, s-Prozess bzw. den schwachen r-Prozess identifiziert.

# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
1.1	Nucleosynthesis and Chemical Enrichment	6
1.2	Substructure in the Stellar Halo	12
1.3	Wide-angle Spectroscopic Surveys	16
1.4	Aims	18
<b>2</b>	<b>4MOST</b>	<b>19</b>
2.1	The 4MOST instrument	19
2.1.1	Telescope and general design	19
2.1.2	The 4MOST subsystems	20
2.1.3	The low- and high-resolution spectrographs	23
2.2	4MOST survey programme	25
2.2.1	Galactic Surveys	25
2.2.2	Extragalactic Surveys	31
<b>3</b>	<b>Selection of input catalogue</b>	<b>34</b>
3.1	The Gaia mission	34
3.2	The Pristine survey	38
3.3	Selection criteria	42
<b>4</b>	<b>Verification of the 4MOST S2 catalogue</b>	<b>44</b>
4.1	Colour-magnitude diagrams	44
4.2	Verification of Pristine-calibrated metallicities	46
4.3	Application of selection criteria to GALAH	56
4.4	Spatial distribution: target catalogue vs. Besançon Galaxy model	60
<b>5</b>	<b>Chemical and kinematic analysis</b>	<b>67</b>
5.1	The 6D-Gaia-data subset of 4MOST S2	67
5.2	Kinematic parameter determination and selection of halo	68
5.3	Verification of the halo selection	74
5.4	The search for kinematic substructure in the halo	80
5.5	Identifying candidate n-capture process enriched halo stars	88
<b>6</b>	<b>Results and conclusions</b>	<b>94</b>
<b>7</b>	<b>Outlook</b>	<b>97</b>
<b>A</b>	<b>Appendix</b>	<b>105</b>
A.1	Reproduction of previous data	105
A.2	On the GSP-Spec abundances	108
A.3	Extension of Section 5.5 to full GALAH DR3	109

## List of Figures

1	Nuclear chart with nucleosynthesis processes	9
2	Ba and Eu abundances as function of metallicity	12
3	Discovery of Gaia-Enceladus	14
4	Substructures within the local stellar halo	15
5	Physical layout of the 4MOST instrument	20
6	Hexagonal array of fibres in focal surface	22
7	Optical layout of the 4MOST spectrographs	23
8	Spectral resolution of the LRS	24
9	Spectral resolution of the HRS	25
10	Synthetic spectrum of dwarf and giant for S2	28
11	Flux map of the Milky Way	34
12	Gaia's focal plane assembly	35
13	Gaia EDR3 passbands	36
14	Gaia EDR3 number of parameters solved as function of $G$	36
15	Median errors of Gaia DR3 astrometry	37
16	Ca H&K lines for different stellar parameters	39
17	SDSS-Pristine colour-colour space	40
18	Gaia-Pristine colour-colour space	42
19	Photometric vs. parallax distances; parallax zero-point offsets	44
20	CMDs for the full 4MOST S2 target catalogue	45
21	Histogram of absolute magnitudes for three sub-surveys	46
22	Target catalogue size for decreasing metallicity cuts	47
23	Coordinate separation for crossmatches	48
24	Metallicity comparison to GALAH	49
25	Metallicity comparison to SAGA	51
26	Metallicity comparison to data-driven metallicities	52
27	Metallicity comparison to data-driven metallicities, RGB sample	53
28	Metallicity comparison to APOGEE DR17	54
29	Impact of carbon on Pristine metallicities	55
30	Galactic coordinate distribution in GALAH	57
31	$\alpha$ vs. iron abundance diagrams for GALAH and crossmatch	58
32	C, Eu, and Ba abundances for GALAH and crossmatch	59
33	Spatial distribution: 4MOST S2 vs. Besançon model	62
34	Spatial distribution: 4MOST S2 vs. Besançon model for $[\text{Fe}/\text{H}] < -1.0$	63
35	$z$ -distribution for $7 \text{ kpc} < R < 9 \text{ kpc}$	64
36	Fraction of star counts between Besançon model and 4MOST S2	65
37	Histograms of radial velocities and distribution in $G$	67
38	Radial velocity uncertainties	68
39	Radial velocity comparison to reference catalogues	69
40	Comparison of different potentials	71
41	Spatial distribution of 6D subset and halo selections	73
42	Heliocentric distances of 6D subset and halo selections	73
43	Space velocities of 6D subset and halo selections	74
44	Stellar number distribution in disk and halo	75
45	Power-law fits to the galactocentric distance distribution	77
46	CMDs of 6D subset and halo selections	78
47	Metallicity distribution function for different velocity cuts	79
48	$v_{\perp}$ vs. $v_y$ for 6D subset and three halo selections	81
49	$L_{\perp}$ vs. $L_z$ for 6D subset and three halo selections	81
50	$E_{\text{orb}}$ vs. $L_z$ for 6D subset and three halo selections	82

51	Smoothed and background-subtracted $E_{\text{orb}} - L_z$ -space . . . . .	83
52	Substructures identified by HDBSCAN . . . . .	84
53	$L_{\perp}$ vs. $L_z$ distribution of selected substructures . . . . .	84
54	$v_{\perp}$ vs. $v_y$ distribution of selected substructures . . . . .	84
55	Distribution in kinematic space for strict selection . . . . .	86
56	$E_{\text{orb}}$ vs. $L_z$ in different metallicity cuts . . . . .	86
57	$\alpha$ -abundance-distribution for prograde and retrograde stars . . . . .	87
58	Mean $\sigma$ -deviations between GALAH and reference pattern . . . . .	89
59	2D plot of deviations from different elemental patterns . . . . .	90
60	Statistical properties of the selected n-capture process candidates . . . . .	91
61	Example of direct comparison between elemental patterns . . . . .	92
62	Reproduction of Helmi et al. (2018) velocity plot . . . . .	106
63	Reproduction of Helmi et al. (2018) $L_z - E_{\text{orb}}$ -plot . . . . .	106
64	Reproduction of Lövdal et al. (2022) kinematic parameters . . . . .	107
65	Direct comparison of kinematic parameters to Lövdal et al. (2022) . . . . .	108
66	GSP-Spec vs. GALAH metallicity and $\alpha$ -abundance . . . . .	109
67	Mean $\sigma$ -deviations from references for full GALAH DR3 . . . . .	110
68	2D plot of deviations from references for full GALAH DR3 . . . . .	110
69	Statistical properties of candidates for full GALAH DR3 . . . . .	111

## List of Tables

1	4MOST instrument specifications . . . . .	21
2	4MOST Consortium and Community Surveys . . . . .	26
3	Target selection criteria of the 4MOST S2 survey . . . . .	42
4	Halo number density fits and literature references . . . . .	76
5	Properties of identified substructures in halo . . . . .	85

# 1 Introduction

The history of the early Universe and especially the nature of the first stars to form are topics to which a lot of attention has been dedicated, both theoretically and observationally, because an understanding of these topics is fundamental in building up a complete picture of our Universe and putting its present-day state into context. In particular, observations play the key role of constraining, supporting, confirming, or even refuting theoretical models and simulations (see, e.g., [Frebel & Norris 2013](#); [Bromm & Larson 2004](#); [Klessen & Glover 2023](#)). Thus, it is desirable to determine the ages of stars in order to identify the oldest ones. Methods to determining the age of a star include asteroseismology, isochrone fitting, and nucleochronometry. Asteroseismology makes use of oscillations within a star to derive the sound speed inside the star, which is related to the composition of the star, and the composition is an indicator of the star's evolutionary status. For stars with precise measurements of stellar parameters such as the effective temperature, surface gravity, and metallicity, the age can be estimated by fitting this data to isochrone models in a Kiel diagram. Nucleochronometry entails the determination of stellar ages using the decay of long-lived isotopes, in particular Th and U, but this method is only viable for r-process-enhanced stars where Th or U has been detected ([Soderblom 2010](#)).

In most cases, however, the determination of the age of a star is neither straightforward nor easy, so the amount of heavy elements found in the star, referred to as its metallicity, is often utilised as a proxy. As the Universe has been enriched in metals by each consecutive generation of stars, more metal-deficient stars are typically older compared to more metal-rich stars (see, e.g., [Frebel & Norris 2015](#)). Specifically the first generation of stars not yet enriched with processed stellar material that started forming at redshifts of around  $z \sim 30$  (commonly referred to as Population III stars) are of prime interest for the above-mentioned topics, but as of today they remain theoretical as not a single completely metal-free star has been identified thus far ([Klessen & Glover 2023](#)).

Accordingly, throughout the past decades, the importance of observing metal-poor stars in regards to a large number of astronomical and astrophysical problems – including but not limited to early metal- and gas-mixing processes, galaxy formation, nucleosynthesis, primordial lithium abundances, the metallicity distribution function, and the first mass function ([Frebel & Norris 2015](#); [Beers & Christlieb 2005](#)) – has become increasingly clear, and since [Beers et al. \(1985\)](#) at the latest, systematic searches for the stars with the lowest metallicities have been carried out. However, due to their rarity, the process of finding those stars can be rather difficult. As [Beers & Christlieb \(2005\)](#) describe, the typical procedure is to apply selection criteria to a catalogue from a wide-angle survey to pick candidates, observe these at intermediate resolution in order to identify the stars which are actually metal-poor, and then finally follow this up with high-resolution observations for a full analysis of promising stars.

While large-scale all-sky surveys contribute greatly to our understanding of our Galaxy's structure, it is in many cases desirable or even necessary to have spectroscopic observations that go beyond the surface level and provide high-resolution spectra for a large number of objects down to faint magnitudes. For instance, while the *Gaia* mission ([Gaia Collaboration et al. 2016](#)) measures the astrometric position, parallax, and proper motion for an unprecedented number of more than a billion stars, the spectroscopic instrument ([Cropper et al. 2018](#)) is restricted to bright objects only. This means that the number of objects within the most recent data release Gaia DR3 ([Gaia Collaboration et al. 2023c](#)) for which spectroscopic data is available is several magnitudes lower compared to the astrometric data.

Similarly, other projects such as Euclid ([Euclid Collaboration et al. 2022](#)), PLATO ([Magrin et al. 2018](#)), eROSITA ([Merloni et al. 2012](#)), as well as numerous non-European projects and surveys would benefit tremendously from high-quality spectroscopic observations, and the combination of these data will enable in-depth studies which would not be possible individually. An example of this interplay between these different projects is the discovery of Gaia-Enceladus ([Belokurov et al. 2018](#); [Helmi et al. 2018](#)). The discovery of Gaia-Enceladus, which is likely the Milky Way's last big merger around 10 Gyr ago,

and the fact that the remains of it appear to be making up a significant fraction of the halo near our Solar system, was facilitated by Gaia’s data release 2 (Gaia Collaboration et al. 2018b) together with numerous spectroscopic surveys including GALAH (De Silva et al. 2015), APOGEE (Majewski et al. 2017; York et al. 2000), RAVE (Kunder et al. 2017) and LAMOST (Deng et al. 2012; Zhao et al. 2012; Cui et al. 2012). This important finding also showcases the significant advances made in the field investigating streams and substructures and the progress in understanding the early history of the Milky Way.

Typically, the Milky Way Galaxy is considered to be made up of several components that are distinct in their kinematics, ages, and chemistry, namely the thin and thick disks, the bulge including the bar, warm and cold gas, as well as the stellar and dark matter halos. The thin disk is characterised by ongoing star formation and a stellar density greater than the thick disk, and its stars rotate in a rather orderly fashion with a small velocity dispersion. As opposed to this, the thick disk has a higher velocity dispersion, although the rotation remains ordered (see, e.g., Bensby 2014; Sanders & Das 2018). Its stars are generally more metal-poor and older, and its chemical properties differ from the thin disk as a result of a comparably short but violent star formation history (see, e.g., Haywood et al. 2015; Schneider 2006). The centre of the Galaxy, where both the bulge and the bar are located, is the densest region overall and within the components individually, but due to dust leading to a substantial extinction in the visible, it is not easy to investigate its exact nature. While neither dominant in regards to the light emitted from our Galaxy nor its mass (Shen et al. 2010; Licquia & Newman 2015), it displays a variation of different populations with different properties (Carroll & Ostlie 2017).

The stellar halo reaches far out beyond the other components, but thins out significantly towards larger distances from the Galactic centre and makes up only a few percent of the Milky Way’s total visible mass (Deason et al. 2019). The stellar halo features a lot of the Galaxy’s most metal-poor as well as oldest stars, and is dominated by random motion (see, e.g., Bland-Hawthorn & Gerhard 2016). All of the visible components described above are situated within a dark matter halo which hosts the majority of the Galaxy’s mass. Because of the distinctness of the components, they likely had their own individual formation processes, but presumably these processes were not always independent of each other. In particular, large-scale mergers can be responsible for the formation of a bulge, a stellar halo, and for contributing to the disk (Barnes 1992; Zolotov et al. 2009; Quinn et al. 1993), and since the same event is experienced by the whole Galaxy, there may be common stellar material within different components (Helmi 2020). In order to probe into the past and uncover the early history of the Milky Way, the detailed analysis of stars is necessary, as they serve as witnesses of their origin and environment. This is where the field of *Galactic archaeology* comes in, which focuses on old, long-lived stars towards that end.

## 1.1 Nucleosynthesis and Chemical Enrichment

Since a star’s lifetime is inversely proportional to its mass, the earliest stars with low masses should still be observable today and depending on their evolutionary status, its properties can give insight into the conditions characterising their natal clouds within the interstellar medium (ISM). In particular, this means that stars that formed from the same natal cloud and experienced a similar life should have patterns of elemental abundances that are alike (Frebel & Norris 2015; Freeman & Bland-Hawthorn 2002). Conversely, distinct environments lead to the production of distinct elements, and the timescales on which this happens can vary dramatically, with abundances of particular elements being characteristic of particular events.

One prominent example is the typical appearance of the plot of the  $\alpha$ -element abundance  $[\alpha/\text{Fe}]$  versus the iron abundance  $[\text{Fe}/\text{H}]$ , which shows a constant level of overabundance of  $\alpha$ -elements at lower iron abundances up to around  $[\text{Fe}/\text{H}] = -1.0$ , from where on  $[\alpha/\text{Fe}]$  drops continuously (e.g. Lee et al. 2011; Bensby et al. 2007). This is the result of the different timescales associated with the primary production mechanisms of these elements: While a type II supernova (which is in most cases the core collapse and subsequent explosion of a massive star) occurs comparably soon following a star’s formation due to

the lifetimes of massive stars being short, a type Ia supernova (in which a white dwarf is pushed to the Chandrasekhar limit by absorbing mass from a companion star) takes much longer because the mass of the star is much smaller. As  $\alpha$ -elements are liberated in large quantities during type II supernovae, while type Ia supernovae produce a lot of iron, there is an overabundance in  $\alpha$ -elements until the onset of type Ia supernovae, from which point onwards  $[\alpha/\text{Fe}]$  starts declining as more and more type Ia supernovae occur (McWilliam 1997; Nissen & Schuster 2010).

Looking at the binding energy per nucleon, it is growing with the number of nucleons up until iron, where it reaches its peak and then slowly declines. Accordingly, stars create elements only up to the iron-peak elements through nuclear burning. As for heavier elements, they are produced by means of neutron capture. The two general types of neutron capture in stars are the slow neutron-capture process and the rapid neutron-capture process (s-process and r-process, respectively), which happen depending on whether the capture of neutrons proceeds slower or quicker than the  $\beta$ -decay of the neutron into a proton, respectively. A typical s-process site is the envelope of an asymptotic giant branch (AGB) star, and especially those with low masses are significant participants in shaping the Galaxy's chemical properties (Käppeler et al. 2011). R-process sites tend to be more violent, as the r-process requires a rapid flux of neutrons to work. This is, for instance, the case in particular kinds of type II supernova, or in the mergers of a neutron star with another neutron star or a black hole, but other scenarios are possible as well (Cowan et al. 2021). Furthermore, it seems that stars within dwarf galaxies of our Local Group show certain trends such as mostly being separated by mass within both the  $[\alpha/\text{Fe}]$  vs.  $[\text{Fe}/\text{H}]$  diagram and a similar diagram using neutron-capture elements instead of  $\alpha$ -elements (Tolstoy et al. 2009). The trend shown in the former diagram hints that the dwarf galaxies generally have lower  $[\alpha/\text{Fe}]$  abundances compared to the Milky Way, or at least that the  $\alpha$  knee occurs earlier for galaxies with lower masses, which might be because of inefficient star formation, or because of star formation being maintained over a shorter period of time (Helmi et al. 2018). In any case, these examples illustrate how studying certain elements of stars in detail yields information on its origin and history.

The r-process in particular appears to be enhanced in many halo stars, notably members of substructures like Gaia-Enceladus and Sequoia (Matsuno et al. 2021; Aguado et al. 2021; Roederer et al. 2018). This merits a closer inspection of the r-process regarding the specific nuclear physics behind it, the conditions necessary to achieve it, and especially which astrophysical sites may provide such conditions, all of which has been laid out in detail in the review by Cowan et al. (2021). Both fusion and photodisintegration happen at considerable rates in the case of high densities and temperatures, and when forward and backward nuclear reactions (including proton and neutron-captures) reach an equilibrium for all reactions, nuclear statistical equilibrium is achieved (Hix & Thielemann 1999).

However, the requirements for this are not always sufficiently present, which can lead to the formation of quasiequilibrium groups in limited areas of the nuclear chart along isotopic chains and the charged-particle reactions are frozen. At high neutron densities, the capture of neutrons occurs much more rapidly than  $\beta$ -decay and nuclei with very low neutron-separation energies are created, all the while photodisintegration keeps going depending on the temperature, and since both of these reactions are also quicker than astrophysical processes, neutron captures and photodisintegration can enter a chemical equilibrium. Deriving the abundance ratios of two neighbouring isotopes from chemical potentials, Clayton (1968) shows that in all isotopic chains there will be an abundance maximum which only depends on the neutron number density as well as the temperature. The path of the r-process in the nuclear chart is then established by the fact that the neutron-separation energy at the maximum is similar in all chains. At the magic neutron numbers  $N = 50$ ,  $N = 82$ , and  $N = 126$  the stability peaks as the shells, within the frame of the atomic shell model, are filled there. This leads to longer  $\beta$ -decays and thus abundance maxima which can be seen at the matching mass number  $A$  after the isotopes reach stability through the  $\beta$ -decay.

Thus, the process can be divided into two stages, the first of which is dominated by neutron captures, whereas in the second one both neutron captures and the  $\beta$ -decays occur at comparable rates during the so-called r-process freeze-out. The second stage can commence once the ratio between free neutrons and nuclei (neutron-seed-ratio) approaches one and the r-process abundances are smoothed, which is something that distinguishes the r-process from the s-process where the  $\beta$ -decays and neutron captures are basically never on the same timescale. Once the nucleon number becomes very high, spontaneous fission can occur and create lighter fission products which can participate in the neutron capture process again to grow, and this can repeat several times over before the r-process terminates (Cowan et al. 2021). As for the required conditions to reach the nuclear statistical equilibrium, it appears that high enough neutron-seed-ratios are only feasible for a very neutron-rich environment at low entropy, or for a moderately neutron-rich environment at very high entropy (Farouqi et al. 2010).

One important aspect of the high temperature environments is that nucleons become free and cooling occurs primarily via neutrinos and antineutrinos, which can lead to matter becoming neutron-rich if it is exposed to them sufficiently long. However, the disparity between neutrino emissions from merging neutron stars and emissions that result from forming neutron stars in core-collapse supernovae should be considered: While in the latter case, a large difference between the average energies of electron antineutrinos and neutrinos is required, which has not been reproduced by any recent neutrino-wind simulations (Martínez-Pinedo et al. 2014; Fischer et al. 2020), the required difference is much smaller in the former case (Foucart et al. 2016). Since for the r-process the presence of a large amount of neutrons is typically necessary, an obvious candidate for supplying neutrons are neutron stars, but in order to enrich its surroundings, the produced material needs to be able to escape the neutron stars' strong gravitational pull, which consequently also necessitates rather violent conditions. Thus, the sites that are most obvious as candidates for the r-process are core-collapse supernovae of massive stars, in particular those with fast rotation and strong magnetic fields, as well as neutron star-black hole mergers or neutron star-neutron star mergers. The observation of the gravitational wave event GW170817 in 2017 (Abbott et al. 2017b), which has been identified as a binary neutron star merger (Abbott et al. 2017c,a), has shifted the focus onto the latter, as a reanalysis of the spectra led to the direct discovery of r-process elements in neutron star mergers (Watson et al. 2019).

In the process of a binary neutron star merger or a neutron star-black hole merger, several kinds of (mass) ejections occur, including dynamic ejecta, neutrino winds, as well as accretion disk outflows. The dynamic ejecta contain both cold tidal ejecta and shock-heated ejecta from the areas of direct contact, although only the former are present in neutron star-black hole mergers, and taking into account weak processes, they can supply light and heavy r-process nuclei (Wanajo et al. 2014; Martin et al. 2018). Beyond the dynamic ejecta arising during the merging process, there will also be ejections following the merger, such as a neutrino wind, which results primarily in the production of r process elements below  $A = 130$  (Martin et al. 2015). Finally, the accretion disk that forms from the merger generates outflows, which can include significant portions of the disk itself, and simulations show that the main process driving the ejections is viscous heating (Just et al. 2015), and that basically all r-process nuclei can be created from disk outflows (Wu et al. 2017). Simulations in the field of neutron star mergers and associated nucleosynthesis have come a long way, with Newtonian approaches to the modeling being superseded by fully relativistic approaches (see, e.g., Baiotti & Rezzolla 2017; Papenfort et al. 2018; Bovard et al. 2017) and additional physical phenomena such as magnetic fields, also in relation with neutrinos, being taken into account (see, e.g., Giacomazzo et al. 2015; Guilet et al. 2017; Giacomazzo et al. 2009). Clearly, further multi-messenger campaigns will help shed further light on the role of neutron star mergers in nucleosynthesis processes and consequently the chemical enrichment of our Universe by providing direct observational constraints.

However, as Skuladottir et al. (2019) point out, it is unlikely that neutron star mergers are the only source of r-process enrichment. Coming back to supernovae-related candidate sites, a number of scenar-

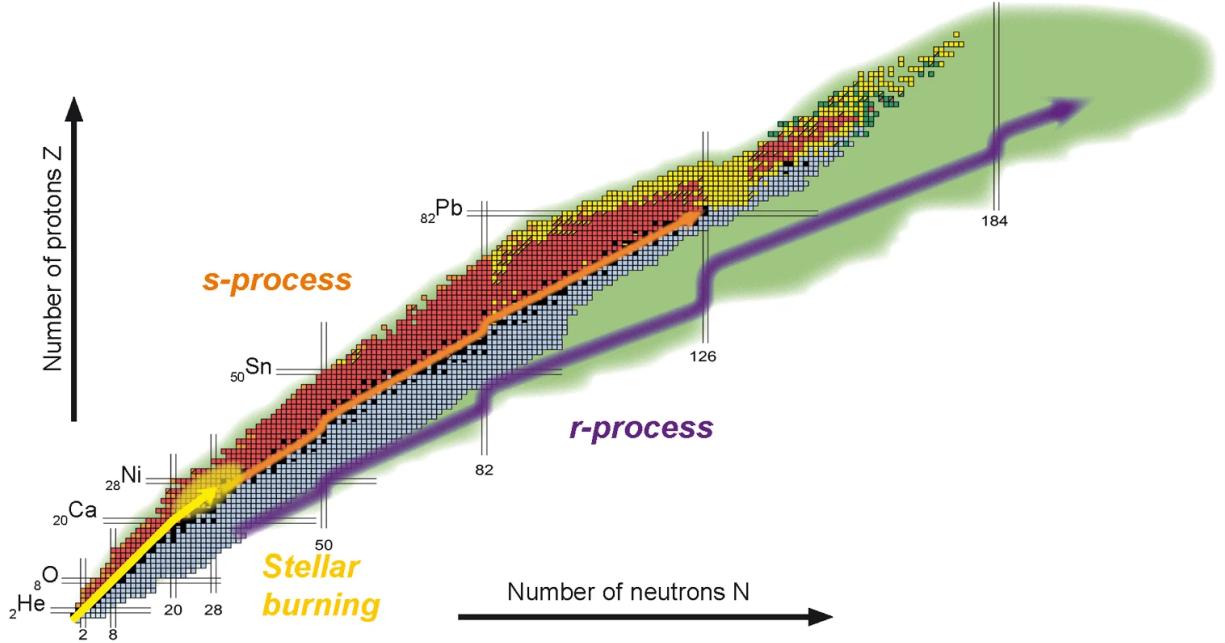


Figure 1: Table of nuclides with the path of the major nucleosynthesis processes marked (Thielemann 2023).

ios that have historically been seen as a possible source of r-process elements – including neutrino winds from core-collapse supernovae, electron-capture supernovae, neutrino-induced r-process in the helium shell, as well as quark deconfinement supernovae – have been ruled out in recent years, or have been identified to only allow the production of lighter r-process elements. On the other hand, magnetorotational supernovae as well as hypernovae seem better suited to the production of r-process elements, but are rather rare occurrences (Cowan et al. 2021). Besides the main r-process, there is most likely another component to the r-process, commonly labeled “weak r-process”, thought to be responsible for the production of light neutron-capture elements (see, e.g., Wanajo & Ishimaru 2006; Truran et al. 2002). This has been discovered and studied by means of detailed chemical analyses, which have shown that some stars have a particular common chemical pattern that shows similarity to the solar system’s r-process enrichment for heavy neutron-capture elements, but deviates in their light neutron-capture elemental abundances (Sneden et al. 2000). Fujibayashi et al. (2023, 2020) study the mass ejections in neutron star mergers and find that the scatter in the light-to-heavy element ratios is dependent on both the binary mass ratio and the total mass, with low-mass binary stars producing abundance patterns in line with the weak r-process abundance pattern. The star HD 122563 in particular is an important probe into the weak r-process (Honda et al. 2006), whereas a typical (strong) r-process star is, for example, CS 22892-052 (Sneden et al. 2003).

As opposed to the r-process, the s-process moves along the stability valley within the table of nuclides (see Figure 1), which is the natural result of its distinguishing feature of slow neutron-capture when compared to the  $\beta$ -decay’s timescale. In order to fully comprehend and trace the s-process and the resulting abundances, an understanding of the nuclear physics behind it is necessary, especially concerning the neutron capture reactions as well as the  $\beta$ -decay rates, but also relating to charged particle reactions influencing energy and neutron production which provide the conditions for the s-process to occur in the first place (Angulo et al. 1999; Heil et al. 2008). To investigate neutron capture rates, a principal parameter of interest are the neutron capture cross sections, and thus it is essential to perform measurements on them in laboratory conditions. Neutron sources for this cause include shooting electron beams at heavy metal targets in accelerators, or causing spallation reactions with highly energetic particle beams (for instance at the CERN’s n\_TOF Facility (n\_TOF Collaboration et al. 2003)). The produced neutrons comprise a range of energies, which is important in order to cover any stellar temperature, and cross sections can be obtained using their time-of-flights. An alternative method of measurement is the activation tech-

nique, although it can only be used if unstable nuclei are generated by neutron capture. Käppeler et al. (2011) outline these methods among others in their detailed review on the s-process. As they also point out, due to the limitations in the experimental approach, theoretical calculations incorporating statistical models must be carried out so that experiment and theory will complement each other and provide a more comprehensive picture. Calculations on s-process  $\beta$ -decay rates were given by Takahashi & Yokoi (1987), and experimentally, the upcoming PANDORA project of the INFN will study decays in plasma in order to better recreate the conditions within the stellar environment (Mascali et al. 2022).

In our current understanding, there are actually two components making up the s-process: the weak s-process, which is responsible for most of the isotope production from Fe to Sr and occurs in massive stars as the convective He core burning as well as the convective C shell burning takes place, and the main s-process, which produces all isotopes from  $A = 90$  onwards up to Pb in low-to-intermediate mass thermal-pulse asymptotic giant branch (TP-AGB) stars (Käppeler et al. 2011). Observations that aim to study s-process abundances typically focus on AGB stars, post-AGB stars, and their planetary nebulae, or on the companion of an AGB star in a binary system that is accreting material from it. In the former case, recurring thermal pulses and so-called third-dredge-up events (during which material, especially carbon, from the inner parts of a star is transported to the surface) provide the heavy s-process abundance at the surface of AGB stars, but due to their rather cool surface temperatures, the spectral determination of s-process elements can be difficult because of a large number of molecular lines. Post-AGB stars present similar difficulties, but analyses have been possible, and neutron-capture element abundances in planetary nebulae have also been measured successfully (Sterling & Dinerstein 2008; Sharpee et al. 2007).

In the case of a binary system where s-process-enriched material is accreted by the companion star of an AGB star, and the AGB star subsequently becomes a white dwarf, the measurements of the corresponding s-process isotope lines are easier as the companion star remains warm and thus showcases less molecular absorption lines. The three magic neutron numbers at 50, 82, and 126 correlate to the three peaks featured in the abundance pattern of the s-process, namely  $^{88}\text{Sr}$ ,  $^{138}\text{Ba}$ , and  $^{208}\text{Pb}$ , with Sr and Ba in particular displaying strong lines in the optical, which allows detection even for very metal-poor stars (Käppeler et al. 2011). Notably, some carbon-enhanced metal-poor (CEMP) stars can also showcase features that indicate r-process enhancement in addition to s-process elements, while others show only one of the two, and some neither, which prompted Beers & Christlieb (2005) to define subclasses of CEMP stars including CEMP-r, CEMP-s, CEMP-r/s, and CEMP-no stars, concerning which Bisterzo et al. (2011) have given detailed interpretations based on calculations with AGB star models (Bisterzo et al. 2010).

Stars at the very low-metallicity end, which typically correspond to very old stars, exhibit abundance patterns that are chiefly in line with r-process enrichment. This shows that the onset of the r-process was early in the history of the Universe and is a further indication for the sites of the r-process being short-lived and evolving on small timescales, as well as being related to massive objects, as the elements were already produced and ejected into space before the halo stars containing them were formed (Sneeden et al. 2008). Shen et al. (2015) employ the high-resolution cosmological zoom-in simulation *Eris* (Guedes et al. 2011) in order to examine the r-process enrichment of our Galaxy (in particular early on) using [Eu/Fe]. They conclude that even at a small delay time of 100 Myr, the material synthesised in compact binary mergers can contribute to stars with very low metallicities at early times. Using the cosmological zoom-in simulation FIRE (Hopkins et al. 2014), van de Voort et al. (2015) obtain similar results.

Generally, material ejected into space is not mixed with the interstellar medium immediately, which leads to local inhomogeneities resulting from individual r-process events. These events are rare, specifically early in the evolution of a galaxy when the interstellar medium was not very well-mixed. Thus, there is a large scatter of about two orders of magnitude for r-process elements at low metallicities, as

stars forming close to where a strong r-process event took place will be significantly more enriched than others, whereas at a later stage of galactic evolution the r-process abundance should tend to approximate an average value as the Galaxy turns more chemically homogeneous (Thielemann et al. 2017; Sneden et al. 2008). One aspect of the r-process enrichment history that remains to be resolved is that r-process elements have been observed in stars with metallicities  $[Fe/H] \leq -3.0$ , but the supernovae that result in the neutron stars necessary for a merger yield a considerable quantity of  $[Fe/H]$  already, making such a low metallicity in the presence of r-process elements rather odd. So far, only significant turbulent mixing of the interstellar medium could replicate these observations within simulations (Thielemann et al. 2017).

The onset of the s-process is delayed compared to the r-process, as the low- and intermediate-mass stars that later develop into AGB stars producing elements via the s-process take longer to reach that evolutionary stage, while the r-process should enrich the Galaxy from earlier on as described in the previous paragraph, meaning early chemical evolution should be dominated by the r-process (Käppeler et al. 2011). As soon as AGB stars of low to intermediate mass emerge, the s-process can start. However, the timing of this onset is not well constrained as there is large scatter, which might be the result of the early Galaxy being unmixed or the stars being contaminated at high metallicity. Simmerer et al. (2004) report s-process activity already at  $[Fe/H] = -2.6$ , which might indicate that the range of masses permitting the s-process to occur is quite wide and includes more massive intermediate-mass stars as well. In any case, despite the large scatter even at later times, the general trend shows a growing ratio of s-process elements towards higher metallicity – although not necessarily monotonically – and the s-process synthesis seems to be occurring gradually and increasing continually, instead of having an abrupt onset. Sneden et al. (2008) also describe the increase of the s-process as the metallicity rises, as well as abundances concordant with the r-process being the only source of neutron-capture enrichment for many stars, and consider significant s-process enrichment to occur from  $[Fe/H] \gtrsim -2.0$  onwards. Utilising detailed studies of high resolution spectra, Battistini & Bensby (2016) find the s-process to become dominant over the r-process at  $[Fe/H] \sim -0.5$ .

An important sub-class of metal-poor stars are carbon-enhanced metal-poor (CEMP) stars, which sometimes exhibit s-process enhancement as well. These so-called CEMP-s stars are likely the result of mass transfer from a more massive binary companion that already experienced the third dredge-up, which transported s-process-enriched material from the inner He shell to the surface (Bisterzo et al. 2011). This companion star might have subsequently become an undetected white dwarf (Sneden et al. 2008). One further conclusion drawn from the neutron-capture element abundances relates to the thick disk of the Galaxy: Thick disk stars at low metallicity have a considerably higher  $[Eu/Fe]$  value, while also displaying a somewhat lower  $[Ba/Fe]$  abundance. This implies that the timescale on which the thick disk formed is shorter than that of supernovae type Ia, and that the r-process is the main source of heavy elements in thick disk stars. Since the s-process appears to play little to no role, it seems that the thick disk formed rapidly (Käppeler et al. 2011).

As most elements are synthesised by both the r- and s-process, it may sometimes be difficult to tell which processes are responsible for a star’s enrichment. Still, some elements seem to be better indicators than others, with Eu being formed almost exclusively by the r-process, while Ba is largely produced by the s-process. For instance, Bisterzo et al. (2014) employ a Galactic chemical evolution model in order to estimate the s-process contributions to heavy element abundances and find a percentage of only  $\sim 6\%$  for Eu, while nearly 90% of Ba appears to be supplied by the s-process. Thus, an important discriminator for estimating whether a star has been enriched primarily via the s-process or the r-process is the  $[Eu/Ba]$  (or  $[Ba/Eu]$ ) abundance ratio (Battistini & Bensby 2016; Beers & Christlieb 2005; Mashonkina & Gehren 2001). In particular,  $[Ba/Eu]$  generally increases as  $[Fe/H]$  increases due to the later onset of the s-process, as was discussed above. Figure 2 shows this trend, containing data from many different studies that have been carried out over the years and compiled here by Käppeler et al. (2011). The figure also shows the abundance ratios  $[Eu/Fe]$  and  $[Ba/Fe]$  as a function of  $[Fe/H]$ . The downward trend

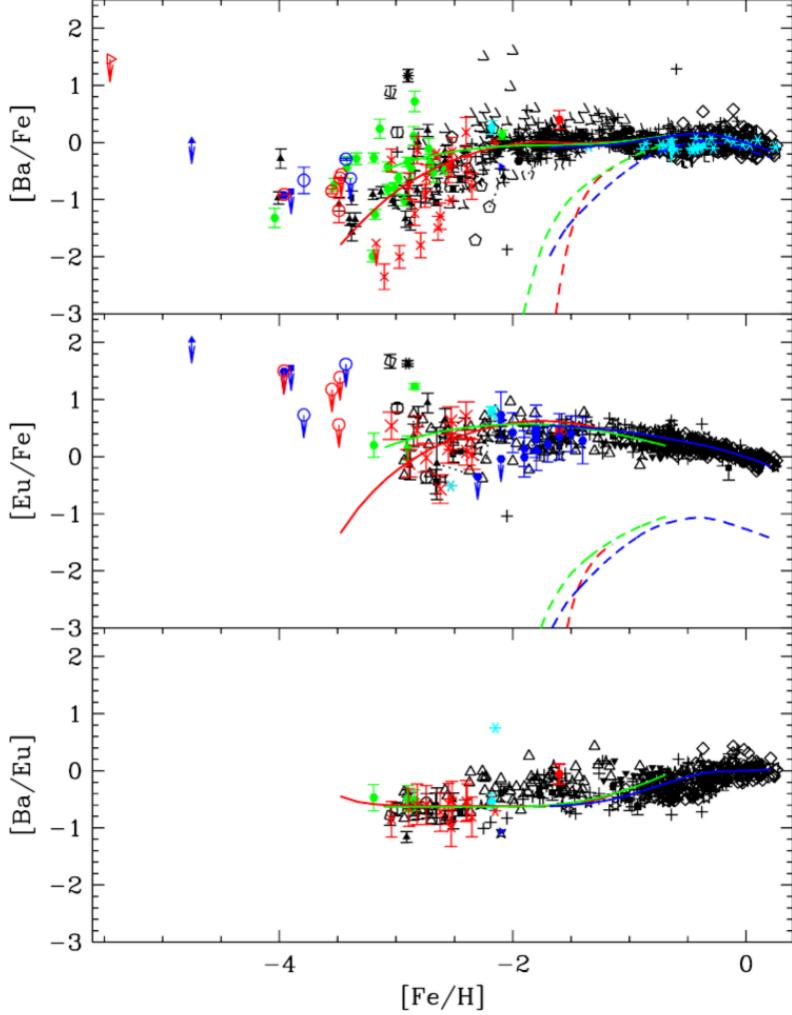


Figure 2: Evolution of  $[\text{Ba}/\text{Fe}]$ ,  $[\text{Eu}/\text{Fe}]$ , and  $[\text{Ba}/\text{Eu}]$  as a function of  $[\text{Fe}/\text{H}]$ . The dashed lines indicate model predictions based only on the s-process, while the solid lines show the predictions for the total s/r-abundances (Käppeler et al. 2011).

of  $[\text{Eu}/\text{Fe}]$  at higher metallicities can be explained by an increasing iron abundance in the interstellar medium due to the late onset of supernovae type Ia, and is thus not due to a change in efficiency of the r-process. Ba is mostly synthesised by the s-process at higher metallicities and consequently  $[\text{Ba}/\text{Fe}]$  does not exhibit a decrease there. The models suggest that in metal-poor stars, Ba stems from the r-process.

## 1.2 Substructure in the Stellar Halo

The stellar halo is particularly interesting in the study of the early Galactic history. One reason for this is that the primary drivers of galaxy evolution and accumulation of dynamical mass – which are mergers and accretion events – leave their debris mostly in the stellar halo, meaning these events can best be disentangled from kinematics and chemical abundance information. Moreover, disk stars heated by such events can also be displaced here (Tissera et al. 2013). Another reason why the stellar halo is such an important component to study in regards to the early Milky Way history is that many of the most metal-poor and oldest stars can be found in the halo. The existence of a relationship between a galaxy’s mass and its metallicity has been studied for a long time, starting with the observations by Lequeux et al. (1979), and the so-called mass-metallicity relation that was subsequently discovered shows that more massive galaxies have higher metallicities (e.g. Tremonti et al. 2004; Maiolino et al. 2008). Since the Milky Way in its early stages was the most massive object among its surrounding neighbours, this means that the lower-mass galaxies that were accreted had a lower metallicity, thus making the stellar halo of

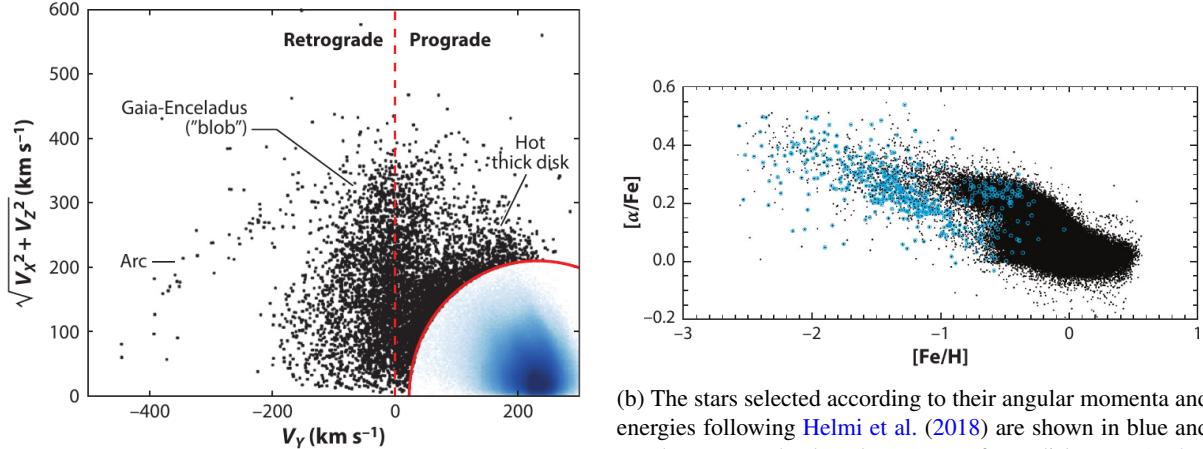
today's Milky Way so metal-poor by leaving their debris from the merger there. In addition to being metal-poor, the halo is also older than the thin disk as mergers occurred much more regularly at earlier times and the thin disk could only fully develop into its current form after this epoch of mergers ended (Helmi 2020). However, some of the oldest stars should also be located in the inner regions or bulge of the Milky Way (El-Badry et al. 2018). In any case, the distinction between accreted stars and stars that formed in-situ is of primary importance, as the study of the different contributions to the stellar halo, for instance accretion from dwarf galaxies or the capture of globular clusters, gives insight into the build-up and evolution of the Galaxy.

A recent major discovery is that of Gaia-Enceladus as the object of the Milky Way's last large-scale merger event, first found by Belokurov et al. (2018). This came about as a result of the high quality of Gaia DR2 data, through which it became apparent that most halo stars nearby are part of one big kinematic structure which is overall slightly retrograde (Koppelman et al. 2018; Myeong et al. 2018b). It was named and characterised in greater detail by Helmi et al. (2018), who combined kinematic and chemical information to identify it within the inner stellar halo and thick disk (see Figure 3a). A comparison to a simulation shows that most stars within the above-mentioned structure could have an extragalactic origin and entered the Milky Way through a merger event. Using chemical abundances from APOGEE, the stars within the structure are for the most part separate from the thick and thin disk stars in a  $[\alpha/\text{Fe}]$  vs.  $[\text{Fe}/\text{H}]$  plot (see Figure 3b). Them having metallicities distributed over a large range and also showing a lower alpha-element abundance at  $[\text{Fe}/\text{H}] \sim -0.6$  than the thick disk means that the stars were not born in a single burst and that the system they formed in has a star formation rate that is lower compared to the thick disk. Arguing based on isochrone models and abundance patterns, the authors find that the merger with this structure, dubbed Gaia-Enceladus, occurred about 10 Gyr ago, and rough estimates yield a mass-ratio of  $\sim 0.24$  for this event, meaning there must have been considerable heating and growth of the thick disk. Gallart et al. (2019) find a similar mass ratio and time.

The age of the low- $\alpha$  stars in Gaia-Enceladus, which contain the most metal and should have formed before the merger event occurred, indicates when the merger happened, and since they have been found to be younger than low- $\alpha$  stars in the thick disk, it appears that the Milky Way already had a disk before the merger. It seems that the merger of the Milky Way with Gaia-Enceladus was the Galaxy's last merger event, which was likely followed by more quiescent thin disk growth (Helmi 2020). The merger with the Sagittarius dwarf galaxy around 8 Gyr ago (e.g., Dierickx & Loeb 2017) and the current one with the Large Magellanic Cloud both have a much smaller mass ratio and thus have a much smaller effect on the Milky Way, although they do result in so-called phase-space spirals as well as waves (Antoja et al. 2018; Laporte et al. 2019).

Typically, kinematic studies concentrate on integral-of-motion (IoM) spaces (for example angular momentum or energy) as objects that were accreted from the same origin stay close together in this space even if they are spatially indistinguishable from other objects. Besides Gaia-Enceladus, a number of other kinematic substructures within the stellar halo that are associated with merger events with dwarf galaxies have been found, among them Sequoia (Myeong et al. 2019), which is even more metal-poor and more retrograde compared to Gaia-Enceladus, and Thamnos (Koppelman et al. 2019a), which is also slightly more retrograde and at lower energies than Gaia-Enceladus. Aside from these, there are also the Helmi streams (Helmi et al. 1999), which despite being accreted by the Milky Way about 5 to 8 Gyr ago (Koppelman et al. 2019b) have already reached the inner halo, although this timescale is still not entirely certain.

In a recent effort, Ruiz-Lara et al. (2022) have set out to characterise the independent substructures that can be found in the stellar halo of the Solar neighbourhood. In their paper, the authors demonstrate how most of these are actually part of 12 extended substructures, with only a few independent ones remaining. In the first paper of their series (Lövdal et al. 2022a), the authors implemented a data-driven



(a) The blob that was named Gaia-Enceladus can be seen to be kinematically distinct from the hot thick disk. The blurred area in the bottom right corner are regular disk stars (Helmi 2020; Koppelman et al. 2018).

Figure 3

hierarchical clustering algorithm and found 67 nearby clusters within the stellar halo. The details of the single linkage algorithm used can be found in Lövdal et al. (2022a), but broadly speaking it makes use of the energy, angular momentum  $L_z$ , and the angular momentum  $L_{\perp}$  perpendicular to that to form cluster candidates by a step-wise connection of data points close to each other in IoM space that have not been connected previously. If the region around these candidates is of greater density in contrast to an artificially created halo, the algorithm considers it a cluster. To identify which clusters can be grouped together, the authors first employ a method of isochrone fitting to estimate the age and metallicity of clusters and then go through a number of steps which take into account the IoM distances between groups of stars, their metallicity distributions, and their colour distributions, resulting in the 67 clusters being divided into 12 extended substructures. Figure 4 shows these substructures in  $E-L_z$ -space, where they have been identified with Gaia-Enceladus, Sequoia, Thamnos, the Helmi streams, as well as a few more. Besides this division, Ruiz-Lara et al. (2022) find that not only are metal-poor stars generally on more retrograde orbits, but also that clusters on retrograde orbits are typically older compared to prograde clusters. Furthermore, the different substructures that were found exhibit distinct metallicity distribution functions, although some aspects that resemble each other hint at the size of the origin systems being similar.

The existence of the thick disk as a feature of the Milky Way is something that has been debated ever since it had first been identified by Gilmore & Reid (1983), with some authors finding that there is no distinct thick disk (Bovy et al. 2012) and different authors deriving differing properties depending on the method used (Kawata & Chiappini 2016). On the other hand, the thin and thick disk have been shown to form chemically separate sequences, for instance in a  $[\alpha/\text{Fe}]$  vs.  $[\text{Fe}/\text{H}]$  diagram (e.g., Hayden et al. 2015). In any case, the thick disk merits closer inspection, as it is sort of the transition population between the halo and the thin disk, with the tail of the thick disk, called the metal-weak or hot thick disk, being particularly interesting. The stellar halo appears to be dual in nature, as can be seen from the bimodality in Figure 22 in Gaia Collaboration et al. (2018a), and it appears that while the more metal-poor part can be explained by Gaia-Enceladus, the other sequence is made up of stars that belong to the tail of the thick disk (Helmi 2020). In fact, some authors suggest that the Milky Way's halo is entirely accreted and any in-situ halo stars originate from the heated thick disk following heating from Galaxy interactions such as accretion events. Di Matteo et al. (2019) find that in the Solar neighbourhood, metal-rich thick disk stars make up around half of their kinematically defined halo stars and that the stellar halo in the inner parts of the Galaxy is dominated by them.

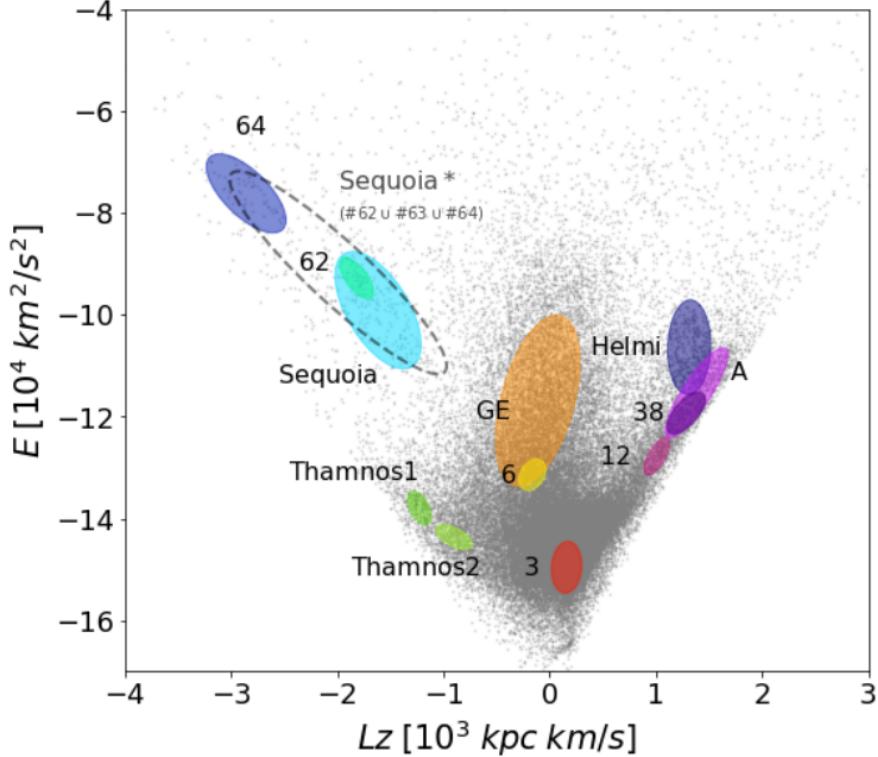


Figure 4: The 12 independent extended substructures including approximate boundaries (Ruiz-Lara et al. 2022).

A number of different formation scenarios for the thick disk exist, for example that a minor merger event occurred and dynamically heated a preexisting disk, giving rise to a hotter component, or that the thick disk stars were formed at earlier times when the gas was very turbulent and hotter, which prevented them from becoming part of the thin disk (Robin et al. 2014). It is likely that radial migration of stars into the thick disk also contributed to its formation, but it is doubtful that this process contributes significantly (Minchev et al. 2012). Rather, the most likely scenario is that the merger with Gaia-Enceladus dynamically heated a gas-rich disk already in place, leading to increased star formation that came to make up most of the thick disk (Gallart et al. 2019; Helmi 2020). It should be noted that substructures and asymmetries observed in our Galaxy do not automatically mean that the objects contained within them were accreted, as mergers can also lead to these phenomena in in-situ populations (e.g., Gomez et al. 2012), and also because they can emerge as a result of dynamical processes within the Galaxy, for instance resonances with the Galactic bar (Helmi 2020).

In general it appears that the inner stellar halo is predominantly made up of just a few building blocks of considerable size such as Sequoia, Thamnos, Sagittarius, but most notably Gaia-Enceladus. In first order, this is in agreement with  $\Lambda$ CDM model predictions such as those of the EAGLE cosmological simulation (Schaye et al. 2015) or IllustrisTNG (Nelson et al. 2019), which include a number of close Milky Way analogues, sometimes even down to the stellar halo (Bignone et al. 2019). As more and more information on the stellar halo and Galactic evolution comes to light, it is increasingly important to implement simulations which incorporate the boundary conditions and constraints provided by that information. At the moment, it is the absence of extensive stellar samples including thorough chemical information that is restricting the investigations into the exact nature of substructures and their pre-accretion properties, most importantly regarding their size as well as star formation and chemical enrichment.

Furthermore, most data available so far is limited to stars in the Solar vicinity. In order to put the substructures discovered into a bigger context, it is thus necessary to extend observations further beyond. In particular, the Galactic centre is the place of the greatest density for the halo, but is still largely uncharted

([Helmi 2020](#)). Several groups of elements can be helpful in identifying and analysing accretion events. For example,  $\alpha$ -poor stars tend to originate from accreted systems, and r-process enhancement can be another indicator of accretion ([Sakari et al. 2018](#)), with field halo populations displaying significant scatter at the low metallicity end, hinting at origins from a number of different small accreted systems ([Roederer et al. 2018](#)). Large surveys can find, for instance, low- $\alpha$  stars which can subsequently be selected and observed with high-resolution spectroscopy ([Helmi 2020](#)). While high-resolution spectroscopic follow-ups are central, it is important to stress the necessity of having large all-sky surveys in order to build up samples that cover all streams and substructures.

### 1.3 Wide-angle Spectroscopic Surveys

The aim of having spectroscopic surveys that cover significant numbers of stars is increasing in priority, as can be seen from the multitude of upcoming projects such as DESI, WEAVE, and MOONS. The Dark Energy Spectroscopic Instrument (DESI; [DESI Collaboration et al. 2016](#)) will carry out ground-based wide-area observations on baryon acoustic oscillations and the growth of structure by measuring distortions in redshift-space. More specifically, the observations will target luminous red galaxies for redshifts of  $z \leq 1.0$ , bright emission line galaxies for redshifts  $z \leq 1.7$ , and quasars in order to discern the dark matter distribution as well as the Ly- $\alpha$  forest absorption characteristics at high redshifts. The redshifts of around 10 million bright galaxies with a median redshift of  $z \approx 0.2$  will be acquired during times when the primary targets listed above cannot be efficiently measured. By the end of the five-year survey, the redshifts of more than 30 million galaxies and quasars will have been determined across a sky area of around 14 000 deg<sup>2</sup>, utilising ten three-armed spectrographs which are being fed by fibres enabling the simultaneous recording of up to 5000 spectra for wavelengths between 360 nm and 980 nm. The resolution of the instrument varies between 2000 and 5500 across the wavelength range. DESI will significantly boost and refine our knowledge of the Universe’s expansion history, thus allowing for a deeper understanding of the nature of dark energy. Further progress is expected regarding the topic of neutrino mass and theories of modified gravity as well as inflation ([DESI Collaboration et al. 2016](#)).

The Multi-Object Optical and Near-infrared Spectrograph (MOONS; [Cirasuolo et al. 2020, 2014](#)) is a spectroscopic facility to be constructed on the ESO’s 8-metre VLT, whose collecting area in the infrared and visible is one of the largest in the world. Two spectrographs are fed by 512 fibres each that can be robotically positioned according to the requirements of the observations, and each spectrograph contains three dispersers for low resolution and two dispersers for high resolution observations. The observed wavelengths lie between 0.64  $\mu\text{m}$  and 1.8  $\mu\text{m}$ , which allows for the observation of strongly dust-obscured regions of the Galaxy and the high-redshift Universe. The spectrograph can be run in low- or high-resolution mode, whereby in the former case the resolution is between 4000 and 7000 across the atmospheric windows in the entire covered wavelength range. For the latter, the two dispersers provide resolutions of more than 9000 and of around 19 000, respectively, which enables detailed chemical abundance determinations. Within the 300 nights of observation time allocated to the developers of the instrument, the aims include probing into strongly obscured regions of the Galactic bulge, the chemodynamical structure of the two Galactic disks, and the evolution of galaxies on cosmic timescales in order to enhance our understanding of the integrated properties of stellar populations and the ionised interstellar medium up to high redshift, which will make it possible to examine the influence of environment and black hole feedback on redshifted galaxies’ formation and evolution. Over the course of 10 years, the instrument will be able to measure tens of millions of radial velocities and precise chemical abundances and in addition millions of galaxy spectra over a redshift range of  $0 < z < 10$  ([Cirasuolo et al. 2020](#)).

The WHT Enhanced Area Velocity Explorer (WEAVE; [Dalton et al. 2014, 2012](#)) is a multi-object spectroscopic facility for the William Herschel Telescope (WHT) on the Canary Islands which provides follow-up for targets identified in optical ground-based and space-based surveys such as LOFAR and Gaia, respectively. It hosts 1000 multi-object fibres feeding a single spectrograph, which features both a low-resolution mode and a high-resolution mode. The low-resolution mode provides a resolution of

$R \sim 5000$  across the complete wavelength range from 370 nm to 1000 nm, whereas the high-resolution mode at  $R \sim 20\,000$  only encompasses sections of this full range within its blue and red arm. It can be utilised to provide accurate radial velocity measurements as well as basic stellar parameters and abundance ratios for Gaia objects down to very faint objects in low-resolution mode, while for brighter objects, high-resolution observations can yield extensive and accurate chemical information which traces stars with a similar formation history. At high redshifts, metallicities and velocity dispersions can be determined for extragalactic sources, and at low to moderate redshifts, the internal structure and kinematics of stellar populations can be observed by WEAVE in order to investigate galaxy evolution in detail (Dalton et al. 2014).

One further upcoming project is 4MOST. The 4-metre Multi-Object Spectroscopic Telescope (4MOST) is a wide-field spectroscopic survey facility to be constructed on the four-metre VISTA telescope at the Paranal Observatory in Chile (de Jong et al. 2019). It has 2436 fibres dispatching the incoming light to three spectrographs with a spectral resolution of around 6500 for two of them and around 20 000 for the third, and can thus carry out observations of around 2400 objects at once, considering that some fibres are utilised to record the background level of sky areas. This means that, in order to use the facility most efficiently, several surveys will be carried out simultaneously. Aside from accomplishing the objectives outlined above, it can also be used for a wide range of additional science cases, and a number of planned surveys carried out by the 4MOST Consortium<sup>1</sup> have already been decided.

Since other large ground-based spectroscopic surveys such as LAMOST (Deng et al. 2012; Zhao et al. 2012; Cui et al. 2012) and the Sloan Digital Sky Survey’s APOGEE (Majewski et al. 2017; York et al. 2000) are situated in the Northern hemisphere, and surveys in the Southern hemisphere like GALAH (De Silva et al. 2015) have a smaller field of view in addition to observing fewer objects, 4MOST is of particular importance in further mapping out the Southern sky. Due to the instrument using the latest technology and the efficiency of the gratings being much higher compared to Echelle gratings, the overall efficiency of the setup allows to include millions of objects down to  $G \sim 17$  in the high-resolution spectroscopic surveys, despite observations being performed on a 4-m telescope. Currently, 4MOST is still being assembled and tested, and science operations are anticipated to begin in 2024.<sup>2</sup>

The majority of the observing time will be taken up by ten surveys conceived of by the 4MOST Consortium, which includes both Galactic (Helmi et al. 2019; Christlieb et al. 2019; Chiappini et al. 2019; Bensby et al. 2019; Cioni et al. 2019) and extragalactic surveys (Finoguenov et al. 2019; Merloni et al. 2019; Driver et al. 2019; Richard et al. 2019; Swann et al. 2019) intended to shine light upon the different components of the Milky Way, and objects such as AGN and galaxy clusters among other things. The remainder of the time will be dedicated to 15 surveys by the ESO community. The second Consortium Survey, “The Milky Way Halo High-Resolution Survey” (Christlieb et al. 2019), from here on referred to as S2, will provide detailed chemical analyses on more than 1.5 million stars, with high precision abundances for many elements being determined from the observed spectra. This will not only make it possible to compare detailed observations to predictions from galaxy formation models and thus disentangle the early (chemical) history of our Galaxy, but will also increase the number of known metal-poor stars considerably. Furthermore, the halo is specifically targeted in order to be able to reconstruct the origin of halo stars, i.e. whether they were formed in situ or stem from accretion events of variable proportions, and to figure out which abundance patterns are associated with which of the aforementioned origins. The target catalogue for the survey will be based on data from Gaia DR3 in conjunction with synthetic Pristine data for the initial metallicity estimates (see Pristine survey, Martin et al. 2023; Starkenburg et al. 2017), but will avoid kinematic selection criteria, and the catalogue size after application of the criteria is a few million stars, of which on the order of one hundred thousand should be actual halo stars.

---

<sup>1</sup><https://www.4most.eu/cms/collaboration/consortium/>, retrieved 22.10.2023

<sup>2</sup><https://www.4most.eu/cms/home/>, retrieved 22.10.2023

## 1.4 Aims

The subject of this thesis is the target catalogue of the upcoming 4MOST S2 survey. Therefore, the 4MOST project will be described in detail in Section 2 regarding its instrumentation and the five-year survey programme carried out by the 4MOST Consortium and the ESO community. As the selection criteria are basically entirely based on Gaia DR3, and metallicity estimates employing the method outlined by Starkenburg et al. (2017) are used in the selection, background information on the Gaia mission – particularly Gaia DR3 – and on the Pristine survey are provided in Section 3, before the selection criteria used to compile the target catalogue for 4MOST S2 are described and explained. The analysis is split into two major parts: the verification of the 4MOST S2 target catalogue (Section 4) and the chemical and kinematic analysis of the catalogue (Section 5). The goal of the first part is to verify that the selection of the target catalogue is consistent, and starts with the creation of colour-magnitude diagrams of the catalogue to give a general characterisation of it. A particularly important aspect is the verification of the metallicities used for the metallicity criterium in the selection, which is why several reference catalogues containing metallicity data are crossmatched to the 4MOST S2 target catalogue for comparison. Another goal is to verify that there are no unexpected selection effects, as potentially interesting objects might be removed otherwise. This is done via a stepwise application of the selection criteria to a reference catalogue, in this case GALAH DR3 (Buder et al. 2021) and briefly SAGA (Suda et al. 2008). Also, a comparison to a simulation of the 4MOST S2 survey via the Besançon Galaxy model aims to check whether the model is consistent with the catalogue or if there seem to some important absences in terms of the catalogue’s content.

The second part starts by outlining and characterising the 6D dataset at the basis of the following analyses and determining a number of parameters relevant to kinematic analyses. The aim is to use these parameters to first separate the halo from the disk in the catalogue, and to then characterise it in detail regarding its stellar number distribution, its appearance in the colour-magnitude diagram, as well as its metallicity distribution function in different kinematic cuts. Once the halo is identified, another goal is to identify and roughly characterise halo substructure present in the 4MOST S2 target catalogue in kinematic space, as substructures are typically the result of mergers and other galaxy interactions (e.g., Helmi 2020) and thus of interest for detailed study using the future 4MOST data in order to better constrain galaxy formation and evolution. The goal of the final subsection is then to select candidate samples within the 4MOST S2 target catalogue containing stars likely enriched by the r-process, the s-process, and the weak r-process. This is done by exploiting the chemical abundances of GALAH DR3 and comparing the neutron-capture abundance patterns to the solar-system r- and s-process components, and to the star HD 122563 (Honda et al. 2006) substituting for the weak r-process. This analysis of the elemental patterns can be extended to the full GALAH DR3 sample, and a long-term goal is to refine the approach outlined here in order to be able to successfully select candidate stars for follow-up observations.

## 2 4MOST

### 2.1 The 4MOST instrument

#### 2.1.1 Telescope and general design

The 4MOST survey facility will be built on the Visible and Infrared Survey Telescope for Astronomy (VISTA) ([Emerson & Sutherland 2002](#); [Emerson et al. 2006](#)), which is part of the European Southern Observatory's (ESO) Paranal Observatory and is situated around 1500 m to the north-east of the Cerro Paranal peak hosting the VLT. It is located in Chile within the Atacama Desert, a region that is both very remote and extremely dry, at a latitude of about  $24^{\circ}40'$  in the Southern hemisphere at an altitude of around 2500 m. Contained within an enclosure building of 19 m diameter that serves a number of purposes, including protection from weather conditions and minimisation of seeing effects, reduction of straylight, regulation of temperature, and housing of additional hardware, VISTA is built as an altitude-azimuth mount which can rotate the telescope around two axes. It consists of two mirrors in a Cassegrain setup, more specifically using a quasi-Ritchey-Chrétien design, where its primary mirror (M1) is a concave hyperboloid that has a diameter of 4.1 metres with an f/1 focal ratio, and the secondary mirror (M2) is a convex hyperboloid that has a diameter of  $\sim 1.24$  metres, with a total focal ratio of f/3.25 for this setup ([Emerson et al. 2004](#)).

The primary mirror has an uncommonly high curvature for a mirror of its size and has a 1.2 m hole for the instrument sitting at the Cassegrain focus to receive the light. In order to implement active optics – which is to reshape the mirror in reaction to external forces – the primary mirror is supported by 81 axial force actuators and 3 additional axial definers, whereas the secondary mirror is fixed to a hexapod with six legs whose lengths can be controlled independently, which can move the secondary mirror around 5 axes. Although not originally designed for this purpose, VISTA is very well-suited for large-multiplex fibre spectrographs, as it has at its Cassegrain focus both a large field-of-view together with a large instrument capability ([Sutherland et al. 2015](#)). In order to exploit the 4MOST instrument to its full potential, the VISTA telescope will be exclusively occupied and utilised by the 4MOST instrument within the scope of a 5-year programme of surveys.

The concept for the design of the 4MOST facility is that it should allow several surveys to be carried out at once, and the capability of the instrument to perform observations for several surveys in parallel enables the execution of surveys that have very low target densities in some regions, which would normally be too expensive. While the key design mission and reason why the 4MOST project was initiated is to complement large-area surveys such as Gaia, Euclid, PLATO, and eROSITA with comprehensive spectroscopic observations ([de Jong et al. 2012, 2014](#)), there are many other issues that may be addressed, and 10 surveys conceived by the 4MOST Consortium have already been approved, which will be discussed in Section 2.2. The 4MOST instrument has to sufficiently perform to meet the requirements of its science missions, which includes within a two-hour observation the measurement of high precision radial velocities ( $\Delta v_r \leq 1$  km/s) down to the faintest Gaia sources ( $G < 20.5$  mag), the determination of the stellar parameters and a few essential elemental abundances at  $G < 18$  mag, as well as up to 15 elemental abundances in the apparent magnitude range brighter than  $G = 15.5$  mag. In addition, the redshifts of galaxies and AGN around  $r \sim 22.5$  mag should be acquired within this timeframe, and the sky coverage during the initial five-year phase encompasses more than  $17\,000 \text{ deg}^2$  that are to be observed twice, which is a significant fraction of the southern sky. During this period of time, millions of spectra will be recorded, whereby the goal is to obtain more than 25 million spectra at low resolution ( $\langle R \rangle \sim 6500$ ) and more than 2 million at high resolution ( $\langle R \rangle \sim 20\,000$ ) ([de Jong et al. 2016](#)).

On the focal surface of the telescope setup, the 4MOST instrument provides a field-of-view of  $2.6 \text{ deg}$  diameter and more than  $4.2 \text{ deg}^2$  in total area with 2436 fibres, which can be arranged according to the demands by a fibre positioner. The light collected by these fibres is put through to three spectro-

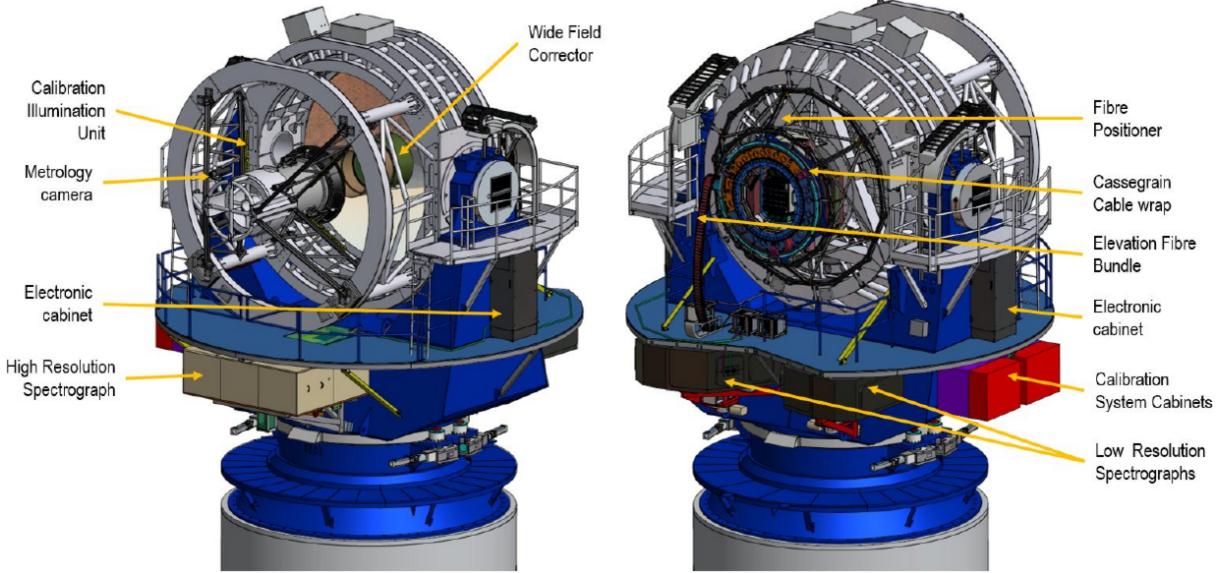


Figure 5: Physical layout of the 4MOST instrument on the VISTA telescope with some subsystems labeled (de Jong et al. 2016).

graphs, two of which perform low-resolution spectroscopy and one of which performs high-resolution spectroscopy. Each spectrograph receives data from one third (812) of the fibres, has three channels for red, green, and blue light, and is fixed in its configuration. Other subsystems besides the spectrographs, the fibre-feed system, and the fibre positioner include metrology cameras, detectors and a cryostat, the calibration system, the Wide Field with Atmospheric Dispersion Corrector, and the Acquisition & Guiding and Wavefront Sensor Systems. Figure 5 depicts the general layout of the 4MOST instrument on the VISTA telescope, with some of the subsystems' locations indicated. The key specifications of the combination of all these subsystems into the 4MOST instrument are summarised in Table 1. In the following paragraphs, the individual subsystems will be briefly described.

### 2.1.2 The 4MOST subsystems

In order to have a pupil-centric and aberration corrected focal surface, the Wide Field Corrector (WFC) system is integrated into the design of the 4MOST facility. As it corrects the images from the VISTA telescopes optics, it is of critical importance to achieve the required field-of-view, plate scale, and image quality, and also incorporates the Atmospheric Dispersion Corrector (ADC; Cunningham et al. 2023). It consists of four elements, out of which two are singlets and two are wedged doublets, the latter of which comprise the ADC part. This part rotates about the central axial axis in order to ensure satisfactory chromatic coupling between stellar images and the fibre optics extending across a large range of zenith angles (Azais et al. 2016). At high zenith angles, the light from the edge of the 4MOST wavelength range may fail to reach the fibre due to dispersion, and the ADC is limited to zenith angles below 55°. The observed target's light passing through the telescope and the WFC/ADC system is coupled into the fibre, and the coupling efficiency sets the condition on the performance of the system. The goal is to reach an efficiency where 85% of a target's light at any wavelength is contained within a circle of maximum diameter 0.8 arcsec, or 47.8  $\mu\text{m}$  on the focal surface, which has been reproduced in simulations across the whole field (de Jong et al. 2016; Azais et al. 2016). In addition to the WFC/ADC, additional elements ensuring the sky image's quality are four wave front sensing (WFS) cameras around the field-of-view in order to control the primary mirror's shape and the secondary mirror's position, as well as two Acquisition & Guiding (A&G) units for pointing and for raising the sky coverage (de Jong et al. 2016, 4MOST User Manual).

Table 1: 4MOST instrument specifications, adapted from the 4MOST User Manual<sup>3</sup>, de Jong et al. (2019), and de Jong et al. (2016). The sensitivities in the final row are given in AB magnitudes.

Parameter	Design value	
Hexagonal field-of-view (FOV)	$>4.2 \text{ deg}^2$ , diameter $2.6 \text{ deg}$	
Accessible sky (zenith angle $<55^\circ$ )	$>30\,000 \text{ deg}^2$	
Number of fibres per area	$>3$ within a circle of 2 arcmin diameter at 3.5 arcmin from the edge of the FOV	
On-sky diameter of fibre aperture	1.45 arcsec circular	
Smallest target separation	15 arcsec on any side	
Expected on-target fibre-hours	LRS: $>3\,200\,000 \text{ h/yr}$ ; HRS: $>1\,600\,000 \text{ h/yr}$	
Multiplex fibre positioner	2436	
	Low Resolution Spectrographs (LRS)	High Resolution Spectrograph (HRS)
Number of fibres	812 fibres	812 fibres
Wavelength range	370-950 nm	392.6-435.5 nm; 516-573 nm; 610-679 nm
Spectral resolving power	$R > \lambda \cdot 10$ for $400 \text{ nm} < \lambda < 500 \text{ nm}$ and $R > 6000$ for $500 \text{ nm} < \lambda < 885 \text{ nm}$	$R \geq 18\,000$
Radial velocity calibration accuracy	$\leq 1 \text{ km/s}$	$\leq 1 \text{ km/s}$
Spectral sampling	PSF FWHM $\geq 2.5 \text{ pixel}$	PSF FWHM $\geq 2.5 \text{ pixel}$
Mean sensitivity $6 \times 20 \text{ min}$ , mean seeing; LRS: $S/N = 10 \text{ \AA}$ ; HRS: $S/N = 100 \text{ \AA}$	400 nm: 20.2; 500 nm: 20.4; 600 nm: 20.4 700 nm: 20.2; 800 nm: 20.2; 900 nm: 19.8	420 nm: 15.7; 540 nm: 15.8; 650 nm: 15.8

In the focal surface, a hexagonal array of 2436  $85 \mu\text{m}$  fibres collects the incoming light, with an additional 12 guide fibre bundles placed around the science field. Using the tilting-spine principle which was first put into practice for FMOS-Echidna (Sheinis et al. 2014; Akiyama et al. 2008), these fibres are positioned according to the targets’ locations by the AESOP Fibre Positioner (Brzeski et al. 2018). The benefit of using a tilting-spine fibre positioner is that a large area can be covered by each fibre: While the separation between two spine tips is 85 mm or  $\sim 161$  arcsec, a circular area of radius  $> 11.8$  mm or  $\sim 200$  arcsec is accessible for each spine, meaning there is some overlap in these areas (although the smallest attainable separation between two fibres is anticipated to be  $\sim 15$  arcsec). In fact, at least three LRS fibres, and one or two HRS fibres cover each position in the science field-of-view, which is particularly relevant for clustered targets, where an increased fibre-to-target allocation completeness can be reached because of this (4MOST User Manual). The AESOP Fibre Positioner must be capable of providing a positional accuracy better than 0.2 arcsec or  $10 \mu\text{m}$  rms radial error, and also to retain these positions during the time of observation (de Jong et al. 2016). The hexagonal array of fibres is shown in Figure 6, which also indicates the mapping of the fibres to each of the three spectrographs. This mapping is quite complex, as each spectrograph’s fibres are supposed to be distributed evenly on the hexagon.

The accuracy of the fibre positions in the focal surface will be ensured by four Metrology Cameras, which determine those positions for the feedback loop control doing the positioning for the setup of observations. Together with the Back Illumination Units, they make up the Metrology System (Winkler et al. 2016, 2018). The Back Illumination Units are sitting at the slits of the spectrographs and serve as the illumination source for the fibres in order for the cameras to determine the fibre position, as well as serving as shutters for the spectrographs. The four cameras are attached to the M2 mirror’s support and it must be ensured that they remain thermally stable. Using four cameras allows for a higher imaging quality when combining them, and is also important in case one camera fails (de Jong et al. 2016). As the accuracy and speed of the control loop positioning the fibres is of prime importance to the time needed to set up an observation configuration, it is also highly important to the overall efficiency of 4MOST. With a fast metrology system, the 2436 fibres can be arranged in the desired configuration with the required accuracy in less than two minutes (4MOST User Manual).

<sup>3</sup>[https://www.4most.eu/cms/files/VIS-MAN-4MOST-47110-9800-0001\\_2\\_00-4MOST-User-Manual.pdf](https://www.4most.eu/cms/files/VIS-MAN-4MOST-47110-9800-0001_2_00-4MOST-User-Manual.pdf), retrieved 22.10.2023

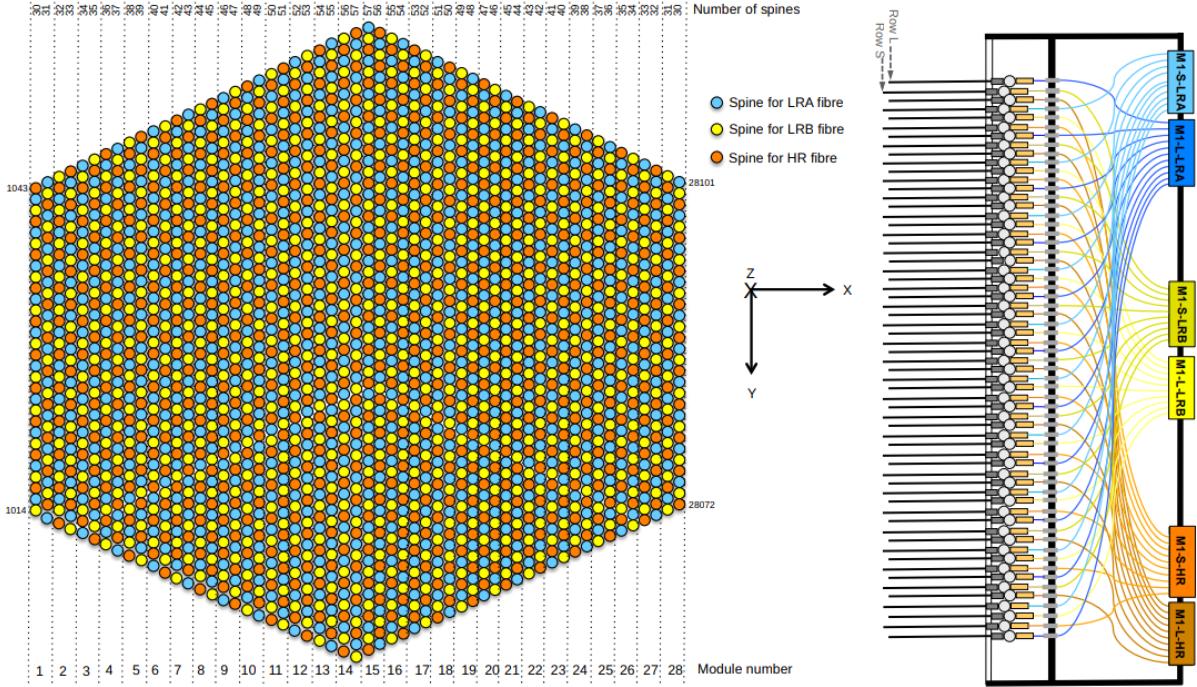


Figure 6: The array of fibres in the focal surface, colour-coded according to which spectrograph they feed. On the right, the module spine rows with spectrograph group and connector are depicted ([Haynes et al. 2016](#)).

In order to transmit the light from the fibres at the Cassegrain focus to the slits of the spectrographs, the Fibre-Feed System is employed. It is made up of two portions: the Positioner Cable, which extends from the fibre positioner to the fibre connector unit, and the Spectrograph Cable, which extends from the fibre connector unit to the slits of the spectrographs and leads the main cable through the de-rotator unit, the elevation fibre chain, and the fibre strain release units ([de Jong et al. 2016](#)). The fibre cable de-rotator unit allows for rotations during the observations of  $\pm 200$  degrees, and the elevation fibre chain accommodates altitude changes all the way from the horizon to the zenith, whereas the fibre strain release unit supplies some additional length and facilitates the assembly. Besides the 2436 science fibres, the fibre-feed system transmits light for the guide fibres, the fibres connecting the positioner to the Back Illumination Units, a number of fibres for the back-illumination of the system fibres, as well as the fibres that deliver the light from the calibration system to the spectrographs' slits. The fibres themselves are silica/silica, step-index, multi-mode fibers with polyimide coating as buffer, which should mean low attenuation as well as a low degree of fibre focal ratio degradation ([Haynes et al. 2016](#)).

The Calibration System of the 4MOST facility contains a broad-band source for continuum light used to obtain flat-fields to calibrate individual fibres, a Fabry-Perot etalon illuminated by a Laser Driven Light Source which creates a regular pattern of spectral lines for the calibration of the wavelengths, and ThAr lamps providing narrow, well-studied lines for additional wavelength and resolution calibration. This light can go both through the telescopes and the science fibres as well as straight through the calibration fibres to the spectrographs to provide precise wavelength calibrations, as the radial velocity accuracies are supposed to be better than 1 km/s ([4MOST User Manual](#)). The Calibration System is made up of a lamp cabinet containing the lamps as well as illumination units on the M2 supports to send the light through the optical train for calibration ([de Jong et al. 2016](#)). The complete calibration of 4MOST data involves the instrument calibration, the data calibration concerned with clearing away instrument signatures, and finally the science calibration at the high level. The control, coordination, and monitoring of the subsystems are carried out by the Facility Control Hardware and the Facility Control Software, whereas the Operations System is involved in the planning and production of observing blocks as well as monitoring the survey program's progress. It contains, among other tools, the 4MOST Facility Simulator,

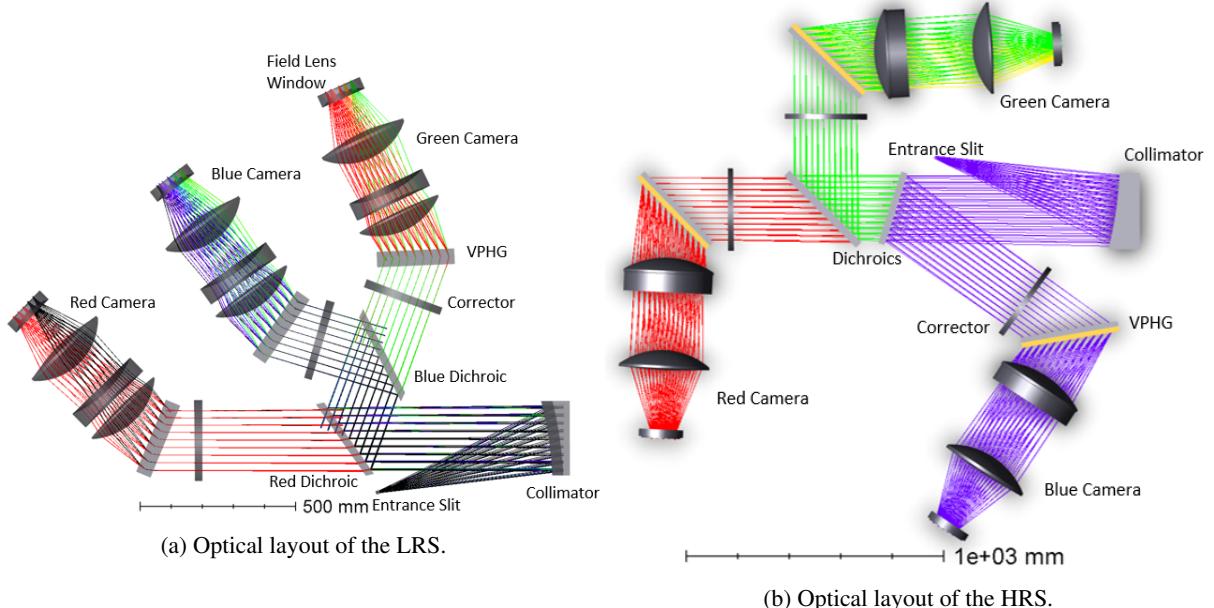


Figure 7: Optical layout of the 4MOST spectrographs.<sup>4</sup>

which allows for the simulation assisting significantly in the planning stages of 4MOST via simulations of instrument properties and surveys (Dwelly et al. 2016). Finally, the data processing, quality control, and archiving is taken care of by the Data Management System.

### 2.1.3 The low- and high-resolution spectrographs

Both the two LRS and the HRS are located beneath the telescope's main platform, which shelters them better from outer conditions compared to the telescope itself and the 4MOST subsystems located in the focal plane. Besides the 812 science fibres connecting to each spectrograph, there are 6 calibration fibres connected to either end of all entrance slits of the spectrographs. The LRS and the HRS both have a blue, green, and red channel in a three-camera fixed setup constituting three wavelength bands (de Jong et al. 2016), and their optical layout is summarised in 7a and 7b for the LRS and HRS, respectively.

In the LRS's entrance slit, the 812 science fibres plus 10 calibration fibres are in a configuration of 28 slitlets of 29 fibres each for the science fibres and 2 slitlets of 5 fibres each for the calibration fibres, whereby the separation between the centres of two fibres within a slplet is  $170\ \mu\text{m}$ , and the distance between neighbouring slitlets is 5.49 mm. The fibres are arranged according to the focal plane's curvature, and the first surface of the slit lens glued to the entrance slit has the same curvature, with the fibres being glued directly onto that surface. After the entrance slit, a collimator mirror reflects the light, turning it into a collimated beam of 200 mm and putting the fibres at infinity, as the light's rays run parallel now. This is also where the Back Illumination Unit serve as a light source for the fibres' position measurements and as a shutter for the spectrograph. To separate the light into a red, green, and blue channel, two dichroic mirrors are used, the first of which only lets the red light pass and reflects the rest of the light at an angle, whereas the second only lets the green light pass and reflects the blue light. The wavelength ranges of these three arms are from 370 nm to 554 nm for the blue arm, from 524 nm to 721 nm for the green arm, and from 691 nm to 950 nm for the red arm, making for a total wavelength coverage from 370 nm to 950 nm for the LRS (Laurent et al. 2016).

Following this splitting of the light, there is a different plano off-axis asphere corrector for each channel correcting the collimator image. Due to their efficiency across a large wavelength range, volume phased

<sup>4</sup><https://www.4most.eu/cms/facility/subsystems/>, retrieved 22.10.2023

holographic gratings are used, which function as transmission gratings and are placed at the average locations of the entrance pupil images. The gratings' aperture is 250 mm and their Bragg angles are  $18.35^\circ/22.49^\circ/22.46^\circ$  with 1351/1220/924 lines per mm for the blue, green, and red channel, respectively. The dichroics, correctors, and gratings are all made of silica. Finally, the three cameras are each made up of three lenses, out of which the second and third are concave and convex aspheric lenses, respectively, which together with a final lens (the Field Lens) make a final separation of all the incoming light according to the wavelength. The Field Lens doubles as the entrance to the detector system, and a CCD detector of 6k by 6k pixels with a pixel size of  $15\ \mu\text{m}$  is located 4 mm behind the back of the Field Lens ([Laurent et al. 2016](#)). As the temperature needs to be kept stable below 150 K, the detector system is contained within a cryostat. The maximum shifts due to thermal variations are estimated to be  $0.6\ \mu\text{m}$  in X and  $0.33\ \mu\text{m}$  in Y. According to calculations, the flux contamination is about 0.75%, mainly due to scattered light, and the total throughput of the LRS is about 73.8% on average ([4MOST User Manual](#)). In Figure 8, an estimate of the spectral resolution  $R = \lambda/\Delta\lambda$  of the LRS as a function of wavelength is shown. The spectral resolution ranges from  $R \approx 4000$  to  $R \approx 8000$ , with a mean of  $R \approx 6500$ .

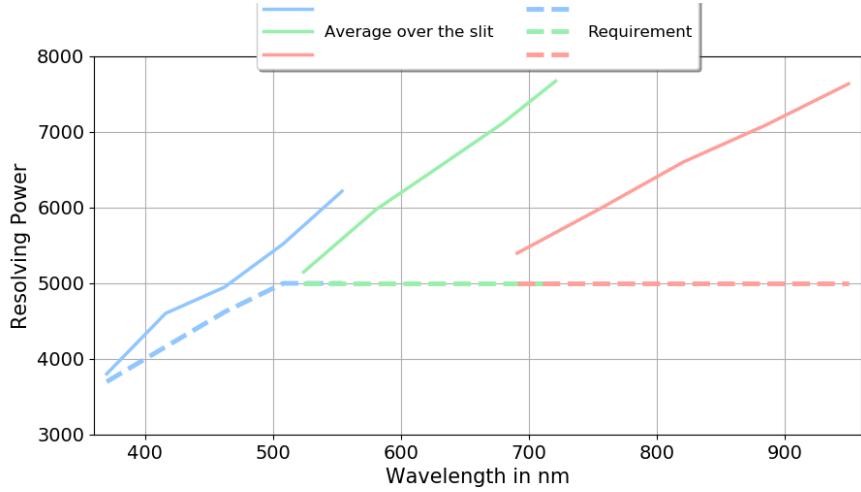


Figure 8: Spectral resolution of the LRS as a function of wavelength.<sup>5</sup>

In many ways, the optical layout of the HRS is similar to that of the LRS. The HRS is also fed by 28 slitlets containing 29 science fibres each, with a separation of  $170\ \mu\text{m}$  between two fibres, plus an additional two calibration slitlets on each end containing 5 fibres each. These 30 slitlets with the 822 fibres are set up as a curved long-slit with a length of 153.2 mm, following the curved focal surface of the subsequent collimator, which is a spherical off-axis mirror. After this, the light is split into three channels by the dichroics, where the first dichroic reflects the blue light at an angle and lets the green and red light pass, while the second dichroic reflects the green light and lets the red light pass. Thus, the light is separated into a blue arm with a wavelength range from 392.6 nm to 435.5 nm, a green arm with a wavelength range from 516.0 nm to 573.0 nm, and a red arm with a wavelength range from 610.0 nm to 679.0 nm ([Seifert et al. 2016](#)).

Following the dichroics, a corrector plate is inserted to correct the collimator image. Volume phase holographic gratings diffract the beams of 250 mm size as a transmission grating, with the direction change of the chief ray being close to  $90^\circ$ . The gratings for the three channels have 3362/2542/2161 lines per mm for the blue, green, and red channel, respectively. The cameras receiving the diffracted light are comprised by three lenses and one doublet in a similar manner to the LRS. The field lens is a doublet which also serves as the entrance to the detector cryostat, and the CCD detector is mounted at the focal distance of 5 mm behind the back of the field lens. As is the case in the LRS, the detector consists of 6k by 6k pixels of size  $15\ \mu\text{m}$  ([Seifert et al. 2016](#)). Thermal variations of  $1^\circ\text{C}/\text{h}$  will lead to average shifts of less than  $0.8\ \mu\text{m}/\text{h}$ , the flux contamination in the HRS due to scattered light ranges between 0.57% and 1.28%, and the total throughput of the HRS ranges between 0.3 and 0.55 for all three

channels depending on wavelength. Figure 9 depicts the spectral resolution of the HRS as a function of wavelength for all three channels, showing the range of the spectral resolution to be from  $R \approx 18\,500$  to  $R \approx 21\,000$ , with a mean of  $R \approx 20\,000$ . For both the LRS and the HRS, spectra are sampled at  $\sim 3$  pixels per resolution element. The CCD detectors are characterised by a low read-out noise of  $< 2.3$  electrons per read, as well as a high quantum efficiency across a broad range of wavelengths ([4MOST User Manual](#)).

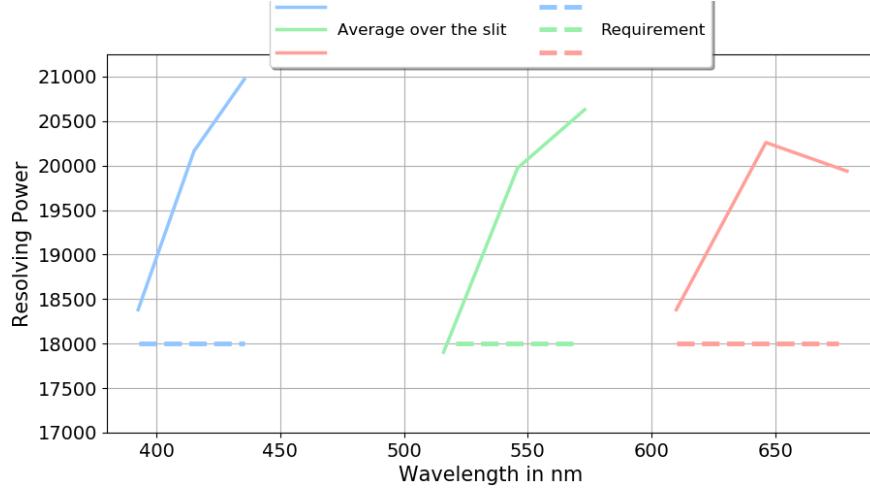


Figure 9: Spectral resolution of the HRS as a function of wavelength.<sup>5</sup>

## 2.2 4MOST survey programme

The special concept of observations within the framework of the 4MOST survey programme is that many different surveys of both low and high resolution share the focal plane, as is enabled by having 2436 fibres available for simultaneous observations. This increases the efficiency of the observations significantly, as generally more fibres can be allocated targets and less fibre hours are wasted. Still, in the direction of the Milky Way, the high target density means that there are not enough high-resolution fibres, whereas at high galactic latitudes there will be many free high-resolution fibres, as extragalactic observations are typically carried out at low resolution. Even when carrying out several surveys at once, a target density that is too small will lead to not all fibres being utilised. Because of this, a detailed survey strategy plan has to be put into place, as additionally the surveys have different target brightnesses, and require different signal-to-noise ratios and different sample completeness levels ([Guiglion et al. 2019](#)). Furthermore, some of the surveys will be working together by combining their data to approach certain scientific questions, such as S1 and S3 working together to study the disk/halo interface, or the data of S1-S4 being combined to produce a comprehensive picture of the merger and assembly history of the Milky Way. An overview of the Consortium Surveys, the Community Surveys and some basic information is given in Table 2. Besides the ones listed, there are seven more smaller surveys which have been merged with a Consortium Survey. The 25 total surveys can be divided into Galactic surveys and extragalactic surveys depending on their target objects. They will be briefly described in the following subsections.

### 2.2.1 Galactic Surveys

**2.2.1.1 The Milky Way Halo Low-Resolution Survey (S1)** S1 ([Helmi et al. 2019](#)) targets the halo of the Galaxy and aims to study its formation and evolution, in particular the 3D distribution of mass as well as its assembly history. More specifically, the goals include the derivation of some characteristic parameters in addition to the shape and density profile of the Milky Way's dark matter halo, investigating how the properties of cold streams are influenced by perturbations due to clumps, and to map out the

<sup>5</sup><https://www.4most.eu/cms/facility/capabilities/>, retrieved 22.10.2023

Table 2: List of the 4MOST Consortium and Community Surveys including their typical magnitude ranges as well as the minimum number of observed targets within the five-year survey programme (Guiglion et al. 2019).

	<b>Consortium Surveys</b>	<b>Spectroscopy</b>	<b>Brightness range</b>	<b># targets</b>
S1	The Milky Way Halo Low-Resolution Survey	LRS	$15.0 \leq G \leq 20.0$	> 1.5M
S2	The Milky Way Halo High-Resolution Survey	HRS	$12.0 \leq G \leq 17.0$	> 1.5M
S3	Milky Way Disc and Bulge Low-Resolution Survey	LRS	$14.0 \leq G \leq 19.0$	> 10.0M
S4	Milky Way Disc and Bulge High-Resolution Survey	HRS	$10.0 \leq G \leq 15.5$	> 2.5M
S5	eROSITA Galaxy Cluster Redshift Survey	LRS	$18.0 \leq r \leq 22.0$	> 1.7M
S6	Active Galactic Nuclei	LRS	$18.0 \leq r \leq 22.8$	> 1.0M
S7	Wide-Area VISTA Extragalactic Survey	LRS	$18.0 \leq r \leq 22.5$	> 1.6M
S8	Cosmology Redshift Survey	LRS	$20.0 \leq r \leq 23.9$	> 8.0M
S9	One Thousand and One Magellanic Fields	LRS/HRS	$10.5 \leq G \leq 19.5$	> 0.5M
S10	The Time-Domain Extragalactic Survey	LRS	$18.0 \leq G \leq 22.5$	> 0.3M
	<b>Community Surveys</b>	<b>Spectroscopy</b>	<b>Brightness range</b>	<b># targets</b>
S11	The White Dwarf Binary Survey	LRS/HRS	$10 \leq G \leq 20$	~170k
S12	The 4MOST Survey of Young Stars	LRS/HRS	$10.0 \leq G \leq 18.5$	~135k
S13	Stellar Clusters in 4MOST	LRS/HRS	$10 \leq G \leq 20$	~160k
S14	The 4MOST Survey of Dwarf Galaxies and their Stellar Streams	LRS/HRS	$13 \leq G \leq 21$	~140k
S15	CHANCES: A CHileAN Cluster galaxy Evolution Survey	LRS	$r \leq 20.5$	~300k
S16	Chilean AGN/Galaxy Extragalactic Survey	LRS	$r \leq 23$	~3.5M
S17	Understanding the Baryon Cycle with HR QSO Spectroscopy	HRS		~1M
S18	The 4MOST Hemisphere Survey of the Nearby Universe	LRS	$J \leq 18$	~5.5M

prevalence of kinematic substructures in three dimensions. Furthermore, both the halo’s metallicity and elemental abundance distribution as well as the distributions of individual substructures are to be investigated, and the overlap of the stellar halo with the thick disk will be studied in conjunction with the S3 survey targeting the Milky Way’s bulge and disk. In order to be able to achieve these goals, all giant stars in the magnitude range  $15 \leq G \leq 20$  are targeted, even in regions of low target density, which is why using a wide-field instrument such as 4MOST is critical to be able to observe an extensive area within a realistic timeframe.

Based on photometry, parallaxes, and proper motions from Gaia, candidate halo giant stars are selected, and known members of streams are also included in addition, even if they are fainter in absolute magnitude. The declination range follows the restrictions of the overall 4MOST project and will thus range from  $5^\circ$  down to  $-80^\circ$ , and in addition, only Galactic latitudes  $|b| > 20^\circ$  are targeted. According to simulations, more than 10 000 square degrees have to be observed to satisfy the scientific requirements, and the S1 survey aims to cover more than 15 000 square degrees. The catalogue produced by this survey will be a complete sample of halo giant stars, being limited only by magnitude and spatial constraints, including radial velocities at precisions around  $1 - 2$  km/s and metallicities with accuracies better than 0.2 dex. Precise  $\alpha$ -abundances will also be derived at sufficient signal-to-noise ratios, which is crucial in order to characterise substructures in the halo in more detail. Since precise radial velocities will be obtained down to  $G \sim 20$  and distances up to 250 kpc (1 Mpc for some AGB stars), the 3D distribution as well as total mass of the Milky Way will be constrained much more accurately than currently, as current estimates include only very few tracers at large distances. Finally, the analysis of the abundance-kinematic space will allow for registering faint stellar streams as well as any other substructures. An interesting consequence of this survey will be the discovery of extremely metal-poor stars.

**2.2.1.2 The Milky Way Halo High-Resolution Survey (S2)** One of the primary science driver for 4MOST’s Galactic surveys is the investigation of the Galactic halo. Beyond S1, which observes halo stars in low resolution, the Milky Way Halo High-Resolution Survey (S2; Christlieb et al. 2019) will provide more detailed, high-resolution information on the stars that will be observed. There are several particular science goals laid out by Christlieb et al. (2019): Firstly, to identify and characterise the elemental abundance patterns of stars from different stellar populations, namely those that were formed

in-situ, those that were accreted during large-scale merger events, and those that were accreted from low-mass systems. Secondly, to investigate to what extent accretion from globular clusters played a role in the build-up of the Galactic halo by discerning the stars with such an origin. Thirdly, to better constrain the chemical evolution of the Milky Way in its earliest stages and to identify which nucleosynthetic processes contributed to chemical enrichment utilising stars at the low-metallicity end. Because of this, S2 will also aim to dramatically increase the number of known metal-poor stars, particularly at the lowest metallicities, in order to find as many second-generation stars as possible. The exploration of these topics will make it possible to uncover the formation history of our Galaxy and its early chemical evolution, as well as further mapping out kinematic substructure in the halo in three spatial dimensions. Further cases that can be addressed within the scope of 4MOST’s Galactic halo surveys concern the dark matter halo’s density profile, shape, and some of its characteristic parameters, as well as the halo’s metallicity distribution function, in particular towards the low-metallicity end ([Feltzing et al. 2018](#)).

In order to achieve these goals, the spectra of more than 1.5 million stars will be recorded in high resolution at high latitude, with the chemical abundances for 10 to 20 elements being determined from these spectra, whereby the exact number of elements determined for each star can be different for differing stellar parameters. As several science cases will be addressed, and as it is desired to obtain a picture of the stars’ elemental patterns that is as complete as possible, it is important that each of the various element groups is represented among the abundances that are being determined. This encompasses light elements such as C,  $\alpha$ -elements such as Mg or Ca, and also some important neutron-capture elements such as Ba, Sr and Eu. To discriminate between separate stellar populations and to be able to pinpoint the possible origin of a star, it is required that the accuracy of elemental abundance ratios is better than about 0.2 dex. Since at high galactic latitude, the number of stars per sky area which can be observed at high resolution is quite small, the sky coverage needs to be rather extensive, which is why the planned area to be covered is more than 14 000 deg<sup>2</sup>. The signal-to-noise ratio per pixel to be achieved in a successful observation for dwarf and subgiant stars is 50 for brighter stars and 25 for fainter stars, whereas for giant stars, the spectral success criteria are 30 for brighter stars and 15 for fainter stars ([Christlieb et al. 2019](#)). In Figure 10, synthetic spectra at the requisite signal-to-noise ratio are depicted for a dwarf and a giant, each at two different metallicities. The spectra show that even at [Fe/H] = −4.0, several important lines – including Fe lines – can be detected.

The selection of targets (see Section 3.3) will consider the full Gaia data, which makes it possible to remove foreground M dwarfs, which might otherwise contaminate samples of red giant stars due to their similar spectral classes. Furthermore, taking into account a rough initial estimate of the metallicities of the stars allows for the selection of only known metal-deficient stars, which should increase the fraction of genuine halo stars ([Feltzing et al. 2018](#)). Using the selection criteria on the Gaia DR2 Mock Stellar Catalog by [Rybicki et al. \(2018\)](#), [Christlieb et al. \(2019\)](#) found that on the order of 200 000 actual halo stars should be part of the target catalogue for S2. As cosmological simulations suggest that there are about a few hundred kinematic substructures within the Galactic halo ([Helmi et al. 2003; Gomez et al. 2013](#)), this number of halo stars is necessary to have a sufficient number of stars for each substructure for a meaningful characterisation of its properties. However, it should be noted that the actual target selection will be carried out using Gaia DR3, which has been published since. As for metal-poor stars, it is expected that S2 will contain on the order of 10<sup>4</sup> stars at [Fe/H] < −2.0, and will add several hundred to the list of known stars in the ultra metal-poor range below [Fe/H] < −4.0, dramatically increasing the number of known metal-poor stars observed at high resolution ([Christlieb et al. 2019](#)).

**2.2.1.3 The Milky Way Disc and Bulge Low-Resolution Survey (S3)** S3/4MIDABLE-LR ([Chiappini et al. 2019](#)) is designed to study the disk and bulge regarding its kinematic and chemical substructure, and aims to combine estimates of the bar’s and spiral arms’ structure as well as asymmetries in the mid-plane with precise age and elemental abundance determinations in order to deduce the Milky Way’s enrichment history across time and Galactic radius. The specific goals of this survey are an en-

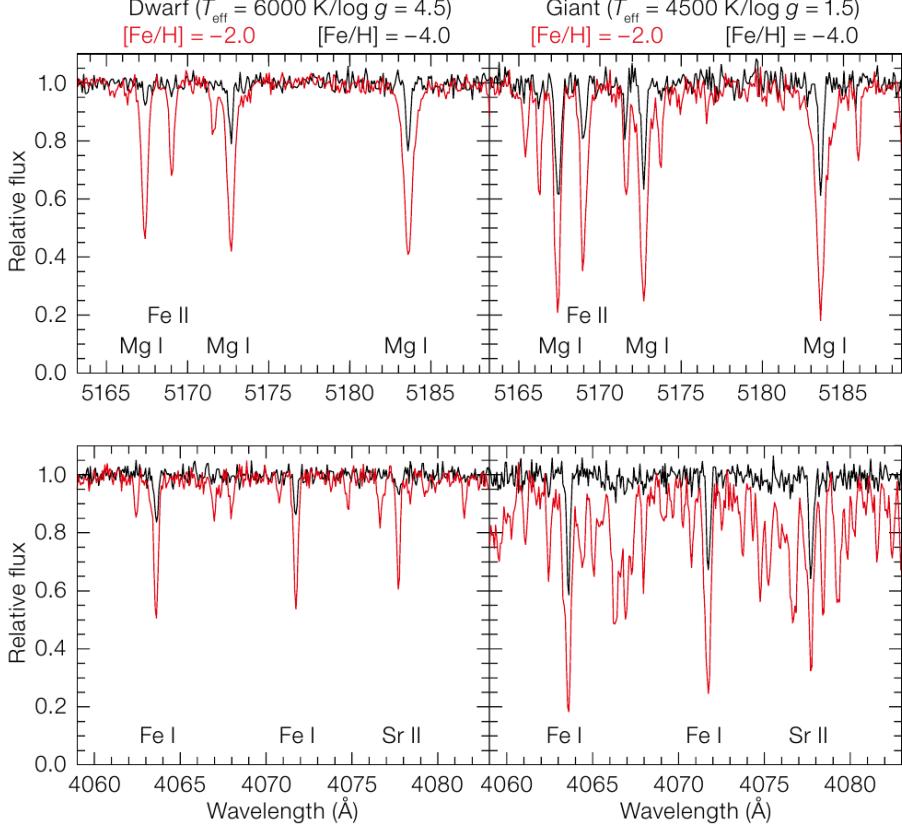


Figure 10: Synthetic spectra of a dwarf and giant at metallicities  $[Fe/H] = -2.0$  and  $[Fe/H] = -4.0$  in two different wavelength ranges containing iron and magnesium lines (Christlieb et al. 2019).

hanced understanding of the Galaxy’s disk structure, disk dynamics, bulge/bar formation, and the disk’s evolutionary history, in other words to obtain a chrono-chemo-kinematical map covering large parts of the Galaxy accurately. In the targeted magnitude range  $14 \leq G \leq 19$  mag, there are hundreds of millions of stars detected by Gaia, but only  $\sim 15$  million stars are selected, with an attempt at being as unbiased as possible while sampling the disk in its full extent. Thus, for the main sub-surveys defined in S3, the selection will be based only on apparent magnitudes, parallaxes, and parallax uncertainties. The main sub-surveys are targeting the Extended Solar Neighbourhood, dynamical disk, faint dynamical disk, chemodynamical disk, and the bulge/inner galaxy, respectively, but there are a number of additional minor sub-surveys focusing on certain classes of stars such as very metal-poor stars or white dwarfs. This survey will produce a sample of stars with high-precision radial velocities at  $1 - 2$  km/s and abundances for a number of elements with precisions of, depending on the signal-to-noise ratio, up to  $0.1 - 0.2$  dex. This sample will have a size that is larger than that of any comparable survey, past or present, and reaches out further in distance and provides better precisions than would be possible with Gaia alone.

**2.2.1.4 The Milky Way Disc and Bulge High-Resolution Survey (S4)** S4/4MIDABLE-HR (Bensby et al. 2019) aims to study the Galaxy’s evolution by taking high-resolution spectra at  $R \sim 20\,000$  and from these derive detailed and accurate abundances for a large number of stars in the Galactic disk and bulge. The specific scientific goals include the detailed study of the disk abundance structure patterns and the nature of the disk components, the study of the chemo-dynamical properties of the Galactic bulge, the formation and evolution of the Galactic bar, as well as the investigation of the Galaxy’s growth history via directly obtained data on kinematic and chemical properties of the outer disk and the inner halo, which yields information on the times of infall as well as masses of merging satellites. The data obtained within the scope of the 4MOST S4 survey may also be used to study secular processes like radial migration, exploring the disk substructure of the Milky Way, and when used in conjunction with complementary

missions such as Gaia can produce additional parameters like masses or ages, allowing for the creation of multi-dimensional datasets for the analysis of the Galaxy’s origin and structure. In order to be able to carry out a detailed study on the formation of the Galactic components, around 2.5 million stars need to be observed, and with about 4 million targets down to a magnitude of  $G = 15.5$  mag, S4 is the most extensive optical high-resolution survey in regards to the survey area, number of targets, and magnitude limit. It will supply rotation, radial velocities with a precision better than 1 km/s, abundances of many elements from all nucleosynthetic channels with an accuracy of 0.05 dex for most elements where the signal-to-noise ratio exceeds 100 per Å, and multi-epoch spectra for some stars.

**2.2.1.5 The One Thousand and One Magellanic Fields Survey (S9)** S9/1001MC ([Cioni et al. 2019](#)) is aimed at studying the Milky Way’s most massive satellites, the two Magellanic Clouds. The specific scientific goals of this survey are the investigation of kinematic and elemental abundance patterns within the Magellanic Clouds, if there are links between these and their location in different stellar populations as well as links to star formation and the system’s dynamical evolution. In order to disentangle their formation and interaction history, the kinematics and chemical abundances will be determined for a range of stellar populations including main sequence stars, red clump giant stars, red giant branch stars, Cepheids, RR Lyrae stars, and more. In total, around half a million stars spanning a magnitude range of almost 10 mag down to  $G = 19.5$  mag distributed across 1000 square degrees are targeted by both the LRS and the HRS, assembling a sample that is around 20 times larger than the largest Magellanic Clouds sample put together in the past. The targets are selected with data from the near-infrared observations of VISTA and 2MASS used in conjunction with Gaia data. Whereas stars for observations with the LRS are selected from VISTA data, those for observations with the HRS are selected from 2MASS data, and specifically parallaxes of Gaia are employed to remove Milky Way stars from the sample and to ensure homogeneous coordinates. The goal of S9 is to measure radial velocities at an accuracy better than 2 km/s, and to obtain metallicities more accurate than 0.2 dex, as this level of precision is normally needed in order to tell different stellar populations apart.

**2.2.1.6 The White Dwarf Binary Survey (S11)** S11/WDB ([Toloza et al. 2023](#)) sets out to investigate stellar evolution and interactions under varying astrophysical conditions using binary systems of stars. It consists of three sub-surveys: The compact white dwarf binary sub-survey, which will study topics such as the mechanisms of angular momentum loss, the formation of magnetic white dwarfs in close binaries, the sources of energy in the common-envelope phase, as well as the eventual fate of close white dwarf binaries; the common proper motion pairs sub-survey, which will gain observational data on the chemical evolution and mechanisms of dynamical heating of the Galaxy by employing white dwarfs as “clocks”; and the central star of planetary nebulae sub-survey, which focuses on the characterising the central star by inspecting its spectra for signs of binarity. The former sub-survey contains the vast majority of the approximately 170 000 targets that were selected using several catalogues in combination with Gaia DR3, and the target catalogue covers a magnitude range from  $G = 10$  mag to  $G = 20$  mag. While most targets will be observed with the LRS, a significant fraction of the spectra will also be obtained with the HRS.

**2.2.1.7 The 4MOST Survey of Young Stars (S12)** S12/4SYS ([Sacco et al. 2023](#)) will use relatively close-by (<500 pc) young stars of low mass in order to gain insights into the spatial and dynamical evolution of star-forming structures, the local star formation rate and chemical inhomogeneities, and to significantly raise the number of known young stars, with the underlying goal being the construction of a catalogue of such stars containing 3D positions with a precision of 1 pc and 3D space motions with a precision better than 1 km/s, a variety of stellar parameters and properties, detailed chemistry and more. The total target catalogue will be split into two samples. The first one will consist of stars in the magnitude range between  $G = 10$  mag and  $G = 18.5$  mag, selected from absolute colour-magnitude diagrams based on Gaia data as those stars that lie above an isochrone with an age of 40 Myr, lie in a colour range that corresponds to spectral types K7 to M6, and have an absolute magnitude  $M_G > 4$ . This

leaves around 100 000 stars, of which the brighter ones will be observed with the HRS, and the fainter ones with the LRS. The second one will consist of stars in the magnitude range between  $G = 10$  mag and  $G = 15.5$  mag that are selected similarly, but instead using an isochrone of age 120 Myr and a colour range for spectral types G7 to K7. An additional rotation-based selection is carried out using TESS data, and the target catalogue of around 40 000 stars will be observed at high resolution. Through this survey, both the volume examined and the number of known young field stars will be raised by two orders of magnitude.

**2.2.1.8 Stellar Clusters in 4MOST (S13)** S13 ([Lucatello et al. 2023](#)) will observe basically all globular clusters, open clusters, and star-forming regions that can be observed with 4MOST. This will make it possible to study the formation, evolution, dissolution of clusters and how they populate the Milky Way, assess the role of star clusters in the evolution of the different Galactic components, and improve the determination of stellar ages by calibrating the physics involved in stellar evolution. Observing about 120 globular clusters, 1800 open clusters, and 80 very young clusters in star-forming regions, S13 will yield spectra for an extensive uniformly selected and analysed sample, covering a metallicity range between  $[\text{Fe}/\text{H}] = -2.5$  dex for some globular clusters to  $[\text{Fe}/\text{H}] > 0.0$  dex for some open clusters, with the ages ranging from just a few Myr up to 13.5 Gyr. This will lead to a full, homogeneous chemo-dynamical description of the approximately 120 globular cluster in the Milky Way and Magellanic Clouds, where for each cluster, tens of stars at  $G < 16.5$  will be observed at high resolution in order to derive abundances for elements of all nucleosynthesis channels. For most identified open clusters that can be observed with 4MOST, detailed chemical information will be obtained for all nucleosynthesis channels, with a total of  $\sim 30\,000$  stars observed with the HRS and  $\sim 75\,000$  with the LRS. The survey is thus divided into six sub-surveys, with each of the three types of cluster having both a low-resolution and a high-resolution sub-survey. The target catalogue is based primarily on the high precision astrometric data from Gaia DR3.

**2.2.1.9 The 4MOST Survey of Dwarf Galaxies and their Stellar Streams (S14)** S14/4DWARFS ([Skuladottir et al. 2023](#)) aims to characterise the chrono-chemo-dynamical properties of dwarf galaxies and stellar streams in the Galaxy. Accordingly, the stellar ages, elemental abundances, and radial velocities of around 140 000 stars will be derived, which will allow for the detailed investigation of a large range of topics, including the creation and distribution of chemical elements, dark matter halos, properties of the first stars, properties of dwarf galaxies, and the small-scale behaviour of hierarchical galaxy formation. The target catalogue will contain stars in all known dwarf galaxies and stellar streams in the Milky Way halo at  $-80^\circ < \delta < +5^\circ$  and latitude  $|b| > 20^\circ$ . As this includes diffuse stellar streams as well as large dwarf galaxies, the target density differs quite drastically from just a few targets per square degree in the former, up to more than 1000 per square degree in the central regions of the latter. The large dwarf spheroidal galaxies Sagittarius, Fornax, and Sculptor are observed both in high resolution and in low-resolution, whereas smaller and ultra-faint dwarf galaxies are only recorded by the LRS, with the selection of member stars of all dwarf galaxies being based on Gaia DR3 data. All stellar streams that have been found in the Southern hemisphere via Gaia DR3 data and the STREAMFINDER algorithm ([Malhan & Ibata 2018](#)) are targeted by both the LRS and the HRS. In total, abundances of 10–25 different elements will be derived from low- and high-resolution spectra for 100 000 dwarf galaxy stars and 30 000 stars in stellar streams, and abundances of around five elements will be derived for a further 10 000 stars at the very faint end, expanding the catalogue of stars in such systems with detailed properties available by several orders of magnitude. S14 in particular lends itself to collaboration with the halo surveys S1 and S2 due to the complementary nature of their target selection.

**2.2.1.10 4MOST Gaia RR Lyrae Survey** 4GRoundS ([Ibata et al. 2023](#)), which has been integrated into S1, is concerned with RR Lyrae variable stars and aims to determine the radial velocities and metallicities of the southern RR Lyrae stars that are present in Gaia DR3. Complementing the high-precision astrometric data of Gaia with accurate radial velocities and metallicities will make it possible to study the

structure of the halo and outer disk, perform dynamical mass modelling for the outer disk and mapping out the Galaxy’s mass out to 100 kpc, search for low-mass satellites to the Galaxy, quantify variations in the velocity coherence throughout the halo to verify cosmological models, and more. Due to Gaia DR3 data being able to provide an extremely clean all-sky sample of RR Lyrae stars, as done by [Clementini et al. \(2023\)](#), the 4GRoundS selection is to simply take this sample of  $\sim 270\,000$  stars as their catalogue, with the goal of observing – as uniformly distributed as possible – up to 100 000 of the stars that can be observed with 4MOST.

**2.2.1.11 Spectroscopic Discovery of Binaries with Dormant Black Holes** The Spectroscopic Discovery of Binaries with Dormant Black Holes survey ([Pawlak et al. 2023](#)) aims to identify tens of previously unknown dormant black hole (or neutron star) binaries via spectroscopic follow-up of photometrically selected candidates. From observations of the Magellanic Clouds performed by the Optical Gravitational Lensing Experiment (OGLE), more than 700 ellipsoidal binaries with orbital periods of less than 10 days are selected, and in addition, over 100 Wolf-Rayet stars are observed. 4MOST will record multi-epoch radial velocity measurements which, in combination with the OGLE photometry, can be used to infer the secondary component mass of the binary, and will thus allow for the discovery of systems with a compact companion. This survey has been merged with S9.

## 2.2 Extragalactic Surveys

**2.2.2.1 The eROSITA Galaxy Cluster Redshift Survey (S5)** S5’s ([Finoguenov et al. 2019](#)) goal is to complement the eROSITA mission’s ([Merloni et al. 2012](#)) X-ray observations of galaxy groups and clusters with spectroscopic redshifts, including 40 000 groups and clusters of galaxies with redshifts of up to 1.4 across a sky area of 10 000 square degrees. At  $z > 0.7$ , redshifts of the brightest cluster galaxies will be determined, and at  $z < 0.7$ , more than 15 galaxies in each cluster will be observed in order to measure dynamical masses, allowing for the calibration of clusters for cosmological experiments. Dynamical masses for around 10 000 clusters with total masses  $> 10^{14} M_{\odot}$  will be derived, which enables to make a connection between cluster observables and their total mass. Furthermore, precise distance measurements will be carried out which are necessary for estimating X-ray luminosities, which can be used as a mass proxy, and filaments are singled out spectroscopically for redshifts up to 0.2. Finally, the S5 survey will supply the S7 survey with high S/N observations of systems not observed in their survey area. The scientific requirements needed to enable these goals are the inclusion of a large observational area so that the number of spectroscopically verified groups and clusters can be significantly raised and that this coverage be uniform across large areas in order to allow for clustering analysis, which will improve and extend cosmological constraints. In addition, the sampling of target galaxies needs to be very complete and high S/N ratios are required for a subsample specifically chosen for investigations on galaxy evolution.

**2.2.2.2 The Active Galactic Nuclei Survey (S6)** S6 ([Merloni et al. 2019](#)) aims to assemble a highly complete sample of active galactic nuclei (AGN) across a significant fraction of the extragalactic sky surveyed by eROSITA in X-ray, with a completeness of 80-90% for AGN selected in the X-ray regime. The scientific goals include investigations into the evolutionary properties of the largest and most powerful supermassive black holes (SMBH), the prevalence of accreting SMBHs in a galaxy population as well as the accretion rate distribution depending on redshift, luminosity, and the host galaxy properties, spectral observations of SMBHs and their host galaxy enabling direct determinations of some SMBH masses, AGN activity in relation to its environment on a large scale, and finally to contribute to fundamental cosmological constraints, for instance the usage of quasars as standard candles. The target selection utilises X-ray data from eROSITA, following up on all point-like extragalactic sources, and in addition makes use of mid-infrared data from WISE for further selection. Up to one million SMBH’s physical properties such as redshift, luminosity, and mass will be derived across a sky area of around 10 000 square degrees and their evolutionary and clustering properties will be analysed, with links between large-scale

structures and AGN being investigated for redshifts up to  $z \sim 6$ . The goals of the survey require a high level of completeness and uniformity in redshift data across a large area, and to determine fairly precise redshifts down to very faint magnitudes of  $r_{\text{AB}} = 23$  mag. Furthermore, the accuracy of the relative flux calibration has to be better than 10% to identify physical features, in particular for high S/N spectra from sources at the brighter end.

**2.2.2.3 The Wide-Area VISTA Extragalactic Survey (S7)** S7/WAVES's ([Driver et al. 2019](#)) purpose is to investigate energy and structure growth across a large range of scales from  $\sim 1$  kpc up to  $\sim 10$  Mpc for a timeline of 7 Gyr. On small scales, determinations of distances and environmental characteristics produced by S7 will be used in combination with existing high-resolution data to infer the bulges', disks', and bars' masses, and size evolutions of galaxies. On intermediate scales, the masses and sizes of galaxy groups are the focus in addition to galaxy merger rates, which allows for direct explorations into the build-up of dark matter halos and stellar mass. On large scales, groups, filaments, but also voids are investigated, in particular concerning their recent emergence. The underlying goal of this survey is identifying and discerning the processes responsible for the evolution of gas to stars, and disentangling how they are related to the host galaxy's halo characteristics as well as the environment on a larger scale. S7 will produce a sample of about 1.6 million galaxies ranging from low-redshift galaxies at  $Z \lesssim 21.1$  mag and photometric redshift  $z \lesssim 0.2$  within the scope of the WAVES-Wide sub-survey, to moderate-redshift galaxies at  $Z \lesssim 21.25$  mag and photometric redshift  $z \lesssim 0.8$  within the scope of the WAVES-Deep sub-survey, with the two sub-surveys containing around 0.9 and 0.75 million galaxies, respectively. The requirements of this survey include a depth of 21.25 mag in the  $Z$  band, spectroscopic observations to be complete to a degree of more than 90%, and adequately-sized continuous areas at all redshifts, with the requisite size being more than 1200 square degrees at  $z \lesssim 0.2$ , more than 50 square degrees at  $z \sim 0.5$ , and four areas with more than 4 square degrees at  $z \sim 0.8$ .

**2.2.2.4 The Cosmology Redshift Survey (S8)** S8/CRS ([Richard et al. 2019](#)) is meant to complement lensing, cosmic microwave background (CMB), and other surveys, carrying out extensive cosmological tests using spectroscopic clustering analyses. By examining samples of bright galaxies (BGs), luminous red galaxies (LRGs), emission-line galaxies (ELGs), and quasars (QSOs), and using their combined data, decisive tests of gravitational physics may be carried out. The survey's specific science goals include such testing via overlapping lensing and spectroscopy, calibrating the source redshift distribution by means of cross-correlation, and producing growth-rate measurements by combining the obtained data with observed CMB data. The survey will lead to the creation of a redshift-space map showcasing large-scale structures by incorporating around 8 million objects from a redshift range between  $z = 0.15$  and  $z = 3.5$  from an area covering 7500 square degrees, where BGs, LRGs and QSOs are distributed over this whole area, whereas ELGs are observed specifically in a 1000 square degree sub-area with the best observational quality. The area was chosen by incorporating the Dark Energy Survey, Kilo-Degree Survey and VST-ATLAS areas, but excluding the DESI areas. Besides a sufficiently large sky coverage, the S8 survey also requires target densities for each of the categories (BG, LRG, ELG, QSO) to be such that the calculations will not be Poisson-noise-limited, and the spectroscopic success rate must be over 95%, over 75%, over 80%, and over 50% for BGs, LRGs, ELGs, and QSOs, respectively.

**2.2.2.5 The Time-Domain Extragalactic Survey (S10)** S10/TiDES ([Swann et al. 2019](#)) observes extragalactic optical transients as well as variable sources found from upcoming large sky surveys, providing spectroscopic follow-up. There are three sub-surveys: TiDES-SN is concerned with the spectroscopic observations of supernova-like transients and their classification, providing spectra for 30 000 live transients up to redshifts of  $z \sim 0.5$ . TiDES-Hosts spectroscopically observes transient host galaxies and provides redshifts for 50 000 such galaxies up to redshifts of  $z \sim 1$  to be applied in cosmological analyses. TiDES-RM keeps up with the reverberation of around 700 AGN via repeated spectroscopic observations up to redshifts of  $z \sim 2.5$ . Most of the targets will stem from the Large Synoptic Survey Telescope's (LSST) survey, with S10 targeting all live transients present in a particular 4MOST field

which is brighter than  $r_{\text{AB}} = 22.5$  mag. The TiDES-SN sub-survey requires that the time between the discovery of a transient and the spectroscopic observation thereof with 4MOST is at most three to four days, whereas the TiDES-Hosts sub-survey simply requires that the current 4MOST field contains supernovae discovered prior by LSST. The S10 survey will have around 30 LRS fibres – which is around 2% of all – assigned to it in each pointing, but will not be considered in the survey strategy deciding the pointing of the instrument, rather utilising whichever relevant targets can be found in each individual pointing.

**2.2.2.6 Extragalactic Community Surveys** In addition to the five Extragalactic Consortium Surveys, there are another nine Extragalactic Community Surveys. CHANCES, the CHileAN Cluster galaxy Evolution Survey ([Haines et al. 2023](#)), will target several hundred thousand galaxies in around 150 massive galaxy clusters out to  $z \sim 0.45$ , and includes massive and dwarf galaxies from the outer filaments to the cores of galaxy clusters. It will study the evolution of galaxy environment and the influence thereof on galaxy evolution itself. The Chilean AGN/Galaxy Extragalactic Survey (ChANGES) ([Bauer et al. 2023](#)) will focus on a sample of millions of AGN to assist in studying black hole accretion rates and black hole mass densities, evolution and host synergies, variations in AGN spectra, tidal disruption events, and more. The Transform our Understanding of the Baryon Cycle with High-Resolution Quasar Spectroscopy (ByCycle) survey ([Peroux et al. 2023](#)) is concerned with the study of the baryons making up the intergalactic gas that cannot be directly observed via starlight, but make up the majority of baryonic matter. The aim of this survey is to gain an understanding of the processes by which the gas makes its way into, through, and out of galaxies. The 4MOST Hemisphere Survey of the Nearby Universe (4HS) ([Taylor et al. 2023](#)) records spectra and derives redshifts for about six million galaxies with high completeness at  $z < 0.15$ . The scientific goals of this survey include to map out the local southern Universe regarding its mass, motion, gravitational growth, and demographic trends, and to build up a benchmark local galaxy reference sample for the future.

The Stellar Population Survey Using 4MOST (4MOST-StePS) ([Iovino et al. 2023](#)) will utilise the spectra of about 3300 galaxies in order to empirically obtain an evolutionary path of massive galaxies in the redshift range  $0.3 < z < 0.7$ . 4MOST-StePS is complementary to WAVES and will be used in combination with WAVES data to study the connection between galaxy properties and environment. The Optical, Radio Continuum and HI Deep Spectroscopic Survey (ORCHIDSS) ([Duncan et al. 2023](#)) will provide spectroscopy on neutral atomic hydrogen and molecular hydrogen to investigate the evolution of neutral gas across most of cosmic history, in particular with regards to feedback processes and the lifecycle of gas in galaxies. The 4MOST Complete Calibration of the Colour-Redshift Relation (4C3R2) survey ([Gruen et al. 2023](#)) aims to use spectroscopic data and redshifts measured by 4MOST in order to, in combination with other programmes, calibrate the colour-redshift relation. Since 4MOST enables to probe the complete redshift distribution given colour, it is very suitable to achieve this goal. The 4MOST–Gaia Purely Astrometric Quasar Survey (4G-PAQS) ([Krogager et al. 2023](#)) is designed as a colour-independent quasar survey, the target catalogue of which is selected based only on astrometric data from Gaia. Being the first large-scale endeavour of this kind, its goal is to uncover and quantify the biases and selection effects within colour-selected samples. The 4MOST Strong Lensing Spectroscopic Legacy Survey (4SLSLS) ([Collett et al. 2023](#)) is a follow-up survey observing strong lens candidates found in the Euclid survey and LSST. Redshifts will be derived for 10 000 pairs of strong-lensing galaxies and background galaxies where the lensed light originates, and velocity dispersions will be derived for 5000 pairs, allowing for the study of galactic evolution and the cosmological model.

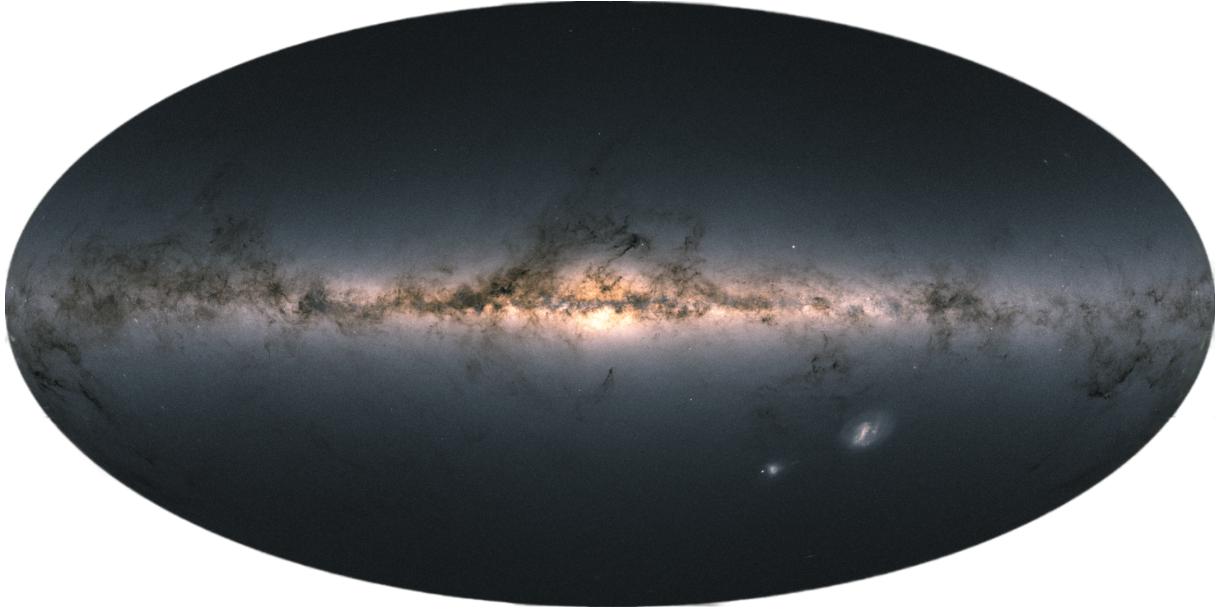


Figure 11: Image of the Milky Way in colour: map of the integrated fluxes observed in Gaia’s three passbands in Gaia (E)DR3 (Credit: ESA/Gaia/DPAC).

### 3 Selection of input catalogue

#### 3.1 The Gaia mission

The Gaia mission forms a core aspect of the European Space Agency’s (ESA) science programme, and its main objective is to provide high-precision astrometric and spectrophotometric information on more than a billion celestial objects, including half a million quasars, a few hundred thousand asteroids, as well as a few million galaxies (Lindegren 2010). Although the concept had been developed since the 1990s, the implementation phase did not commence until 2006, with the Gaia spacecraft launching in 2013 and science operations beginning in the summer of 2014. The accuracy, sensitivity, dynamic range, and sky coverage that are reached by the instrument are unattainable from the ground for photometric and spectroscopic surveys with a scientific scope of the same magnitude as Gaia. In particular, in order to achieve global astrometry on a microarcsecond scale of accuracy, a survey must be conducted from space. Due to this, the science case laid out by Perryman et al. (2001) remained relevant more than a decade later at the start of science operations, and is still valid today, as the highly accurate three-dimensional spatial and velocity distributions of stars obtained by Gaia, aided also by the additional astrophysical parameters contained within the Gaia data releases, are crucial for the understanding of the structure, history and future of our Galaxy. The list of scientific topics that can be addressed is thus really extensive, and broadly speaking it includes studies of the structure, dynamics, evolution, and star formation history of the Galaxy, stellar physics and evolution, variability studies and distance scales, binary and multiple star systems, as well as extragalactic topics for objects in the Local Group and quasars. The Gaia survey does not use an input catalogue in order to remain unbiased in its detections (Gaia Collaboration et al. 2016).

The Gaia spacecraft is located at the second Lagrange point ( $L_2$ ) of the Sun-Earth-Moon system and moves around  $L_2$  in a Lissajous orbit. The scanning law is chosen such that the uniformity of the sky coverage is maximised, and it is defined by the spin rate around its spin axis of  $\omega = 60''\text{s}^{-1}$ , the solar-aspect angle of  $\xi = 45^\circ$ , and the precession of the spin axis with a precession period of 63 days. The astrometric measurements are carried out using *scanning space astrometry*, wherein positional information is converted into timing data. An in-depth characterisation of this method and its implementation in Gaia can be found in (Lindegren & Bastian 2010). The payload module of the spacecraft (Holl 2012) contains the focal plane assembly and two identical three-mirror anastigmatic telescopes that are guided

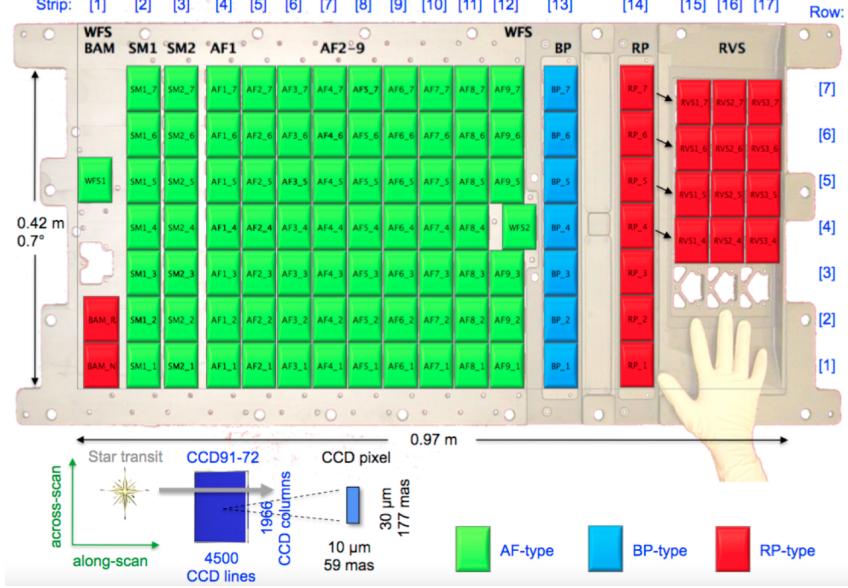


Figure 12: Gaia’s focal plane assembly ([Gaia Collaboration et al. 2016](#)).

into a common path by another set of mirrors. The incoming beams are then folded twice by two additional reflectors to accommodate the focal length of 35 m ([Gaia Collaboration et al. 2016](#)).

The focal plane assembly consists of 106 CCD detectors totaling 938 million pixels in total in order to register the incoming light. There are three types of CCDs, indicated by colour in Figure 12: the default broad-band CCD, the blue-enhanced CCD for short wavelengths, and the red-enhanced CCD for long wavelengths, and each individual CCD is made up of 4500 by 1966 pixels. The astrometric measurements are carried out by first detecting objects via the 7 + 7 Sky Mapper CCDs (SM1 & SM2), and then measuring the astrometric properties using the area of 62 CCDs making up the astrometric field. The focal plane also contains two wave-front sensors of Shack-Hartmann type (WFS1/WFS2), a basic angle monitor (BAM), seven blue CCDs (BP) and seven red CCDs (RP) for photometric measurements, and 12 red CCDs which perform spectroscopy of bright stars (Jordi et al. 2010).

The aforementioned three types of CCDs provide Gaia with three different passbands: the  $G$  band with a wavelength range from 330 to 1050 nm, the  $G_{\text{BP}}$  band with a wavelength range from 330 to 680 nm, and the  $G_{\text{RP}}$  band with a wavelength range from 640 to 1050 nm. Whereas  $G$  band measurements are performed by the CCDs that also carry out the astrometric measurements,  $G_{\text{BP}}$  and  $G_{\text{RP}}$  band measurements are performed using the CCDs specifically dedicated to blue and red photometry (Evans et al. 2018). A fused-silica prism is located in front of each of the two strips of photometric CCDs, and coating is applied to these two prisms to obtain the respective wavelength ranges of the  $G_{\text{BP}}$  and the  $G_{\text{RP}}$  bands. The final form of the passbands is determined by the coating, the telescope transmission as well as detector quantum efficiency, and is shown in Figure 13. While the maximum object density that can be handled by the astrometric detectors is 1 050 000 objects per square degree, the photometric detectors can observe all objects up to a density of 750 000 objects per square degree, beyond which only the brightest sources are measured ([Gaia Collaboration et al. 2016](#)).

The Early Data Release 3 (EDR3) of the Gaia mission ([Gaia Collaboration et al. 2021](#)) included only some basic astrometric and photometric information. The astrometric parameters given are the right ascension, declination, parallax, and proper motions for around 1.5 billion objects as well as a further 400 million objects with only the right ascension and declination. Figure 14 shows how many sources have two-parameter solutions (right ascension  $\alpha$  and declination  $\delta$ ), five-parameter solutions ( $\delta, \alpha, \text{parallax } \pi$ , and proper motion  $\mu_\alpha$  and  $\mu_\delta$ ), or six-parameter solutions ( $\delta, \alpha, \pi, \mu_\alpha, \mu_\delta$ , and pseudocolour) as a

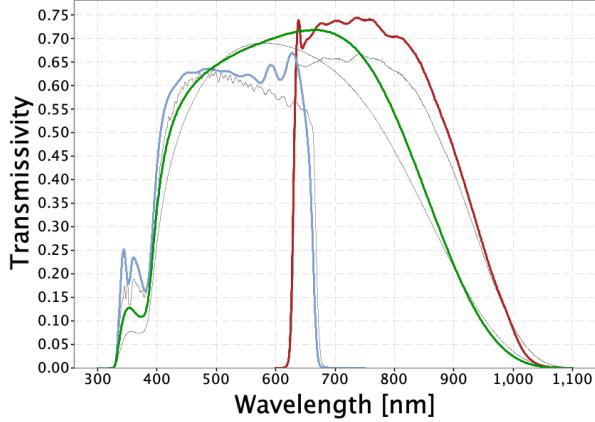


Figure 13: Gaia passbands in (E)DR3 for  $G$  (green),  $G_{\text{BP}}$  (blue), and  $G_{\text{RP}}$  (red) (Riello et al. 2021). The grey lines are the nominal, pre-launch passbands by Jordi et al. (2010).

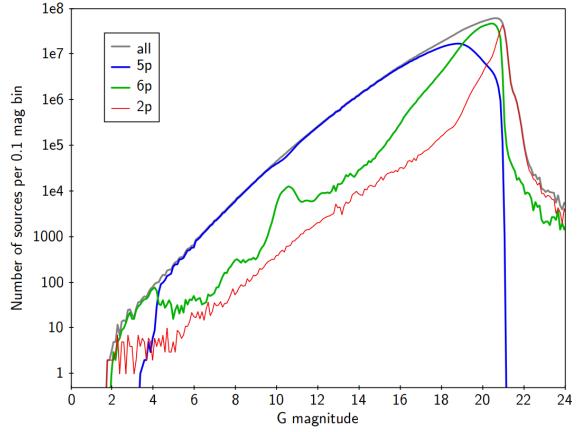


Figure 14: Magnitude distribution of sources separated according to two-, five-, and six-parameter solution (Lindegren et al. 2021b).

function of the Gaia magnitude  $G$ . This figure also showcases Gaia’s limiting magnitude at approximately  $G \simeq 21$ . The photometric parameters given are the broad-band magnitudes  $G$  for all objects, and  $G_{\text{BP}}$  and  $G_{\text{RP}}$  for about 1.5 billion objects each. The full Gaia DR3 catalogue (Gaia Collaboration et al. 2023c), released in June 2022, extends Gaia EDR3 by adding a plethora of further parameters, mostly from analysis of the spectra of the BP/RP photometers and the Radial Velocity Spectrometer. In particular, astrophysical parameters such as the effective temperature  $T_{\text{eff}}$ , surface gravity  $\log g$ , metallicity [M/H],  $G$  band extinction  $A_G$ , distance, etc. are derived for almost half a billion objects from BP/RP spectra recorded by the BP/RP instrument, and for about 33 million objects, RVS-derived radial velocities are available. In addition, more than 200 million time-averaged low-resolution BP/RP spectra are included in the form of 110 coefficients to be used in a set of basis functions along with around 1 million mean RVS spectra. Further data published in Gaia DR3 are QSOs, galaxies, and Solar System objects.

For objects with only positions, the median positional uncertainties are 1-3 mas. For the objects which have full astrometric solutions, the median uncertainties for positions, parallax, and proper motion are 0.01-0.02 mas, 0.02-0.03 mas, and 0.02-0.03 mas per year at  $G < 15$  mag, 0.05 mas, 0.07 mas, and 0.07 mas per year at  $G = 17$  mag, 0.4 mas, 0.5 mas, and 0.5 mas per year at  $G = 20$  mag, and 1.0 mas, 1.3 mas, and 1.4 mas per year at  $G = 21$  mag. There is a global systematic parallax bias of  $-17 \mu\text{as}$ , which is discussed in detail by Lindegren et al. (2021a), who also provide a recipe for the zero-point correction of the parallax. The source position and proper motion of objects are given within the celestial frame defined from more than 1.5 million QSO-like objects, the Gaia-CRF3 system, which is aligned with the ICRF3 to about 0.01 mas root mean square, and is non-rotating according to quasar measurements to within 0.005 mas per year root mean square at epoch J2016.0.

As for the photometry, the uncertainties in the  $G$  band,  $G_{\text{BP}}$  band, and  $G_{\text{RP}}$  band are  $\sim 0.3$  mmag,  $\sim 0.9$  mmag, and  $\sim 0.6$  mmag at  $G < 13$  mag, 1 mmag, 12 mmag, and 6 mmag at  $G = 17$  mag, and 6 mmag, 108 mmag, and 52 mmag at  $G = 20$  mag, respectively.<sup>6</sup> In Figure 15, the distribution of errors of the astrometric parameters as well as the radial velocities is shown across galactic coordinates, showcasing the overall very high accuracy of the Gaia data, although it should be noted that for the right ascension, declination, and parallax, this figure shows only the sources that have five-parameter solutions, and the sources with six-parameter solutions tend to be the ones with higher median errors, as shown by the other plots in the Gaia DR3 Documentation, Chapter 17.2. The large-scale structures that are visible in these plots are a result of the scanning law of the observations.

<sup>6</sup><https://www.cosmos.esa.int/web/gaia/dr3>, retrieved 22.10.2023

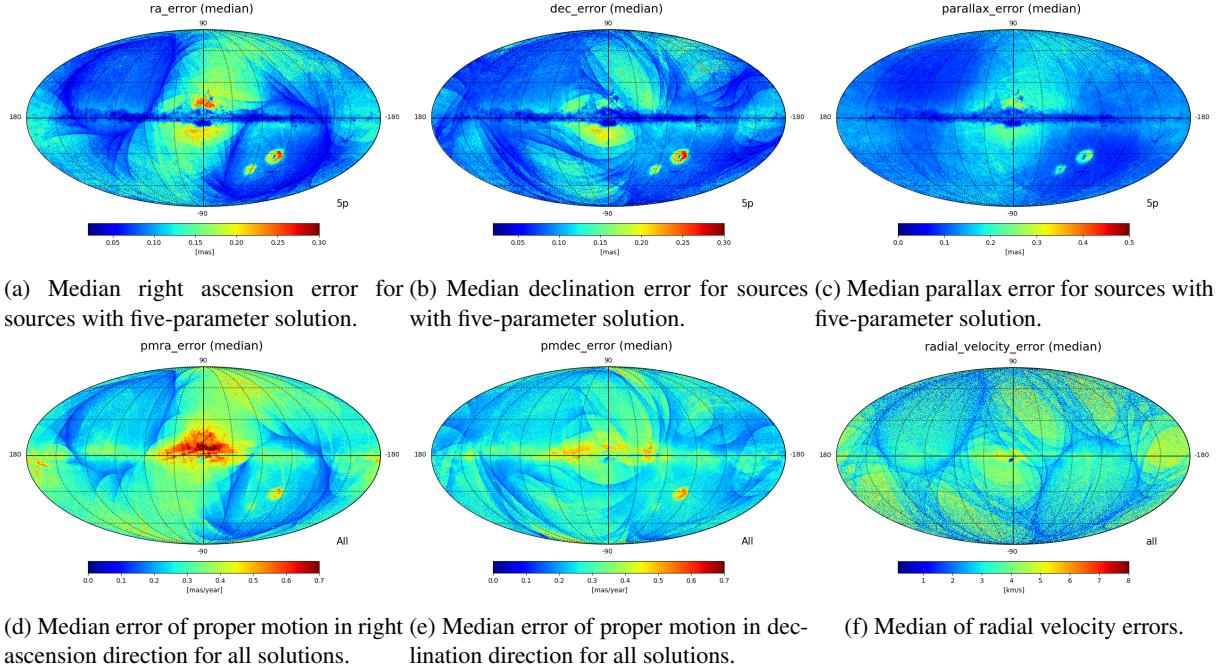


Figure 15: HEALPix maps of the median errors of positions, parallaxes, proper motions, and radial velocities ([Gaia DR3 Documentation: 17.2. Catalogue plots](#)).

In order to derive a variety of astrophysical data products that are included in Gaia DR3, the Astrophysical parameters inference system (APSIS) was developed ([Creevey et al. 2023](#); [Bailer-Jones et al. 2013](#)). This includes source classifications, a 2D total Galactic extinction map, redshifts of galaxy candidates ([Delchambre et al. 2023](#)), as well as stellar  $T_{\text{eff}}$ ,  $\log g$ , [M/H], distances, radii, masses, ages, chemical abundances, and more for varying numbers of objects ([Fouesneau et al. 2023](#)). The Apsis pipeline is made up of 13 modules producing in total 43 astrophysical parameters from the observed photometry, BP/RP spectra, RVS spectra, parallaxes, and proper motions of the objects. Among these are the General Stellar Parametrizer from Photometry (GSP-Phot) and the General Stellar Parametrizer from Spectroscopy (GSP-Spec).

GSP-Phot is a set of algorithms which aims to determine stellar parameters such as  $T_{\text{eff}}$ ,  $\log g$ , [M/H] and others for single stars using their low-resolution BP/RP spectra in tandem with their astrometry and photometry ([Andrae et al. 2023a](#)). The main algorithm utilises the BP/RP spectra, parallax, and  $G$  band magnitude to fit the four parameters stellar age, mass, metallicity, and the line-of-sight extinction  $A_0$  via a Markov chain Monte Carlo method (see q-method in [Bailer-Jones 2011](#)). The initial guesses for these parameters are determined from two additional algorithms, one based on machine learning ([Geurts et al. 2006](#)) and the other based on the gradient-descent method ([Bailer-Jones 2010](#)). PARSEC stellar isochrones are used in a very fine grid to infer the absolute magnitude  $M_G$ , stellar radius  $R$ , effective temperature  $T_{\text{eff}}$ , and surface gravity  $\log g$  from the fit parameters age, initial mass, and [M/H]. From model BP/RP spectra calculated based on  $A_0$  and [M/H] as well as  $T_{\text{eff}}$  and  $\log g$  from the isochrones, the distance can be directly inferred, as the amplitude  $R^2/d^2$  which is required to scale the model spectrum to the observed BP/RP spectrum contains, aside from  $d$ , only the stellar radius  $R$ , which is provided by the isochrones. As  $M_G$  and  $A_0$  are known,  $A_G$  can be derived as well due to the tight constraint given by the accurately observed apparent magnitude  $G$ , and  $A_{\text{BP}}$  as well as  $A_{\text{RP}}$  can be determined analogously to  $A_G$ , which finally yields the reddening  $E(G_{\text{BP}} - G_{\text{RP}}) = A_{\text{BP}} - A_{\text{RP}}$ . The photometric distances determined by GSP-Phot for objects further away than  $\sim 3$  kpc turn out to be systematically underestimated, which may be due to the choice of distance prior. Still, for very accurate parallax measurements,

the distances are more dependable up to 10 kpc. The reddening  $E(G_{\text{BP}} - G_{\text{RP}})$  is validated, for instance, by (Gaia Collaboration et al. 2023b) through comparison to Gaia RVS data of diffuse interstellar bands.

Using the RVS spectra captured by the Gaia satellite at  $R \sim 11500$ , GSP-Spec provides a number of chemo-physical parameters including the stellar atmospheric parameters as well as elemental abundances of N, Mg, Si, S, Ca, Ti, Cr, FeI, FeII, Ni, Zr, Ce, and Nd (Recio-Blanco et al. 2023). The overall metallicity [M/H] and the alpha-element abundance [ $\alpha/\text{Fe}$ ] is estimated as well. The estimates are based solely on the spectroscopic observations of the RVS instrument without input regarding astrometry or photometry. Utilising MARCS atmosphere models (Gustafsson et al. 2008) and the TURBOSPECTRUM code (Plez 2012) to derive extensive grids of stellar spectra, these are then used to train the local multi-linear regression algorithm MATISSE as well as being a reference model for GAUGUIN, which is a local optimisation procedure based on the Gauss-Newton algorithm (Recio-Blanco et al. 2016). The analysis is carried out by a combination of the DEGAS decision-tree method, the MATISSE algorithm and the GAUGUIN algorithm to derive the stellar atmospheric parameters iteratively, and then using just GAUGUIN for the derivation of the individual elemental abundances. An extensive validation produces a detailed quality flag scheme accompanying the final data.

Whereas the validation of the astrometric and photometric content of Gaia DR3 was already performed within the scope of Gaia EDR3 (Fabricius et al. 2021), the validation of the new products released with the Gaia DR3 catalogue was carried out by Babusiaux et al. (2023) using statistical analyses, checks for internal consistency, as well as external data and models in order to point out limitations and caveats when using Gaia DR3. Extinction and reddening are computed as described in Andrae et al. (2023a) and mentioned above, but Gaia DR3 contains two more extinction estimates as well. Effects such as the temperature-extinction degeneracy, the demand for a positive extinction in low-extinction areas, as well as the extinction being overvalued in comparison to BAYESTAR19 (Green et al. 2019) are all discussed by Andrae et al. (2023a), and the GSP-Phot extinctions should be used over the others. The GSP-Phot effective temperatures generally show a good agreement with external samples, but around the library borders of the spectral libraries used in the analysis, there is a greater density of objects, which is not physical. The GSP-Phot metallicity estimates in particular are rather uncertain, as a zero-point disparity of -0.55 is found when comparing to literature values. The GSP-Phot distances as well can be rather unreliable, with a comparison to values derived from taking the inverse parallax for objects with small parallax errors showing that 37% of estimates are outside the  $5\sigma$  range. The overestimation of distances for objects that are far away is discussed in Andrae et al. (2023a). Since Gaia DR3 also contains quality flags for many parameters, it is recommended to use these in order to make the best possible use of Gaia's wealth of groundbreaking data.

### 3.2 The Pristine survey

Making use of narrow-band photometry, Starkenburg et al. (2017) focus on the Ca H&K lines in order to derive metallicities of halo stars, with some of the aims being to make progress in the search for metal-poor stars, in particular within the Galactic halo, and the characterisation of faint dwarf galaxies. Not only can objects be observed down to fainter levels than with spectroscopic methods, it is also more efficient observationally. The method relies on the fact that the strength of the Ca H&K absorption lines correlates with the stars' metallicity at a given temperature or colour. This is shown in Figure 16, where the Ca H&K lines in synthetic spectra computed with MARCS stellar atmospheres (Gustafsson et al. 2008) and the TURBOSPECTRUM code (Plez 2012) can be seen for different stellar parameters. In particular, the line strength is indicated by the different colours, starting with red at  $[\text{Fe}/\text{H}] = 0.0$  and going to blue at  $[\text{Fe}/\text{H}] = -3.0$ , with the black solid line indicating a star with no metals. By examining the relative fluxes of a narrow band around the Ca H&K lines at comparable broad-band colours, the metallicity can be derived.

The Pristine survey is conducted in the Northern hemisphere on the 3.6-metre Canada-France-Hawaii

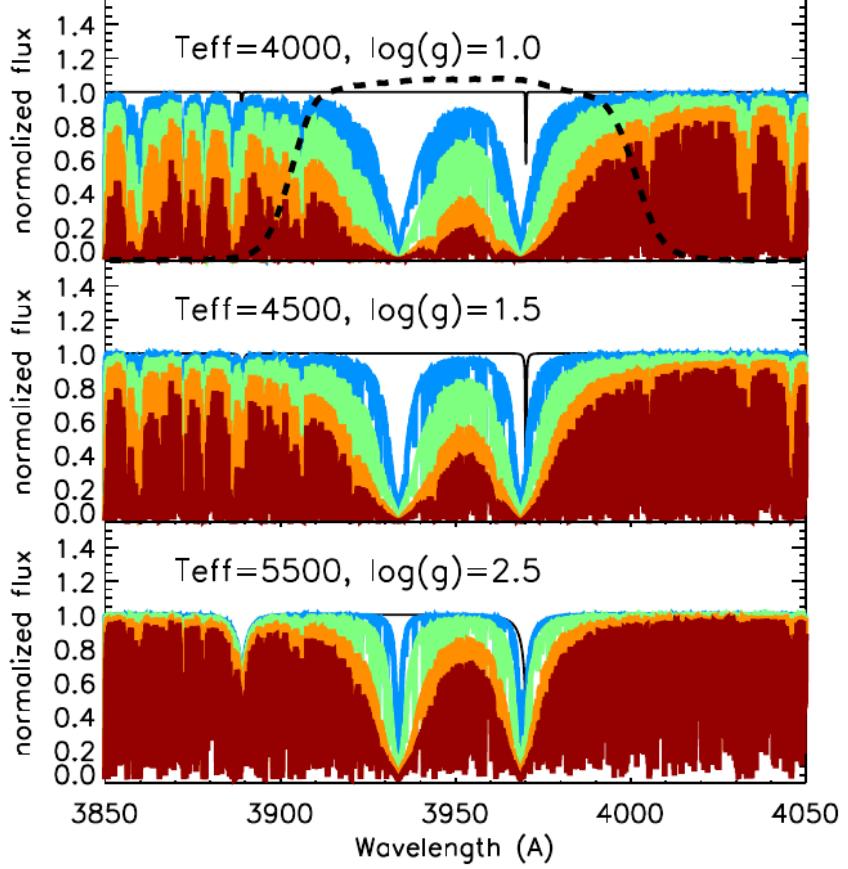


Figure 16: Synthetic spectra around the Ca H&K lines for different stellar parameters, where the metallicities are  $[Fe/H] = 0.0$  for the red part,  $[Fe/H] = -1.0$  for the orange part,  $[Fe/H] = -2.0$  for the green part, and  $[Fe/H] = -3.0$  for the blue part. The solid black line indicates a spectrum for a star without any metals, whereas the dashed line shows the Ca H&K narrow-band filter used in the Pristine survey (Starkenburg et al. 2017).

Telescope's MegaCam, covering more than  $3000 \text{ deg}^2$  of the Galactic halo down to a Sloan Digital Sky Survey (SDSS) magnitude of  $g \sim 21.0$  at  $S/N \sim 10$ . The observations utilise a Ca H&K filter, the throughput of which can be seen in the top panel of Figure 16, and the narrow-band photometry from these observations is used together with SDSS photometry (e.g., York et al. 2000). This filter has a width of about  $100 \text{ Å}$  in order to account for radial velocities of stars shifting the Ca H&K lines from their rest wavelengths. Owing to a lack of good photometric standard stars, the individual fields have to be calibrated internally, with the reference field chosen from the observations themselves. The calibration involves calculating the zero-point offset for each field with regards to the reference field, and then making a flat-field correction due to the magnitudes being position-dependent on the field of view.

For around 17500 stars observed, there are spectra available from the SDSS/SEGUE catalogue, published with their associated radial velocities and metallicities (Yanny et al. 2009; Eisenstein et al. 2011). In order to only include data of high quality, several selection criteria are employed, leaving 7673 stars for the calibration of the Pristine metallicity scale. Contaminants such as variable stars, quasars, and white dwarfs make up about 7% of the whole sample and are consequently removed. To derive metallicities from the photometric data, all pixels in the  $[CaHK - g - 1.5 \cdot (g - i)]_0 - (g - i)_0$  space (each sized 0.025) are given a metallicity value based on SDSS/SEGUE data and Pristine synthetic spectra, where the metallicity estimate for the SDSS/SEGUE spectra is derived using neural networks trained on real spectra (see Lee et al. 2008). In case a pixel has no associated metallicity, the metallicity is interpolated from nearby pixels, and in the sparsely populated range where the synthetic spectra predict no more metal content, any pixel is given a metallicity of  $[Fe/H] = -4.0$ . Figure 17 shows the colour-

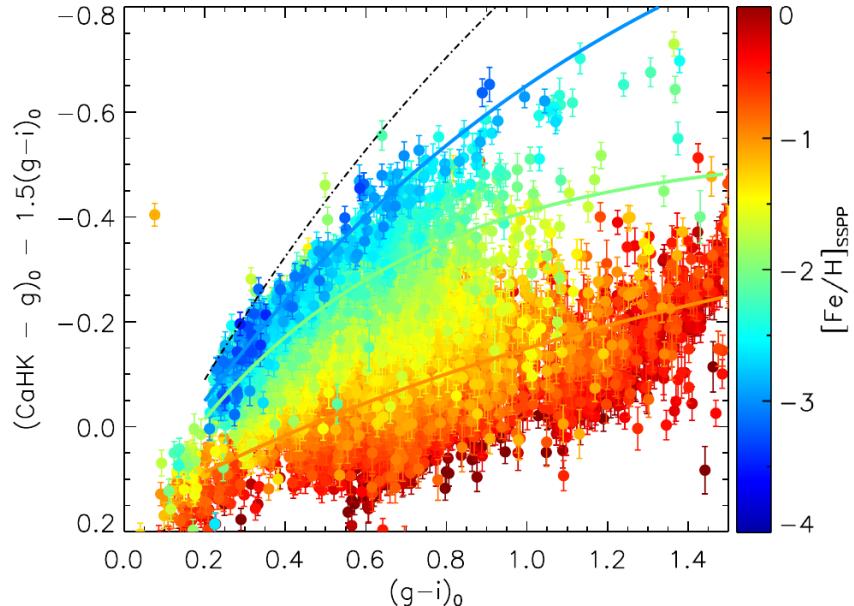


Figure 17: SDSS-Pristine colour-colour space with colour-coding according to SDSS stellar parameter pipeline (SSPP) metallicities. The coloured lines indicate synthetic spectra model predictions in the case of  $[Fe/H] = -1.0$ ,  $[Fe/H] = -2.0$ ,  $[Fe/H] = -3.0$ , and no metal lines for the orange, green, blue, and black line, respectively (Starkenburg et al. 2017).

colour space, with the metallicity from the SDSS stellar parameter pipeline (SSPP) colour-coded and the coloured lines indicating synthetic spectra model predictions for different metallicities. Lastly, in order to achieve less noise, a 2D Gaussian is used to smooth the grid. As shown in Starkenburg et al. (2017), comparisons to other spectroscopic data sets such as LAMOST, APOGEE, and some high resolution studies at low metallicities show that the Pristine metallicities agree well with existing studies down to very low metallicities. The Pristine survey demonstrates that using narrow-band photometry around the Ca H&K lines together with broad-band photometry yields very good metallicity results.

The low-resolution BP spectra released with Gaia DR3 contain the wavelength range in which the Ca H&K lines lie, making them well-suited for deriving metallicity estimates following the method laid out in Starkenburg et al. (2017), even in the extremely metal-poor regime. As demonstrated by Gaia Collaboration et al. (2023a), the BP/RP data can be used to calculate synthetic photometry for any other filter which is within the wavelength range 330 to 1050 nm covered by the red and blue photometers, and the synthetic Pristine  $CaHK_{syn}$  magnitudes agree well with observed magnitudes of the Pristine survey, as shown in their Figure 32. Thus, Martin et al. (2023) set out to replace the relative calibration based on SDSS photometry by calibrating on the absolute scale of the  $CaHK_{syn}$  magnitudes, and to construct a catalogue of synthetic  $CaHK_{syn}$  from Gaia data across the whole sky. For this, Martin et al. (2023) derive synthetic  $CaHK$  magnitudes for the full set of 219 million spectra for which BP/RP spectra coefficients are provided using GAIAXPY<sup>7</sup>. As described by Gaia Collaboration et al. (2023a), this is done by integrating the product between the spectral energy distribution (BP/RP spectra) with the transmission curve of the Pristine (MegaCam CaHK) filter<sup>8</sup>.

The main application of the newly calibrated photometry and the photometric catalogue by Martin et al. (2023) is the derivation of photometric metallicities following the Pristine model ( $CaHK, G, G_{BP}, G_{RP}$ )  $\rightarrow [Fe/H]_{phot}$ , where only the stars are included for which the synthetic magnitude uncertainties are  $\sigma_{CaHK_{syn}} < 0.1$ , which leaves mostly the brighter end of the sample up to  $G_{BP} \sim 15$  to 17. Due to

<sup>7</sup><https://pypi.org/project/GaiaXPY/>, retrieved 22.10.2023

<sup>8</sup><https://www.cfht.hawaii.edu/Instruments/Filters/megaprime.html>, retrieved 22.10.2023

Gaia's scanning law, some stars are observed more often than others, leading to a non-uniform distribution of the stars whose spectra were published in Gaia DR3, and thus of the stars which are in the Pristine-Gaia synthetic metallicity catalogue.

The Pristine method relies on correcting the Gaia broad-band photometry for extinction effects, the extent of which also depends on the properties of the star ( $T_{\text{eff}}$ ,  $\log g$ ,  $[Fe/H]$ ) besides just the extinction towards that star. [Martin et al. \(2023\)](#) derive a relation for the extinction coefficients  $k_f$  in filter  $f$  based on the relation provided by the Gaia team<sup>9</sup>, but designed to include a dependence on  $[Fe/H]$ . This is done using the DUSTAPPROX package ([Fouesneau et al. 2022](#)) with synthetic spectra by [Castelli & Kurucz \(2003\)](#) and the [Fitzpatrick \(1999\)](#) extinction law, considering the [Riello et al. \(2021\)](#) Gaia passbands as well as the Pristine CaHK filter passband.  $k_f$ , which is a polynomial relation in  $A_0$ ,  $T = T_{\text{eff}}/5040 \text{ K}$ , and  $[Fe/H]$ , is fit to synthetic photometry  $0.01 \leq A_0 \leq 2.99$  calculated by DUSTAPPROX on a synthetic grid in  $T_{\text{eff}}$ ,  $\log g$ , and  $[Fe/H]$ . As the [Schlegel et al. \(1998\)](#) extinction map is used,  $E(B - V)$  is given rather than  $A_0$ , so  $k_f$  is also determined for  $A_0/A_V$ , and finally the extinctions are given as  $A_f = k_f \frac{A_0}{A_V} \cdot 3.1E(B - V)$  for filter  $f$ . The best extinction coefficients,  $T_{\text{eff}}$ , and  $[Fe/H]$  are determined iteratively for each star, with  $T_{\text{eff}}$  being calculated according to the [Casagrande et al. \(2021\)](#) relation of  $T_{\text{eff}}$  as a function of  $(G_{\text{BP}} - G_{\text{RP}})_0$ ,  $[Fe/H]$ , and  $\log g$ .

The method used to derive photometric metallicities from Gaia photometry is very similar to the original method ([Starkenburg et al. 2017](#)), but replaces the SDSS photometry by  $(G, G_{\text{BP}}, G_{\text{RP}})$ . Its use is only viable for FGK stars, as the strength of the Ca H&K absorption lines decreases for hotter stars to the point where they can no longer provide reliable estimates for the metallicity, and at the cool end, stars contain strong molecular bands that decrease the level of the pseudo-continuum too much. For this reason, only the colour range  $0.5 < (G_{\text{BP}} - G_{\text{RP}})_0 < 1.5$  is analysed, which is a temperature range of approximately  $3900 < T_{\text{eff}} < 7000 \text{ K}$ . The training sample used by [Martin et al. \(2023\)](#) to map from the extinction-corrected  $(CaHK, G, G_{\text{BP}}, G_{\text{RP}})$  colour space to the photometric metallicity  $[Fe/H]_{\text{phot}}$  is built up from SDSS/SEGUE ([Yanny et al. 2009; Smee et al. 2013](#)), complemented by known metal-poor stars, APOGEE at  $\log g < 3.9$  and  $T_{\text{eff}} < 5800 \text{ K}$  ([Majewski et al. 2017; Abdurro'uf et al. 2022; Wilson et al. 2019](#)), the PASTEL sample ([Soubiran et al. 2016](#)), and finally targets followed up from the Pristine survey that have been confirmed to be metal-poor. Using MARCS atmospheres ([Gustafsson et al. 2008](#)) alongside the TURBOSPECTRUM code ([Plez 2012](#)), a grid of stellar spectra is added at  $[Fe/H] = -3.0$  and  $[\alpha/Fe] = 0.4$ , to improve the sampling in the extremely metal-poor regime.

The metallicity estimates are derived from colour-colour space including de-reddened Gaia colour  $(G_{\text{BP}} - G_{\text{RP}})_0$  as a temperature proxy, as well as the Gaia broad-band photometry combined with the Pristine narrow-band photometry  $(CaHK - G)_0 - 2.5(G_{\text{BP}} - G_{\text{RP}})_0$ . This space is shown in Figure 18, where  $(CaHK - G)_0 - 2.5(G_{\text{BP}} - G_{\text{RP}})_0$  is plotted against  $(G_{\text{BP}} - G_{\text{RP}})_0$  for the training sample, with the metallicity colour-coded into the plot. This figure is analogous to Figure 17, with the SDSS colours replaced by Gaia photometry and the training sample being extended. A detailed description of the algorithm assigning a metallicity estimate to a star's position in this colour-colour space is given by [Starkenburg et al. \(2017\)](#) and [Fernández-Alvar et al. \(2021\)](#). It essentially divides this space into a grid and assigns the mean metallicity of the training sample stars within a cell to this cell. It should be noted that the metallicity defaults to  $[Fe/H] = -4.0$  for stars close to the black no-metals line. As the metallicity depends on the de-reddened colour and thus the extinction, and the extinction depends on the metallicity, both the extinction coefficients and the metallicity of the stars are determined iteratively (see also the paragraph about extinction).

The input parameters  $(CaHK, G, G_{\text{BP}}, G_{\text{RP}})$  each have Gaussian uncertainties, and through 100 Monte-Carlo samplings of these, the probability distribution function of the photometric metallicity estimates is derived. The Pristine-Gaia synthetic catalogue thus provides, besides the metallicity estimate, the 16th,

---

<sup>9</sup><https://www.cosmos.esa.int/web/gaia/edr3-extinction-law>, retrieved 22.10.2023

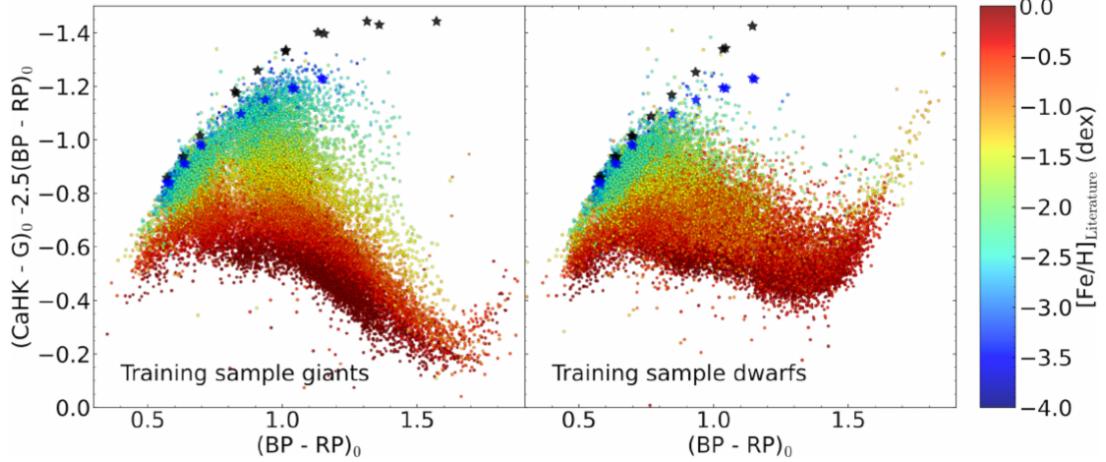


Figure 18: Training sample in the Gaia-Pristine colour-colour space with colour-coding according to metallicities from the literature. The blue stars are the synthetic colours from the synthetic spectra calculated for stars at  $[Fe/H] = -3.0$ , and the black stars signify hypothetical stars with no heavy metals at all (Martin et al. 2023).

50th, and 84th percentile of this distribution as well. As the majority of stars is above  $[Fe/H] = -1.0$ , Martin et al. (2023) focus on making the metallicities as accurate as possible in the range  $[Fe/H] < -1.0$ , so the method is not dominated by those comparably metal-richer stars. Thus, they warn that at  $[Fe/H] > -1.0$ , the metallicities can be less optimised and may include biases. Furthermore, they recommend to use a cut in the uncertainty of the metallicity estimates such as  $\sigma_{[Fe/H]} < 0.5$  dex or even 0.3 dex to remove very unreliable measurements. In summary, the results are designed and most reliable for FGK stars, in particular in the temperature range  $4000 < T_{\text{eff}} < 6000$  K, in the low-metallicity regime, and for relatively bright stars at  $G_{\text{BP}} \leq 16$  mag.

### 3.3 Selection criteria

The selection of targets for the input catalogue of the 4MOST S2 survey is based on the astrometry and photometry from Gaia DR3 and on the metallicity estimates derived from applying the Pristine survey's method to the low-resolution BP/RP spectra released with Gaia DR3. The selection criteria are summed up in Table 3 and described in detail by Christlieb et al. (2019).

Table 3: Target selection criteria split up into the three sub-surveys. Note that compared to the table given in (Christlieb et al. 2019), the actual upper limit in declination of  $+5^{\circ}$  and the updated metallicity cut  $[Fe/H] < -0.4$  are used.

#	<b>Bright survey</b>	<b>Faint survey</b>	<b>Deep survey</b>
1	$+5^{\circ} \geq \delta \geq -80^{\circ}$	$ b  > 20^{\circ}$	selected areas
2			
3	$[Fe/H] < -0.4$		
4	$12.0 \leq G \leq 14.5$	$14.5 < G \leq 15.5$	$15.5 < G \leq 17.0$
5	$0.15 \leq (G_{\text{BP}} - G_{\text{RP}})_0 \leq 1.10$		
6	$(1.10 < (G_{\text{BP}} - G_{\text{RP}})_0 \leq 1.60) \& (M_G < 3.5)$		
# stars	$\sim 750\,000$	$\sim 960\,000$	$\sim 70\,000$

The declination criterium was originally given as  $+20^{\circ} \geq \delta \geq -80^{\circ}$  in Christlieb et al. (2019), but the survey area that will be accessible to 4MOST will be limited to  $+5^{\circ} \geq \delta \geq -80^{\circ}$ , whereas the condition for the galactic latitude  $|b| > 20^{\circ}$  is chosen because S2 targets halo stars. For the deep survey, only stars in the WAVES fields (Driver et al. 2019) and members of globular clusters will be observed, which will be around 70 000. The extragalactic surveys only observe with the low-resolution fibres, leading

to a favourable situation regarding the availability of high-resolution fibres at high galactic latitude. As many stars in the range  $20 < |b| < 30^\circ$  are also targeted by 4MIDABLE-HR/S4, with the number of common targets estimated in the hundreds of thousands, spectra will be exchanged between S2 and S4. The purpose of the metallicity cut at  $[\text{Fe}/\text{H}] = -0.4 \text{ dex}$  is to remove many stars that belong to the disk, while still being high enough to retain the stars in the metallicity range where it is possible to identify inner and outer halo populations from  $[\alpha/\text{Fe}]$  measurements.

The division of the total apparent magnitude range  $12.0 \leq G \leq 17.0 \text{ mag}$  into three sub-surveys is motivated by the overall survey strategy, making the exposure times compatible with those required by the low-resolution surveys in the corresponding fields. The bright survey covers the range  $12.0 \leq G \leq 14.5 \text{ mag}$ , for which spectra with  $\text{S/N} > 50$  per pixel can be recorded in less than 2 hours. For both the faint survey ( $14.5 < G \leq 15.5 \text{ mag}$ ) and the deep survey ( $15.5 < G \leq 17.0 \text{ mag}$ ), the  $\text{S/N}$  that is needed is only half compared to the bright survey. The deep survey in particular will be focusing on sky regions which have longer exposures dedicated to them within the scope of other surveys. As the 4MOST surveys targeting the halo in low (S1) and high resolution (S2) cover different apparent magnitude ranges, they complement each other, with S1 allowing for the exploration of the outer halo, whereas S2 will contain a higher fraction of inner halo stars.

Regarding the colour cuts, the limit  $(G_{\text{BP}} - G_{\text{RP}})_0 \geq 0.15 \text{ mag}$  at the blue end was chosen to still contain main sequence turnoff stars at the lowest metallicity end with an age of 13 Gyr, whereas the limit  $(G_{\text{BP}} - G_{\text{RP}})_0 \leq 1.60 \text{ mag}$  discards M-type main sequence stars while keeping metal-poor giants of spectral type K in the target catalogue. In order to also get rid of K-type main sequence stars that are part of the galactic disk, an absolute-magnitude limit of  $M_G = 3.5 \text{ mag}$  is chosen in the colour range  $1.10 \leq (G_{\text{BP}} - G_{\text{RP}})_0 \leq 1.60 \text{ mag}$  to distinguish between giants and dwarfs. All stars in this range with  $M_G > 3.5 \text{ mag}$  are discarded. In order to make this selection, the absolute magnitudes were calculated using the Gaia parallaxes  $\pi$ , corrected for the zero point offset according to [Lindegren et al. \(2021a\)](#), in the case that  $\pi > 3\sigma_\pi$ , and using the GSP's photometric distances given in Gaia DR3 otherwise.

The target catalogue is built up by applying all selection criteria to the Pristine-Gaia synthetic metallicity catalogue, for which a cross-identification with the Gaia DR3 catalogue has to be carried out via the source identifiers, as the Pristine-Gaia-Synthetic catalogue does not contain all parameters needed for this. At the same time, relevant Gaia parameters for the following analysis can also be obtained. The declinations for the working version of the catalogue still go up to  $+20^\circ$  as stated in the original table in [Christlieb et al. \(2019\)](#), but aside from the verification of the metallicity estimates (Section 4.2) and the application of the criteria to GALAH (Section 4.3), only objects below  $\delta = +5^\circ$  will be retained in this work. Thus, the “full” version of the target catalogue contains about 4.6 million stars, but this number sinks to 3 809 536 when taking into account that the upper limit in declination of 4MOST will be  $+5^\circ$  for the final catalogue, with around 750 000 in the bright survey, 960 000 in the faint survey, and 2 100 000 in the deep survey. The following analysis will not be restricted to the WAVES fields and globular clusters for the deep sub-survey, but will instead make use of the full 2.1 million stars that were selected. It should be noted that, in spite of the colour range shown in Table 3, the Pristine-Gaia synthetic metallicity catalogue only includes the colour range  $0.5 < (G_{\text{BP}} - G_{\text{RP}})_0 < 1.5$  as described in Section 3.2.

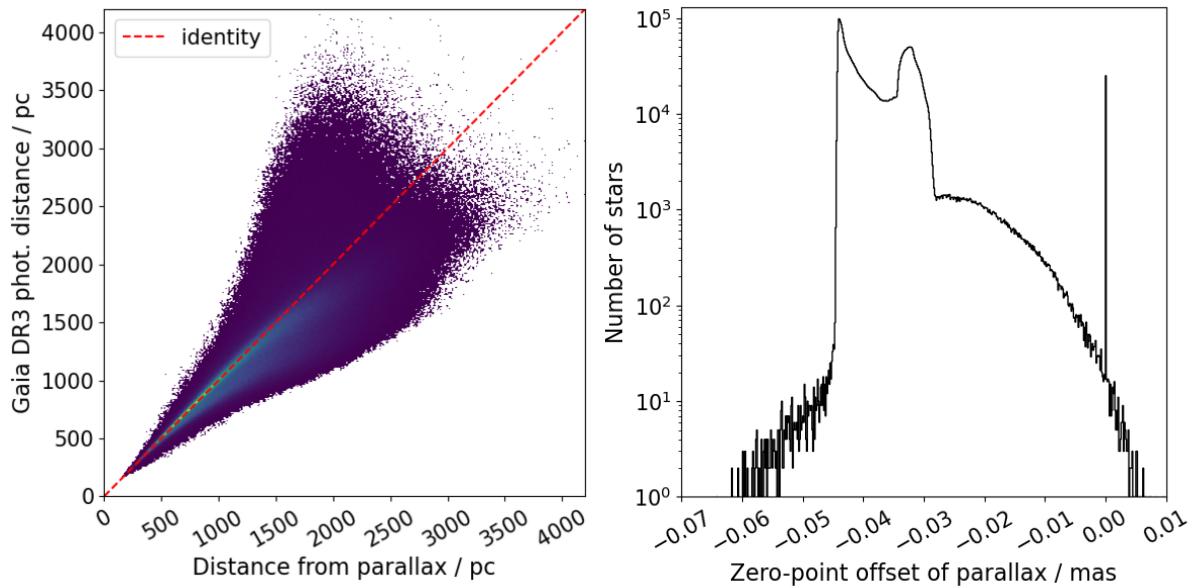
## 4 Verification of the 4MOST S2 catalogue

### 4.1 Colour-magnitude diagrams

In order to showcase the rough distribution in absolute magnitude–colour space and to check which metallicity trend they follow via comparison to isochrones, a colour-magnitude diagram (CMD) containing the around 3.8 million stars is produced. The absolute magnitude  $M_G$  is calculated from the distance modulus  $m - M = 5 \log d - 5 + A$  to be

$$M_G = G + 5 - 5 \log d - A_G \quad (1)$$

where  $G$  is the apparent magnitude in Gaia’s  $G$  band,  $d$  is the distance in parsec, and  $A_G$  is the extinction in the  $G$  band. The distance  $d$  is calculated from the parallaxes via  $d = \frac{1}{\pi}$  for all stars where  $\frac{\pi}{\sigma_\pi} \geq 3$  for the fraction of parallax to parallax uncertainty, which is the case for about 94.5% of the sample. Using the photometric distance estimates provided in Gaia DR3 would increase the percentage of stars for which the distance, and thus the absolute magnitude, is obtained to 98.2%. However, as pointed out by [Babusiaux et al. \(2023\)](#), the reliability of these estimates is not always certain, as can also be appreciated in Figure 19a: Here, the photometric distances are plotted against the distance obtained from the inverse parallax for those stars where the fractional error of both the parallax and the photometric distances is smaller than 10%. Particularly for large distances, the deviations become quite dramatic, up to more than 1000 pc. Thus, only the distances derived from the parallax according to the condition above are retained for the analysis. Observations of quasars in Gaia data have shown that the parallaxes have a bias and are offset by tens of microarcseconds, the extent of which depends on the magnitude, colour, and ecliptic latitude of the objects ([Lindegren et al. 2021a](#)). Using the Python script provided by [Lindegren et al. \(2021a\)](#), the parallaxes of stars with 5-parameter solutions are corrected to  $\pi_{\text{true}} = \pi_{\text{observed}} - zp$ . A histogram showing the range over which these zero-point offsets  $zp$  are distributed is shown in Figure 19b, barring a few outliers.



(a) The comparison between the photometric distances and inverted-parallax distances in the sample for sources where the fractional error on both quantities is at most 10%. (b) Histogram of the parallax zero-point offsets  $zp = \pi_{\text{observed}} - \pi_{\text{true}}$ , calculated according to the recipe by [Lindegren et al. \(2021a\)](#).

Figure 19

As the Pristine-Synthetic catalogue contains the dereddened  $G_{\text{BP},0}$  and  $G_{\text{RP},0}$  magnitudes, the dereddened colour  $(G_{\text{BP}} - G_{\text{RP}})_0$  is adopted from there ([Martin et al. 2023](#)). In order to take into account the extinction in the  $G$  band  $A_G$ , the Total Galactic Extinction map from Gaia DR3 is used, which is

based on Gaia BP/RP spectra (Delchambre et al. 2023) and provides the monochromatic extinction  $A_0$  as adopted in the Fitzpatrick (1999) extinction law for each galactic coordinate pair ( $l, b$ ). Using  $A_0$ , derived as described using the DUSTMAPS Python module (Green 2018), the corresponding extinction in the  $G$  band is calculated as  $A_G = k_G A_0$  using the Gaia DR3 extinction law provided by the Gaia team<sup>10</sup>, where

$$k_G = a_1 + a_2 X + a_3 X^2 + a_4 X^3 + a_5 A_0 + a_6 A_0^2 + a_7 A_0^3 + a_8 A_0 X + a_9 A_0 X^2 + a_{10} X A_0^2. \quad (2)$$

The coefficients were determined via fits on an extinction grid, with the Gaia (E)DR3 passbands (Riello et al. 2021) being convolved with the spectra by Castelli & Kurucz (2003) and the solar metallicity extinction law by Fitzpatrick et al. (2019). These fits were carried out for  $X = (G_{\text{BP}} - G_{\text{RP}})_0$ ,  $X = (G - K)_0$ , and  $X = T_{\text{eff}}/5040 \text{ K}$ . Here, the colour  $(G_{\text{BP}} - G_{\text{RP}})_0$  is used to calculate  $A_G$ . Furthermore, the coefficients were determined separately for main sequence stars and for giants, with the coefficients for giant stars being used up to  $M_G \sim 5$ , and the coefficients for main sequence stars for the rest. Finally, the CMDs can be produced for the full sample and the three sub-surveys and overplotted with several isochrones (Figure 20). In this work, the PARSEC isochrones (Bressan et al. 2012; Marigo et al. 2017) are used, which are available for metallicities down to  $[\text{M}/\text{H}] = -2.2 \text{ dex}$ . Several isochrones of age 12 Gyr with metallicities ranging from  $[\text{M}/\text{H}] = 0.0 \text{ dex}$  to  $[\text{M}/\text{H}] = -2.2 \text{ dex}$  were computed for the Gaia photometric system using the PARSEC web interface<sup>11</sup>.

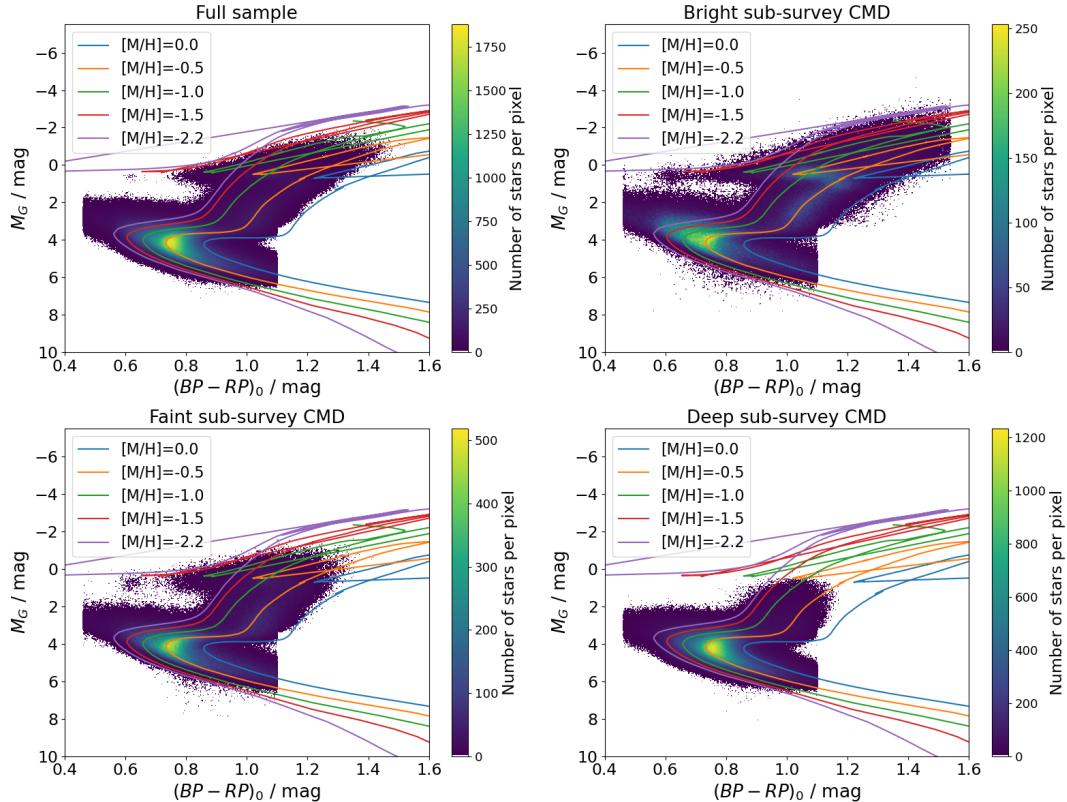


Figure 20: Colour-magnitude diagrams of the full 4MOST S2 catalogue, including a set of PARSEC isochrones. The top left panel shows the CMD of the full catalogue, whereas the other three panels show the CMDs of the three sub-surveys. A density plot is chosen due to the large number of stars.

As can be seen, the sample is dominated by main sequence turnoff stars, centered at around  $(G_{\text{BP}} - G_{\text{RP}})_0 \sim 0.75$  and  $M_G \sim 4$ . Particularly towards the faint end, the fraction of giant stars decreases dramatically, with the bottom right panel of the deep sub-survey showing no horizontal-branch stars anymore. In the top right panel of the bright sub-survey, however, features such as the red giant branch or

<sup>10</sup><https://www.cosmos.esa.int/web/gaia/edr3-extinction-law>, retrieved 22.10.2023

<sup>11</sup><http://stev.oapd.inaf.it/cgi-bin/cmd>, retrieved 22.10.2023

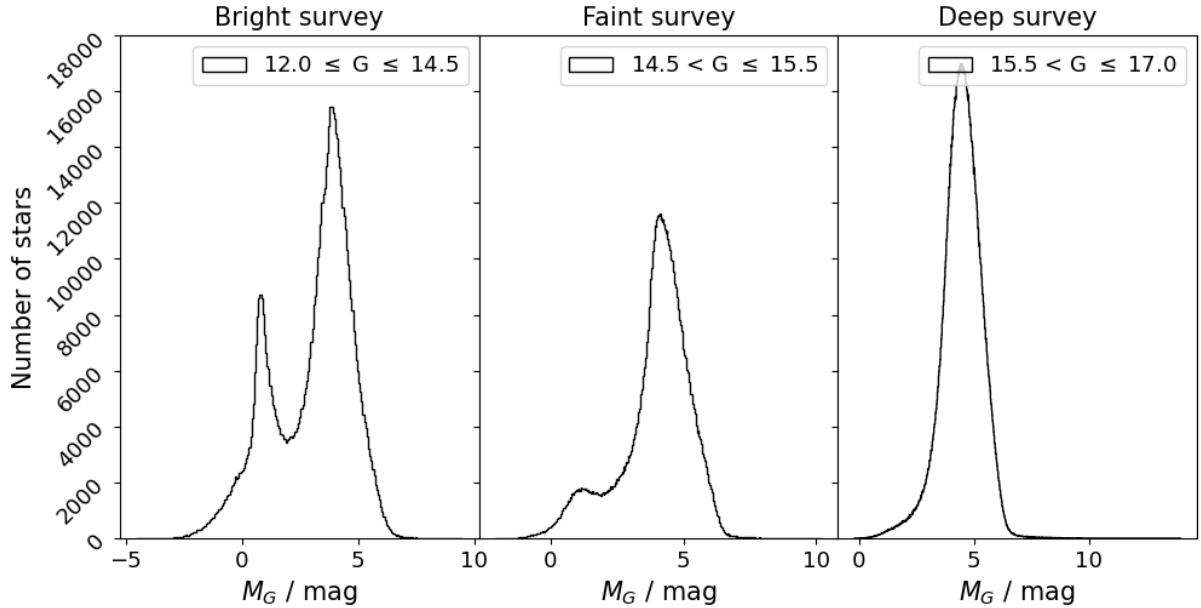


Figure 21: Histogram of absolute magnitudes for the three sub-surveys showcasing the decrease in giant stars towards the faint end of the sample.

the horizontal branch are clearly visible. The lack of giant stars in the sample at the faint end can also be seen in Figure 21 showing the distribution of absolute magnitudes for the three individual sub-surveys. Whereas for the bright sub-survey there are still two clear peaks for the giant stars and the main sequence turnoff stars, respectively, the histogram of the faint sub-survey contains only a tiny peak at absolute magnitudes typical for giants, and in the deep sub-survey, there is only a single peak. As giant stars in the apparent magnitude range of the deep sub-survey would be at very large distances, it is possible that the lack of giant stars at the faint end of the sample is due to the stellar density within the halo dropping so low that there are basically no stars anymore to be included.

The isochrones show that the sample appears to be most well-represented by the isochrone of metallicity  $[M/H] = -0.5$  dex, judging from the position of the main sequence turnoff, and also from the additional features in the CMD of the bright sub-survey. However, a large amount of scatter is present. This may in part be due to measurement uncertainties, as the parallax condition  $\frac{\pi}{\sigma_\pi} \geq 3$  is rather liberal and still allows for a decent amount of uncertainty within the distance determination. Furthermore, while the Gaia photometry is very accurate, the reddening and extinction are based on models which also come with their own uncertainties and caveats. Another effect widening the CMD is cosmic scatter, and since the catalogue includes stars with a large variety of metallicities, it is also expected that the CMD is rather spread out, which is also shown by the isochrones of different metallicities in the CMD being distributed across a large colour range. Finally, it should be noted that at the blue end at  $(G_{\text{BP}} - G_{\text{RP}})_0 \sim 0.5$  and in the lower right corner of the CMDs, cutoffs from the photometric selection effects described in Section 3.3 are visible.

## 4.2 Verification of Pristine-calibrated metallicities

The metallicity cut is one of the central criteria involved in the selection of the target catalogue, as it will have a direct influence on its composition, i.e. the amount of disk and halo stars, and also on the fraction of very low-metallicity stars. Figure 22 shows how the number of objects in the target catalogue changes for each sub-survey when different metallicity cuts in the range from  $[\text{Fe}/\text{H}] = -0.4$  dex to  $-1.2$  dex are chosen. This figure shows that even a difference of 0.1 dex changes the number of targets by hundreds of thousands at the higher metallicities, and particularly the bright and faint surveys are affected more dramatically compared to the deep survey. Thus, it is of great importance that the reliability and accuracy

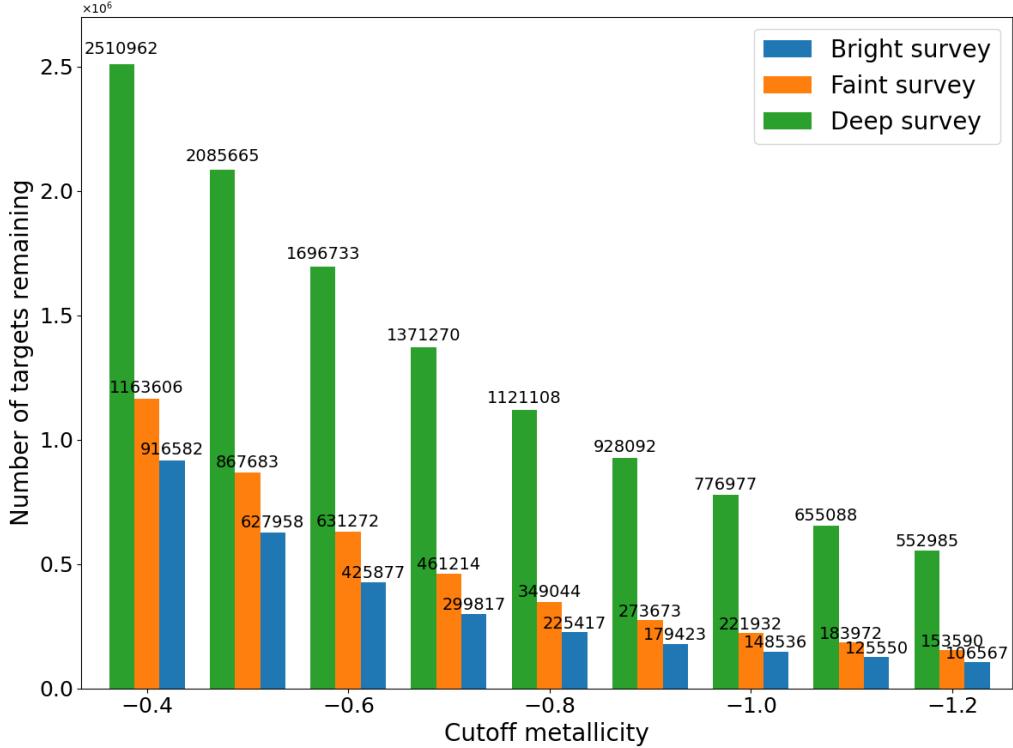


Figure 22: The size of the target catalogue for decreasing metallicity cuts. The number remaining in each individual sub-survey is plotted on top of the corresponding bar.

of the chosen metallicity estimates is ensured. In order to verify the Pristine-Gaia-Synthetic metallicities, they are compared to several reference catalogues, namely GALAH DR3, SAGA, the data-driven metallicity estimates by [Andrae et al. \(2023b\)](#), and APOGEE.

For GALAH DR3, the recommended table from the GALAH website<sup>12</sup> is downloaded. Using the tool TOPCAT ([Taylor 2005](#)), the crossmatch between the 4MOST S2 target catalogue is carried out with a search radius of 0.07 arcsec. No correction for proper motion had to be implemented, since GALAH DR3 includes the Gaia DR2 coordinates, and Gaia DR2 coordinates are given for epoch J2015.5 whereas Gaia DR3 coordinates are given for epoch J2016.0. Such a small radius is also sufficient because the Gaia coordinates tend to be very similar across the data releases, as the same instrument recorded the data, and the processing of data is similar. In fact, the left panel in Figure 23 shows that an even smaller radius would have been sufficient, and basically no more matches are found beyond  $\sim 0.04$  arcsec. In total, 62 988 stars out of the 588 571 stars in the full GALAH DR3 catalogue are found in the crossmatch. For the SAGA database, the recommended table containing 12 094 objects<sup>13</sup> is used. Before the crossmatch is carried out, the Gaia DR3 coordinates are transformed to epoch J2000.0 with the TOPCAT tool using their right ascension, declination, parallax, and proper motion, as the SAGA coordinates are given in epoch J2000.0. A search radius of 3.0 arcsec is adopted to allow for greater deviations (see Figure 23, centre panel), as the coordinates in the table are gathered from many different datasets across several decades. The crossmatch identifies 1250 stars that appear in both the SAGA database and the 4MOST S2 target catalogue.

For the crossmatch with APOGEE, the procedure is similar, with the downloaded table of Data Release 17<sup>14</sup> being crossmatched to the 4MOST S2 target catalogue after transforming the Gaia coordinates to

<sup>12</sup>[https://www.galah-survey.org/dr3/the\\_catalogues/](https://www.galah-survey.org/dr3/the_catalogues/), retrieved 22.10.2023

<sup>13</sup><http://sagadatabase.jp/>, retrieved 22.10.2023

<sup>14</sup>[https://www.sdss4.org/dr17/irspec/spectro\\_data/](https://www.sdss4.org/dr17/irspec/spectro_data/), retrieved 22.10.2023

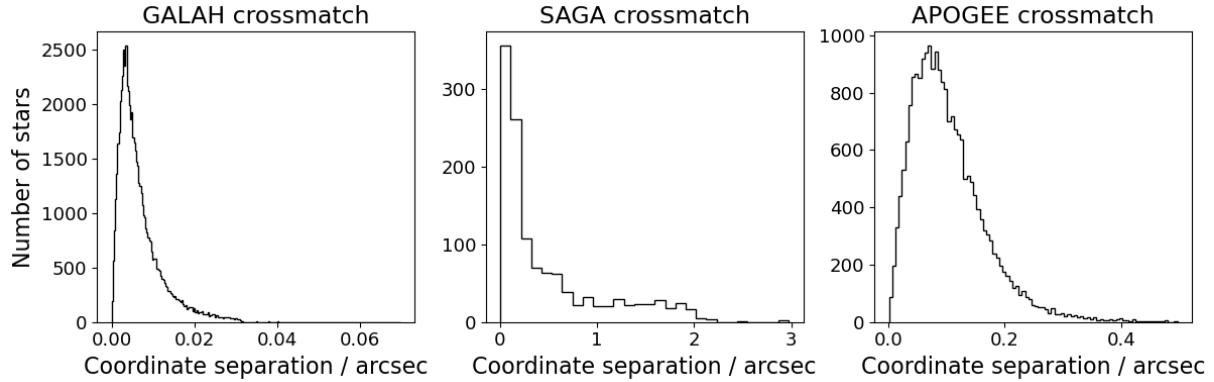


Figure 23: Coordinate separation of crossmatched objects to Gaia DR3 coordinates for GALAH, SAGA, and APOGEE.

epoch J2000.0, but using a search radius of 0.5 instead (Figure 23, right panel). As the SDSS observes from the Northern hemisphere, relatively few objects are recovered in the crossmatch, at only 22 181 out of the total 733 901 in APOGEE DR17. As for the data-driven metallicities derived by Andrae et al. (2023b)<sup>15</sup>, the Gaia DR3 source identifiers were given in the catalogue, so it is enough to carry out a cross-identification based on these. For the full sample, 4 463 361 out of 174 922 161 stars are found, whereas for the vetted sample of RGB stars which are deemed reliable, 488 218 out of 17 558 141 stars are found.

## GALAH

As the GALAH data will keep being used throughout the thesis, it shall be described here briefly. The Galactic Archaeology with HERMES (GALAH) survey (De Silva et al. 2015; Buder et al. 2021) is a spectroscopic survey aimed at obtaining high-resolution spectra of around one million objects in order to acquire stellar parameters and abundances, which can be used in combination with astrometry and kinematics to disentangle the history of our Galaxy. In particular, GALAH data is of interest for the method of Chemical Tagging, which is “the identification of stars that formed together using their chemical composition and an understanding of the astrophysics driving the dimensionality of chemical space” (Buder et al. 2021, p. 151). The third data release of GALAH includes stellar parameters for almost 600 000 objects, such as the effective temperature, surface gravity, [Fe/H], as well as the radial velocity among others, abundance ratios [X/Fe] for a large number of elements, and complementary astrometric data from Gaia DR2 as well as photometric data from 2MASS. About 2/3 of it consists of dwarf stars, about 1/3 of giant stars, and about 2% of stars have a metallicity  $[\text{Fe}/\text{H}] < -1$ .

The observations, carried out with the HERMES spectrograph at the Anglo-Australian Telescope (Sheinis et al. 2015), are restricted to  $\delta < +10^\circ$  and  $|b| > 10^\circ$ , with the additional condition that an observed area must have 400 objects or more per  $\pi \text{ deg}^2$ . Furthermore, the selection of targets includes only bright objects at  $12.0 < V_{\text{JK}} < 14.3$  mag. As a result, the dataset consists mainly of stars close to us, with around 80% of the sample being within 2 kpc, which also means that for almost the whole sample, high-quality 5D astrometry from Gaia DR3 is available (and for a lot of the stars, even 6D astrometry including the radial velocity measured by Gaia).

The data analysis for GALAH DR3 is essentially split into two parts. First, the stellar parameters are determined iteratively, and second, the elemental abundances are fit one by one while leaving the stellar parameters fixed. During this iterative process, a comparison with synthetic spectra reveals which lines are unblended enough to be used to get reliable abundance estimates. Notably, the spectra are not anal-

<sup>15</sup><https://zenodo.org/record/7945154>, retrieved 22.10.2023

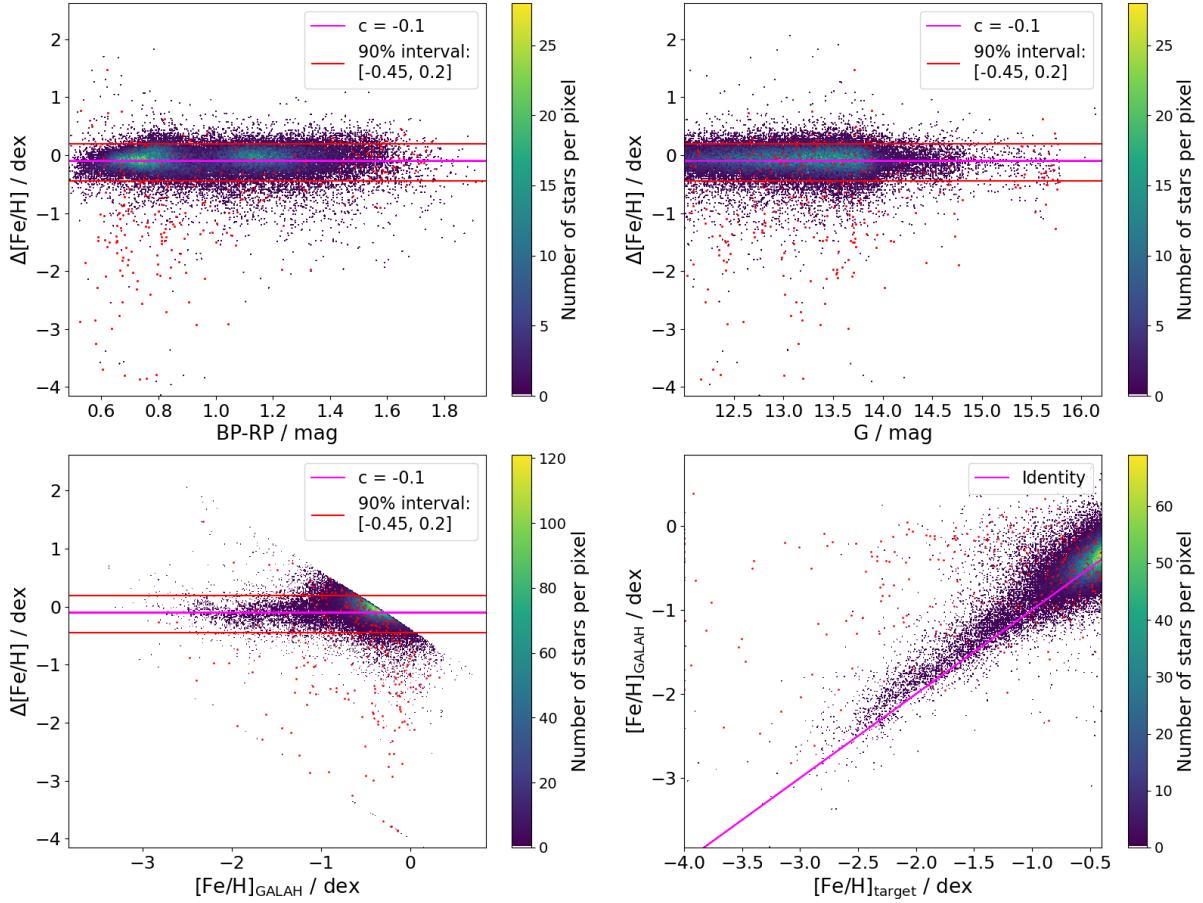


Figure 24: Comparison between the Pristine-Gaia-Synthetic metallicities and the metallicities from GALAH DR3. Due to the number of objects, the data is shown as a density plot. The red points are those for which  $(\sigma_{[\text{Fe}/\text{H}],84} - \sigma_{[\text{Fe}/\text{H}],16})/2 > 0.35$ , and were not taken into account for the fit.

ysed using a data-driven approach, but instead employing the spectrum synthesis code SPECTROSCOPY MADE EASY (SME, Piskunov & Valenti 2017), as there can be issues with the former such as difficulties to inter- and extrapolate, in particular in the case of limited training samples. Also, the authors exclude stars from their analysis which do not have an accuracy better than 5 km/s in radial velocity. The model atmospheres are based on the MARCS grid (Gustafsson et al. 2008) and encompass the ranges  $2500 \leq T_{\text{eff}} \leq 8000$  K,  $-0.5 \leq \log g \leq 5.5$  dex, and  $-5 \leq [\text{Fe}/\text{H}] \leq 1$  dex. They are in principle 1D hydrostatic models, but for some elements, 1D non-LTE departure grids are computed as detailed by Amarsi et al. (2020).

Figure 24 shows the comparison between the Pristine-Gaia-Synthetic metallicities and the GALAH DR3 metallicities. For GALAH, the stellar parameter flags are cleaned by requiring that `flag_sp = 0` to only include reliable estimates, and for Pristine-Gaia-Synthetic metallicities, only those that satisfy  $(\sigma_{[\text{Fe}/\text{H}],84} - \sigma_{[\text{Fe}/\text{H}],16})/2 < 0.35$  were taken into account for the comparison as recommended by Martin et al. (2023), where  $\sigma_{[\text{Fe}/\text{H}],16}$  and  $\sigma_{[\text{Fe}/\text{H}],84}$  are the lower and upper bound of the 68% confidence interval for the metallicity given. The stars for which  $(\sigma_{[\text{Fe}/\text{H}],84} - \sigma_{[\text{Fe}/\text{H}],16})/2 > 0.35$  are still shown in the figure in red, in order to illustrate how even that rather liberal cut removes most of the extreme outliers.

In the top left panel, the difference between the two metallicity estimates  $[\text{Fe}/\text{H}]_{\text{target}} - [\text{Fe}/\text{H}]_{\text{GALAH}}$  is plotted against  $G_{\text{BP}} - G_{\text{RP}}$ , in the top right against  $G$ , in the bottom left against  $[\text{Fe}/\text{H}]_{\text{GALAH}}$  itself, and in the bottom right corner, the two metallicities are plotted directly against each other. The

plot shows that the metallicities are in good agreement overall, with an offset of -0.10 in total (meaning that the GALAH estimates are on average 0.1 dex higher) and 90% of objects falling within the range  $[-0.45, 0.20]$  for the difference between the metallicity estimates. The general shape of the comparison does not change significantly across the  $G_{\text{BP}} - G_{\text{RP}}$  or the  $G$  range, and although the number of outliers seems to be increased at smaller  $G_{\text{BP}} - G_{\text{RP}}$  values and also at smaller  $G$  values, this could also be due to the fact that in these ranges, there are simply more data points, as the colour coding of the density plot shows. The lower two panels show that the metallicity estimates are in good agreement down to  $[\text{Fe}/\text{H}] \sim -2.5$  to  $-3.0$ , after which the number of stars becomes too small to make a meaningful comparison.

## SAGA

[Suda et al. \(2008\)](#) collect previously studied metal-poor stars from the literature into the SAGA (Stellar Abundances for Galactic Archaeology) database in order to facilitate further research in astronomy at the metal-poor end. Although they originally chose a metallicity limit of  $[\text{Fe}/\text{H}] \leq -2.5$  as their limit for which studies to include, they also add additional objects such as reference stars that are included in the same paper, meaning there are many objects above that limit as well, and remove this criterion for later releases ([Suda et al. 2017](#)). In the database, the abundances available from the different papers are listed, alongside other information such as equivalent widths, atmospheric parameters, and photometry in whichever bands the study it is extracted from uses. Since many of the stars included have been studied multiple times, the authors alert the reader regarding possible discrepancies between different analyses. However, in order to have a data set where each star has only one abundance or stellar parameter value despite being included in several studies, a fiducial value is chosen by either using the average of the available data, using the median of the available data, or picking a representative value based on a priority parameter. The priority parameter itself is calculated taking into account criteria such as the resolving power and the year of publication amongst others ([Suda et al. 2017](#)).

Figure 25 shows the comparison between the Pristine-Gaia-Synthetic metallicities and the SAGA database metallicities. For the SAGA crossmatch, most objects have  $[\text{M}/\text{H}]$  given as a metallicity estimate, and only just over half give  $[\text{Fe}/\text{H}]$  as well. For this comparison,  $[\text{Fe}/\text{H}]$  was used, as the Pristine-Gaia-Synthetic metallicities are also given as  $[\text{Fe}/\text{H}]$ . Once again, only the Pristine-Gaia-Synthetic metallicities for which  $(\sigma_{[\text{Fe}/\text{H}],84} - \sigma_{[\text{Fe}/\text{H}],16})/2 < 0.35$  are used for the fit, and the others were only added to the plot in red for additional information. In the top left panel, the difference between the two metallicity estimates  $[\text{Fe}/\text{H}]_{\text{target}} - [\text{Fe}/\text{H}]_{\text{SAGA}}$  is plotted against  $G_{\text{BP}} - G_{\text{RP}}$ , in the top right against  $G$ , in the bottom right against  $[\text{Fe}/\text{H}]_{\text{SAGA}}$  itself, and in the bottom right corner, the two metallicities are plotted directly against each other. Here as well, there is generally a good agreement between the two metallicities, with an offset of 0.11 (meaning that the SAGA estimates are on average 0.11 dex lower), but the spread around the fit line is a lot larger, with the range encompassing 90% of stars being  $[-0.70, 0.79]$  in the deviation between metallicities.

The bottom panels show that there is an overall good agreement in the metallicities in the range  $[\text{Fe}/\text{H}] \sim -1.5$  to  $-3.0$  dex, before and after which the general trend is similar, but the scatter is much higher, although the number of stars in these ranges are also much reduced. Interestingly, as seen in the top panels, it seems that almost all the stars that were discarded according to the condition  $(\sigma_{[\text{Fe}/\text{H}],84} - \sigma_{[\text{Fe}/\text{H}],16})/2 > 0.35$  are at the faint end below  $G \sim 15.5$  mag, where they outweigh the stars that were left in, and most of them are on the blue end around  $G_{\text{BP}} - G_{\text{RP}} \sim 0.6$  to  $0.8$  mag. This is simply because the stars in the whole target catalogue which fulfill that condition are largely located on the faint end, likely because the quality of the BP/RP spectra used to derive the Pristine-Gaia-Synthetic metallicities is lower there. In fact, the top right panel in Figure 24 also showed that at the faint end, a rather high number of stars were discarded when considering the total number of data points in that region. Also, [Martin et al. \(2023\)](#) warn that the Pristine-Gaia-Synthetic metallicity estimates are most reliable for  $G_{\text{BP}} < 16$  mag.

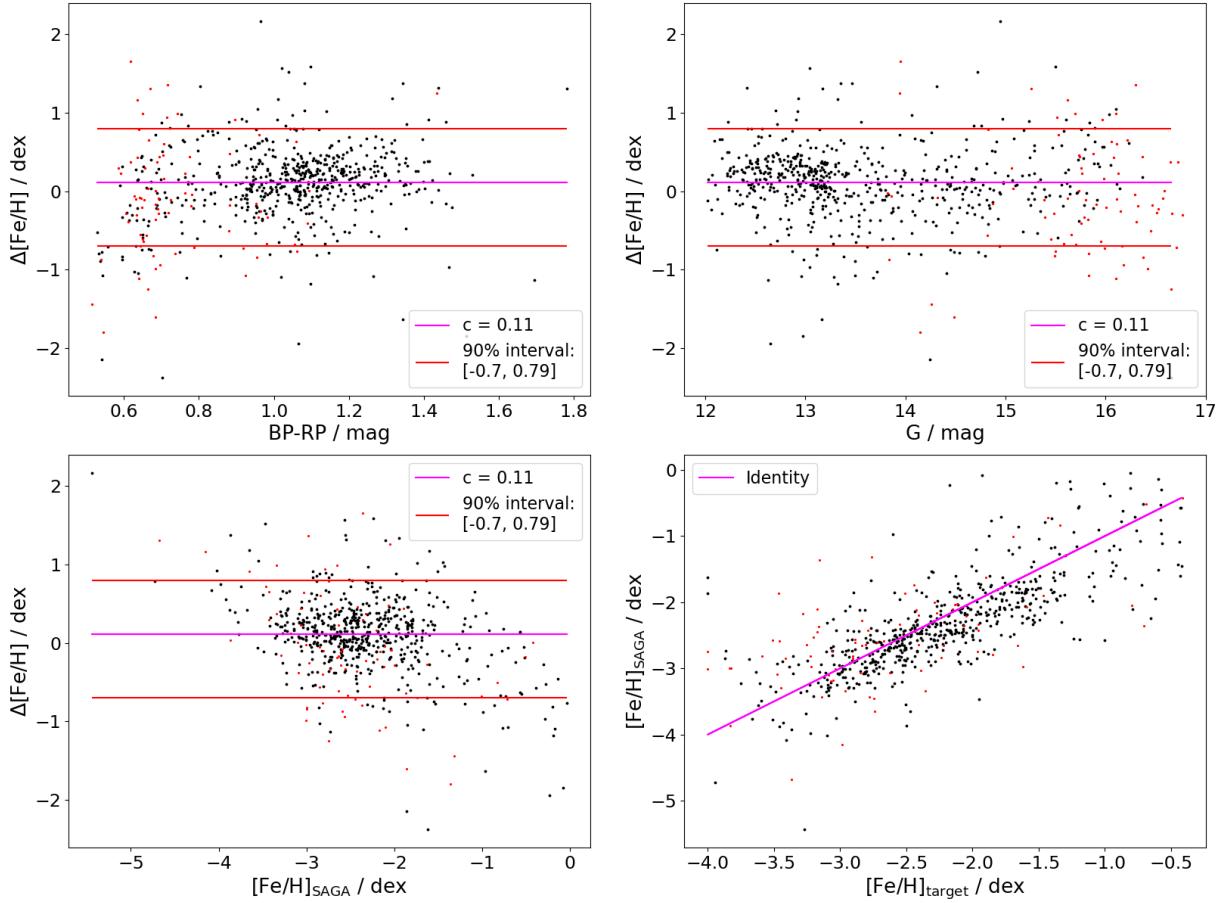


Figure 25: Comparison between the Pristine-Gaia-Synthetic metallicities and the metallicities from the SAGA database. The red points are those for which  $(\sigma_{[\text{Fe}/\text{H}],84} - \sigma_{[\text{Fe}/\text{H}],16})/2 > 0.35$ , and were not taken into account for the fit.

### Andrae et al. (2023b) data-driven metallicities

Andrae et al. (2023b) make use of the low-resolution spectra made available by Gaia DR3 in the BP and RP bands in order to derive metallicities, as well as effective temperatures and surface gravities. They aim to improve over metallicities published within Gaia DR3 by employing data-driven methods and concentrating on stellar types that have proven effective in revealing their metallicities even for low resolutions, as demonstrated by Rix et al. (2022). However, in order to have a catalogue that is as complete as possible, they derive metallicities for almost all stars where low-resolution XP spectra are available, noting that for those stars that do not fall within the above-mentioned category (essentially low S/N and high effective temperature), the estimates might be more uncertain. The XGBoost models are trained on stars from APOGEE DR17 (Abdurro'uf et al. 2022), which extends down to  $[\text{M}/\text{H}] \sim -2.5$ , and on the metal-poor stars from Li et al. (2022) in order to have some training data in the very metal-poor/extremely metal-poor regimes. Some ranges of stellar parameters have little to no training objects (with the effective temperature being limited to the range 3107 K to 6867 K, for example), and the completeness of the predictions decreases towards the fainter end, which is due to the CatWISE  $W_2$  magnitudes' uncertainty becoming higher, although the usage of CatWISE photometry (Marocco et al. 2021) already marks an improvement over the previously used ALLWISE photometry. The input parameters consist of 38 colours from Gaia and CatWISE photometry and the 110 XP spectra coefficients, and since the parallax can improve estimates, information on it was also included.

Figure 26 shows the comparison between the Pristine-Gaia-Synthetic metallicities and the catalogue of metallicities by Andrae et al. (2023b). As was done for GALAH and SAGA, only stars with  $(\sigma_{[\text{Fe}/\text{H}],84} - \sigma_{[\text{Fe}/\text{H}],16})/2 < 0.35$

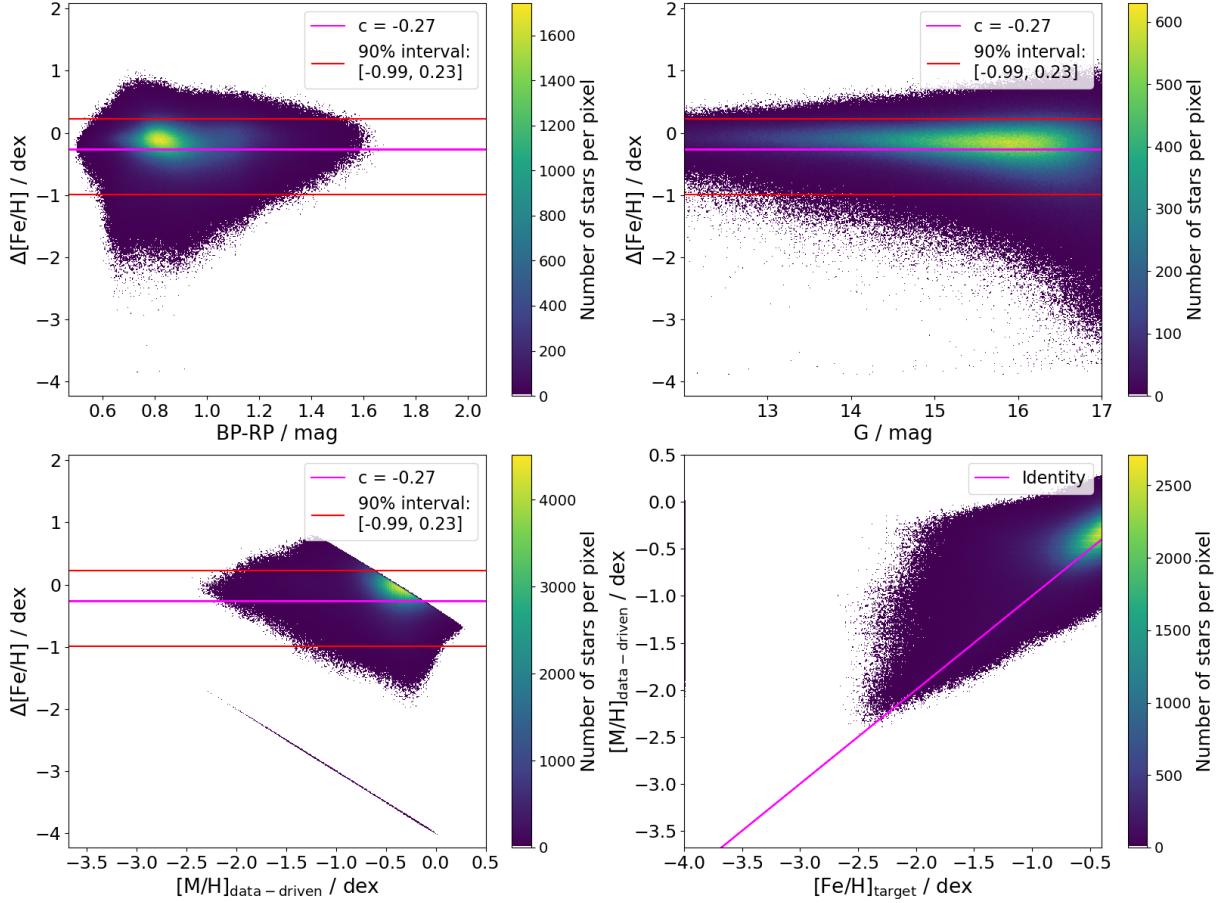


Figure 26: Comparison between the Pristine-Gaia-Synthetic metallicities and the data-driven metallicities by [Andrae et al. \(2023b\)](#). As can be seen, the offset line is slightly below the bulk of data points indicated by the colour coding, probably due to the increasingly more extreme outliers towards the faint end dragging it down.

$\sigma_{[\text{Fe}/\text{H}],16})/2 < 0.35$  are considered in the fit, but here, the other stars are not shown to not limit visibility in the plot. The total offset is -0.27 (meaning that the data-driven estimates are on average 0.27 dex higher), and 90% of objects are within the range of  $[-0.99, 0.23]$  for the deviation of metallicities, although it is visible that the offset line does not pass centrally through the bulk of data points but is instead a bit below, likely pulled down by the increasingly growing deviations at the faint end which can be seen in the top right panel. This increasing deviation is most likely the result of the decreasing completeness of the predictions towards the faint end mentioned above (see also Figure 3 in [Andrae et al. 2023b](#)). On the other hand, the Pristine-Gaia-Synthetic metallicities are also reported to be less reliable for stars fainter than  $G_{\text{BP}} = 16$ , so it is possible that this contributes to the large deviations on the faint end, although the trend seen in the other reference comparisons suggests that it is mainly due to uncertainties of the data-driven metallicities. Although there is a significant fraction of the sample with large deviations from the Pristine-Gaia-Synthetic metallicities, the majority of the data-driven metallicities follow the general trend rather well. The obvious cut in the bottom left panel arises from the metallicity cut at  $[\text{Fe}/\text{H}] = -0.4$ , whereas the line in the bottom half of the image (which can also be seen at the leftmost edge in the bottom right panel) is the result of the Pristine method defaulting to  $[\text{Fe}/\text{H}] = -4.0$  for no detected features in the Ca H&K lines. These two effects could also be seen in Figure 24 and, albeit much less obviously, in Figure 25.

[Andrae et al. \(2023b\)](#) also provided a vetted sample of RGB stars which was created using rather strict cuts in the parameters such as  $G < 16$  mag,  $\frac{\pi}{\sigma_\pi} > 4$ ,  $\log g < 3.5$ , and  $T_{\text{eff}} < 5200$  K among others in order to build up a sample of RGB stars with metallicity estimates that are deemed reliable. The com-

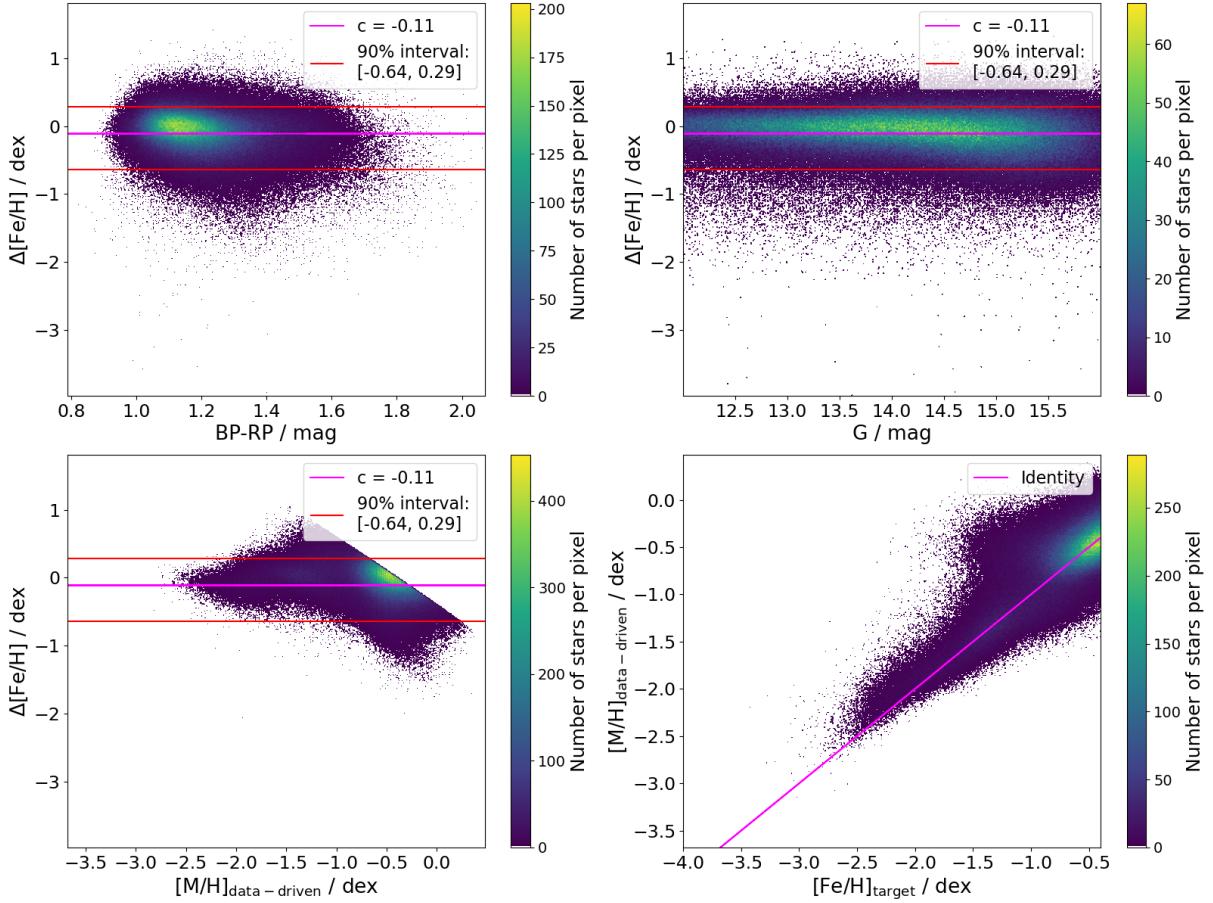


Figure 27: Comparison between the Pristine-Gaia-Synthetic metallicities and the vetted RGB sample of the data-driven metallicities by [Andrae et al. \(2023b\)](#). This sub-sample deviates less from the Pristine-Gaia-Synthetic metallicities and the deviations are more regularly distributed.

parison to this sample is shown in Figure 27, and immediately shows much better agreement with the Pristine-Gaia-Synthetic metallicities, and a basically constant shape of the deviation distribution across the whole magnitude range. The offset is very similar to GALAH at -0.11, although the spread is a bit larger, with 90% of objects being in the range  $[-0.64, 0.29]$ . As in Figure 26, a significant number of objects is available only down to  $[Fe/H] \sim -2.5$  dex, with the very few objects below following the identity line in the bottom right panel approximately, but being almost entirely above it, showing that at this point, the data-driven estimates are higher compared to the Pristine-Gaia-Synthetic metallicities. The fact that the more reliable vetted RGB sample follows the Pristine method's data more closely indicates that the Pristine method remains more useful in determining accurate metallicities for a large number of objects from the BP/RP spectra, and confirms it as the right choice to obtain metallicities for the target catalogue selection over a data-driven approach such as the one by ([Andrae et al. 2023b](#)). It should be noted, however, that they provided estimates for  $[M/H]$  rather than  $[Fe/H]$ , which does tend to be slightly higher than  $[Fe/H]$ , as it takes other elements into account as well.

## APOGEE

Finally, Figure 28 shows the comparison between the Pristine-Gaia-Synthetic metallicities and APOGEE DR17's  $[M/H]$  measurements based on near-infrared observations. The APOGEE flags are cleaned by asking that  $STARFLAG = 0$ , and as is done for GALAH and SAGA, only objects with  $(\sigma_{[Fe/H],84} - \sigma_{[Fe/H],16})/2 < 0.35$  are taken into account in the fit, and those that do not satisfy that condition are plotted into the figure in red, which shows many of the most severe outliers are removed by that condition. The offset is in line with GALAH and the data-driven metallicities' offsets at -0.13, and the interval which

contains 90% of data points is  $[-0.72, 0.24]$ . Although there is a significant number of stars for which much lower metallicities are reported by the Pristine method, the general agreement holds well down to APOGEE’s lower metallicity limit of  $[M/H] = -2.5$ . As is the case for the data-driven metallicities in the previous section, the metallicity estimates are actually  $[M/H]$ , which should be kept in mind when making the comparison.

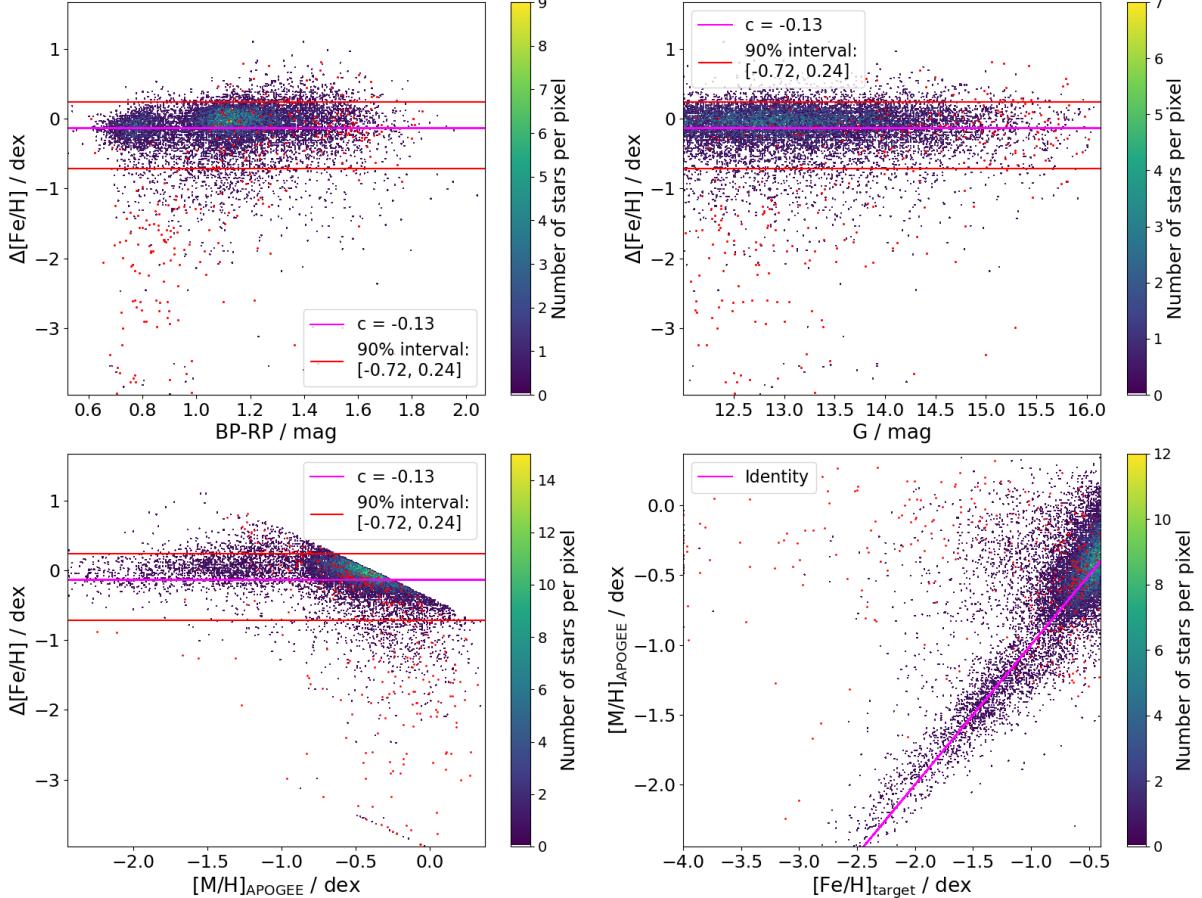


Figure 28: Comparison between the Pristine-Gaia-Synthetic metallicities and the APOGEE DR17 metallicity estimates. The red points are those for which  $(\sigma_{[Fe/H],84} - \sigma_{[Fe/H],16})/2 > 0.35$ , and were not taken into account for the fit.

### Effect of carbon enhancement

Due to strong contamination by carbon bands in spectra of stars that are carbon-enhanced, the effect of carbon on the metallicity estimates provided in the Pristine-Gaia-Synthetic catalogue is considered. For this purpose, the stars in the SAGA crossmatch which also have carbon measurements available are used to investigate a systematic offset between C-enhanced and C-normal stars. Figure 29 shows the comparison in metallicity estimates between SAGA and Pristine, but this time including colour-coding indicating whether a star has an elevated carbon abundance ( $[C/Fe] > 0.5$ ) or not ( $[C/Fe] < 0.5$ ) according to the SAGA data. As can be seen, the offset for the C-enhanced stars is 0.36 dex higher compared to the C-normal stars and almost 90% of C-normal stars are below the fit line of the C-enhanced stars, while more than 80% of C-enhanced stars are above the fit line of the C-normal stars. This indicates that C-enhanced stars do on average have their metallicities overestimated in the Pristine-Gaia-Synthetic catalogue. This effect does not play a role in the verification with the other reference catalogues, as they contain mostly higher-metallicity stars, which are much more rarely carbon-enhanced (e.g., [Arentsen et al. 2022](#)).

While C-enhanced stars are in general rather rare, they become increasingly common towards the low-

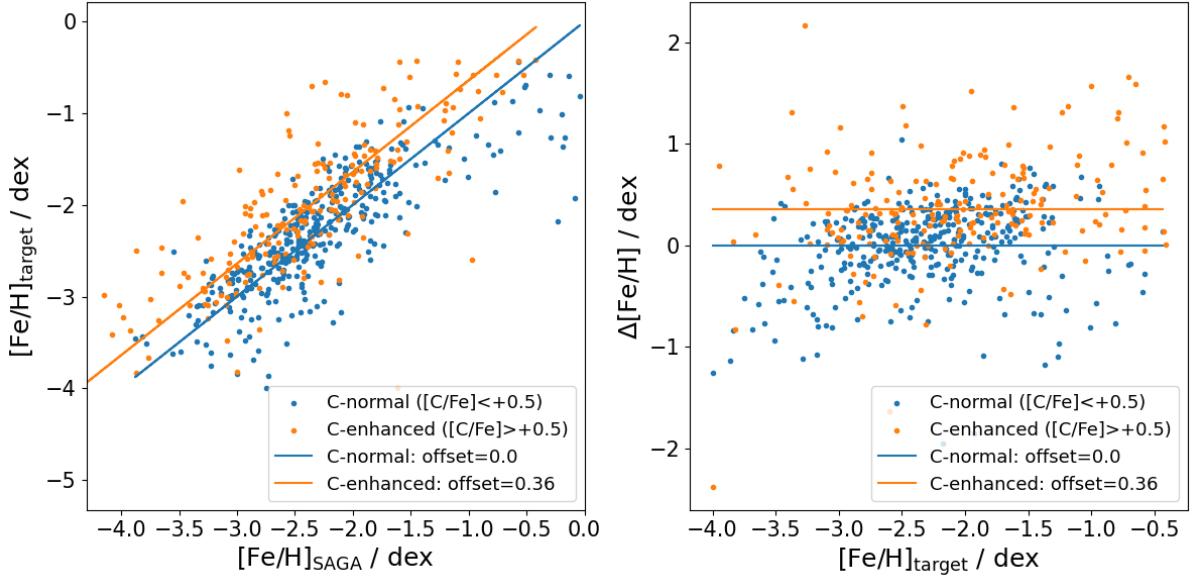


Figure 29: The impact of carbon on Pristine-Gaia-Synthetic metallicities. Whereas the offset from SAGA for C-normal stars is 0, there is an offset of 0.36 dex for C-enhanced stars, suggesting a systematic effect of C-enhancement on the Pristine metallicities.

metallicity end, and while different studies report a large range of fractions of C-enhanced stars in the metal-poor regime from a few percent up to more than 40%, the fractions are always significant (e.g., [Arentsen et al. 2022](#)). To get an idea of the presence of C-enhanced stars within the 4MOST S2 target catalogue, the catalogue of C-enhanced metal-poor (CEMP) stars by [Lucey et al. \(2023\)](#) is employed. With the machine learning algorithm XGBoost, the authors aim to detect CEMP stars ( $[C/\text{Fe}] > 0.7$ ) using the over 200 million BP/RP spectra published within the scope of Gaia DR3. The model is trained on data from SEGUE and AEGIS. Notably, the stars' evolutionary status has not been taken into account in the carbon abundances used. To avoid extrapolation, the colour and magnitude range of the sample is restricted to the range of the training sample. [Lucey et al. \(2023\)](#) find around 59 000 CEMP stars, the Galactic distribution of which hints that many of them are halo stars, as they tend to be farther from the Galactic plane. Doing a cross-identification with their data, a total of 8764 predicted CEMP stars is identified in the 4MOST S2 target catalogue, which is 0.18% of the sample. It should be noted that not all 4.6 million stars in the target catalogue were also in the catalogue by [Lucey et al. \(2023\)](#), and that this fraction increases for lower metallicity cuts: at  $[Fe/H] < -1.0$ , the fraction is 0.45%, and at  $[Fe/H] < -1.5$ , the fraction is 0.51%. As the total catalogue of [Lucey et al. \(2023\)](#) contains around 183 million stars, of which around 59 000 are predicted to be C-enhanced, the total fraction of CEMP stars in their work is only 0.03%, which shows that C-enhancement does play a role in the 4MOST S2 target catalogue at low metallicities.

On one hand, a strong presence of CH lines would result in a lower pseudo-continuum, thereby weakening the line strengths and decreasing the metallicity estimates derived from it. On the other hand, if there are CH features within the pixel of the low-resolution XP spectrum which also contains the Ca K line, it would result in an overestimation of metallicity for that star, which would be an explanation for the systematic offset seen in Figure 29, and is likely what is happening there. [Martin et al. \(2023\)](#) also discuss the impact of carbon on their data and reach similar conclusions using the SAGA database, and they urge to be cautious when using Pristine catalogues to investigate C-enhancement-related topics. A paper going into more detail regarding the causes of the bias is announced by the authors (M. Montelius et al., in prep.).

### 4.3 Application of selection criteria to GALAH

As the selection of interesting targets for a survey is among the most important aspects to be considered when deciding on selection criteria, an overview over which objects were not selected by these criteria is also very valuable. Furthermore, comparing the size of the crossmatched reference catalogue with the size of the subset of the reference catalogue which is left when applying the same selection criteria to it that were used to obtain the target catalogue also indicates whether there may have been some unexpected, undesired selection effects at play, and whether the selection is overall consistent. Here, GALAH DR3 is used as a reference catalogue, as its observations were also carried out in the Southern hemisphere within the optical wavelength range, and contain accurate metallicity estimates. The crossmatch between the target catalogue and GALAH DR3 was already described in the previous section, where 62 988 stars were found, while the complete GALAH DR3 catalogue contains 588 571 stars.

Figure 30 shows the distribution of galactic coordinates for the crossmatched stars and for the full GALAH DR3 catalogue. The black dots are the observation fields of the GALAH survey, and several of the different HERMES observations can be seen, such as the K2-HERMES follow-up showing the characteristic tile-pattern of the Kepler telescope (centrally from bottom left to top right), and the TESS-HERMES observation, which is the big blob at around  $(l, b) = (100^\circ, -30^\circ)$ . Comparing the distribution of the crossmatched sample to the full GALAH DR3 catalogue, the selection criterium  $|b| > 20^\circ$  is immediately visible. Aside from this, however, it appears that the stars in the target catalogue are very homogeneously selected across all galactic coordinates, as the top panel shows almost all the observation fields that are also seen in the bottom panel at  $|b| > 20^\circ$ , just with reduced thickness of the dots, i.e. reduced number of objects per field.

In order to get an overview over how many stars are removed by each of the selection criteria in Table 3.3, they are applied to the full GALAH DR3 catalogue step by step in the order survey area, apparent magnitude, colour and absolute magnitude, and finally metallicity. As basically all stars in the full GALAH DR3 catalogue can be identified in Gaia DR3, the photometric information and parallax is added to be able to appropriately apply the selection criteria to GALAH. After applying the declination ( $\delta < 20^\circ$ ) and galactic latitude criterium, the number of objects drops quite dramatically from the full  $\sim 588\,000$  to  $\sim 336\,000$ . This makes sense, as the GALAH criteria are  $\delta < 10^\circ$ ,  $|b| > 10^\circ$ , and to have at least 400 targets within an area of  $\pi \text{ deg}^2$ , meaning that the range between  $10^\circ < |b| < 20^\circ$  is only covered by GALAH (see also Figure 30). Since this is close to the disk, there should be a high density of stars here, so it is not surprising that a significant fraction of GALAH DR3's stars is not in the 4MOST S2 catalogue. Likewise, the objects that lie between  $10^\circ < \delta < 20^\circ$  in the working version of the 4MOST S2 catalogue do not appear in GALAH, just like those that lie in very low target density areas of the sky.

As for the apparent magnitude requirement, GALAH only observes stars with  $12.0 < V_{JK} < 14.3$ , where  $V_{JK}$  is a synthetic magnitude calculated from  $J$  and  $K$  2MASS photometry to approximate the  $V$  band magnitude. This corresponds mostly to the bright survey within 4MOST S2, and only few stars should be below  $G = 12$  mag, so the apparent magnitude selection criterium should not remove that many objects. This is indeed the case, as the number of stars drops to  $\sim 300\,000$  after applying the apparent magnitude criterium  $12.0 \leq G \leq 17.0$ . It should also be kept in mind for future uses that GALAH data is only really relevant for the bright sub-survey, as was also seen in Figure 24, where the comparison of metallicities only included a significant amount of data points down to  $G \sim 14$  mag.

In the next step, applying the colour and absolute magnitude criterium,  $\sim 281,000$  stars remain, as the lower main sequence and some very blue as well as very red stars are removed. In order to derive an approximate value of the colour excess  $E(\text{BP} - \text{RP})$  to be able to use  $(G_{\text{BP}} - G_{\text{RP}})_0$  in the criterium, the values of  $E(B - V)$  given in GALAH DR3 are transformed according to the photometric relationship between  $G_{\text{BP}} - G_{\text{RP}}$  and  $B - V$  provided by DPAC (Jordi 2018). The accuracy of this transformation for calculating the colour excess in particular is not very high, as the uncertainty given by Jordi (2018)

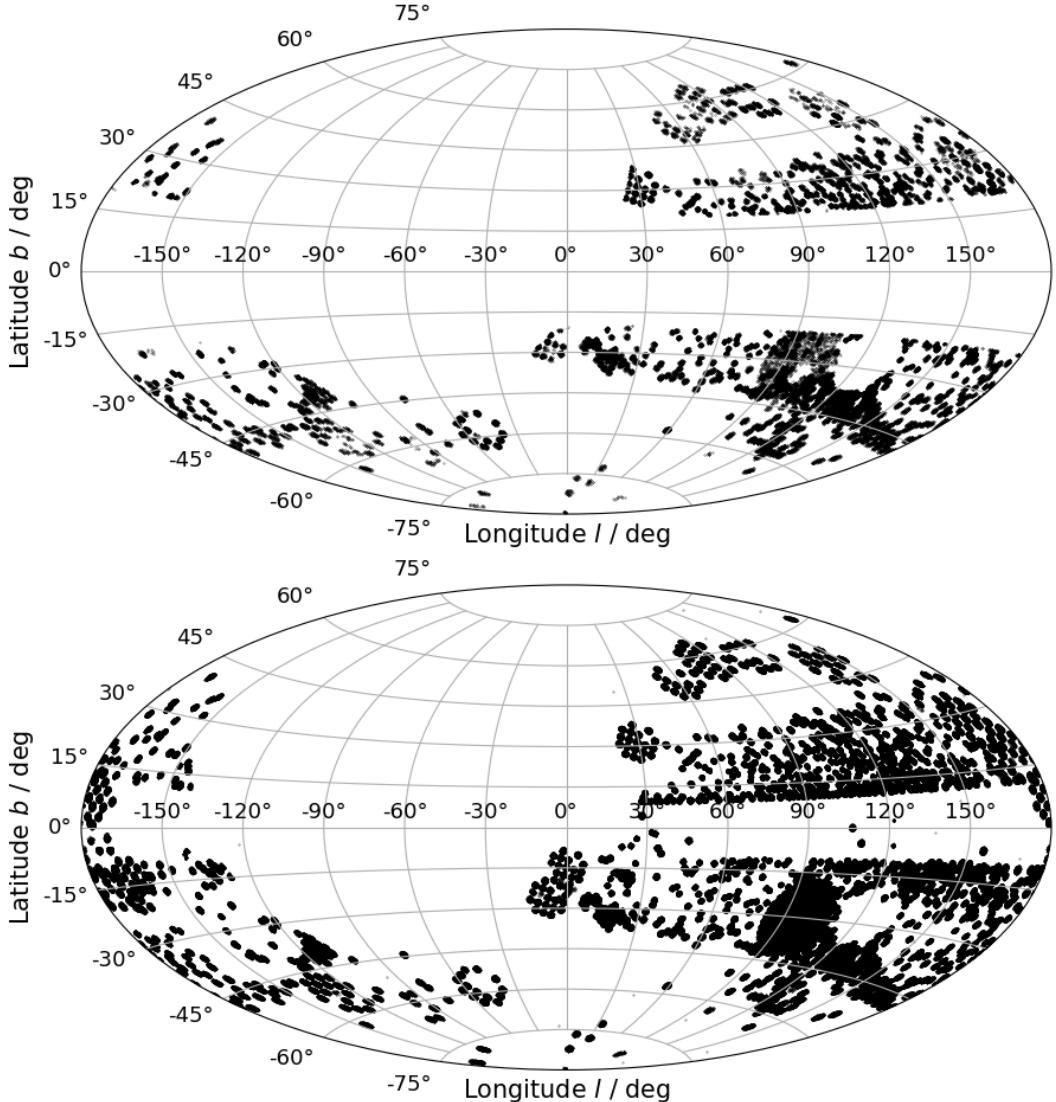


Figure 30: Comparison between the distribution of stars in galactic coordinates for the crossmatched sample and the full GALAH DR3 catalogue. Each of the black dots is one of the few thousand observation fields that make up the GALAH survey.

is  $\sigma = 0.60$ , but the number of stars in GALAH bluer than 0.15 mag or redder than 1.6 mag should be rather small, so it should suffice for the purpose of this check of the selection criteria. As for the absolute magnitude, it was simply calculated using the inverse parallax as distance, as the overwhelming majority of the stars has a well-measured parallax.

The final criterium that the metallicity be  $[\text{Fe}/\text{H}] < -0.4$  reduces the number of objects dramatically to only  $\sim 65\,000$ . Thus, this is the number of objects in GALAH DR3 that is expected to be observed, and with the number of crossmatched stars between 4MOST S2 and GALAH DR3 being in very good agreement at 63 000, there are apparently no surprises or unexpected selection effects at play. However, it should be noted that according to Figure 24, there is a systematic difference in metallicities between GALAH DR3 and the Pristine-Gaia-Synthetic estimates in the 4MOST S2 target catalogue, with GALAH metallicities tending to be higher. This means that the final metallicity cut made here would translate to a lower cut in the Pristine estimates, and the actual number to be expected should be higher than 65 000 if the cut of -0.4 for the 4MOST S2 catalogue was accurately represented. Properly taking this into account would be a challenging matter, and is thus not pursued further, but in any case,

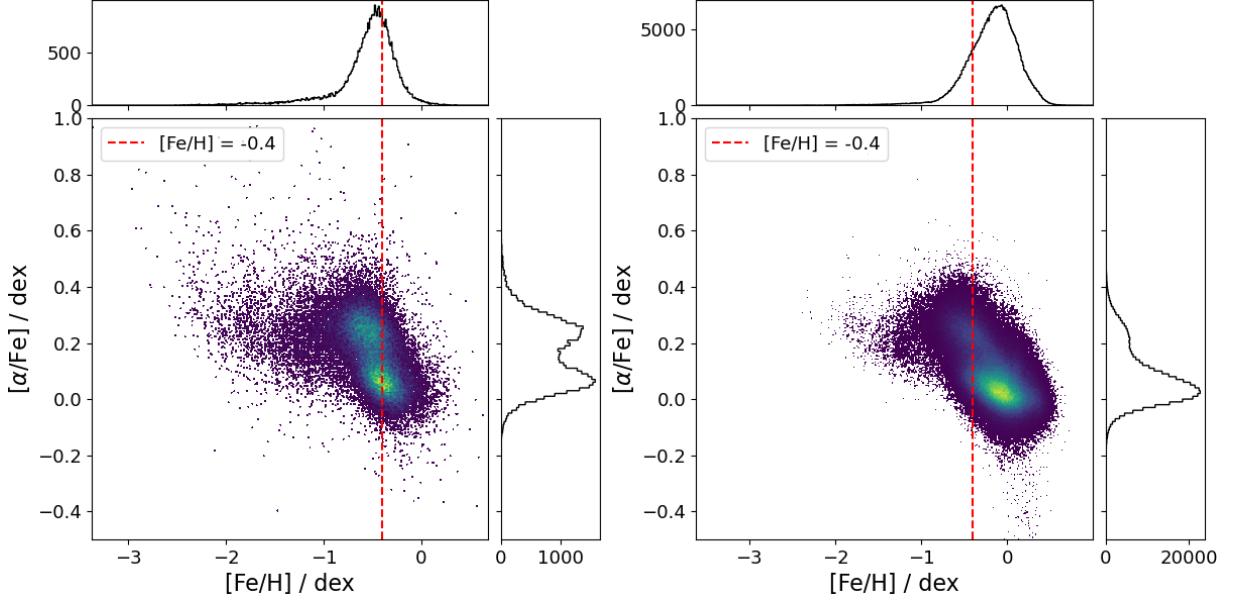


Figure 31:  $[\alpha/\text{Fe}] - [\text{Fe}/\text{H}]$  diagram for the crossmatch between GALAH DR3 and the 4MOST S2 target catalogue (left), and for the full GALAH DR3 catalogue (right). The histograms of the  $[\text{Fe}/\text{H}]$  and  $[\alpha/\text{Fe}]$  distributions are included in the plots opposite the corresponding axes.

the conclusion remains that the numbers are of similar order.

Making use of the  $\alpha$ -element abundances provided in GALAH DR3, an  $[\alpha/\text{Fe}] - [\text{Fe}/\text{H}]$  diagram is created for both the crossmatched sample and the full GALAH DR3 catalogue (Figure 31). For this, the flags for stellar parameters as well as the flags specific to  $[\alpha/\text{Fe}]$  and  $[\text{Fe}/\text{H}]$  are cleaned by requiring `flag_sp = 0`, `flag_fe_h = 0`, and `flag_alpha_fe = 0`. In the diagram of the full GALAH DR3 catalogue (right panel), two peaks are visible in the  $[\alpha/\text{Fe}]$ -distribution, with a large one around  $([\text{Fe}/\text{H}], [\alpha/\text{Fe}]) = (0, 0)$  and a small one which almost disappears in the tail of the first one. This second peak is much more pronounced in the diagram of the crossmatched sample (left panel), which is due to the first peak being significantly reduced as a result of the  $[\text{Fe}/\text{H}] < -0.4$  cut. It likely corresponds to the  $\alpha$ -knee, although its position is slightly below  $[\alpha/\text{Fe}] = +0.4$ . Since the  $[\alpha/\text{Fe}] - [\text{Fe}/\text{H}]$  diagram of a galaxy such as the Milky Way follows the characteristic form where  $[\alpha/\text{Fe}]$  is at a higher value at lower  $[\text{Fe}/\text{H}]$ , and the 4MOST S2 target catalogue is chosen specifically to contain a large number of low-metallicity stars, it is to be expected that it will contain comparably more objects with a higher  $\alpha$ -abundance when compared to a sample without such a selection criterium. In other words, the second peak is not well visible in the right panel in Figure 31 because all the metal-rich stars dominate, whereas the contrast between the peaks is reduced by a lot in the crossmatched sample.

The distributions of abundances for a few additional elements are plotted for the crossmatched sample and for the full GALAH DR3 catalogue in Figure 32. Although its shape looks similar overall, the distribution of carbon abundances seems to be shifted to higher values in the crossmatched sample, with the peak being located at  $\sim 0.2$  as opposed to 0, and the tail towards the highly C-enhanced stars becoming slightly more prominent. This is in line with the previously discussed increasing fraction of CEMP stars in metal-poor samples. While the distribution of barium abundances remains basically the same, the europium abundances appear significantly higher in the crossmatched sample. In fact, the plot displaying the full GALAH data appears to have a tiny protrusion in its distribution of europium abundances at  $\sim 0.25$ , whereas the peak in the crossmatched sample is located at approximately 0.25 to 0.3, suggesting that perhaps there are two components with different characteristic europium abundances in the GALAH data for different metallicity ranges. Furthermore, the distribution of the  $[\text{Ba}/\text{Eu}]$  distribution is shifted into the negative range, suggesting that the majority of stars in the crossmatched sample were enriched

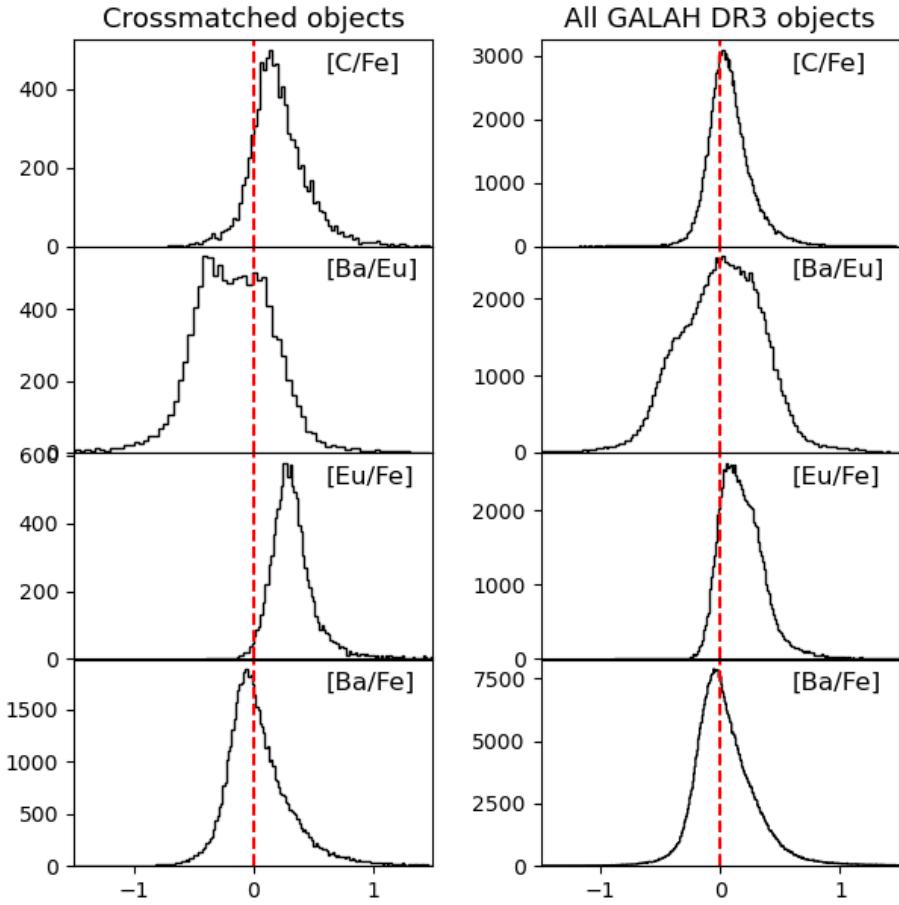


Figure 32: C, Eu, and Ba abundance distribution for the full GALAH DR3 catalogue (right) and the crossmatch between GALAH and the 4MOST S2 catalogue (left). The red line signifies an abundance of 0.0.

by the r-process, following the definitions laid out by [Beers & Christlieb \(2005\)](#) wherein  $[\text{Ba/Eu}] < 0$  and  $[\text{Eu/Fe}] \geq 0.3$  are indicators of r-process enrichment. As described in Section 1.1, it has also been found by many authors that there are clear signs of Eu-overabundance and elevated r-process enrichment in prominent substructures in the stellar halo of the Milky Way (see, e.g., [Matsuno et al. 2021](#); [Aguado et al. 2021](#); [Roederer et al. 2018](#)), and the dominance of r-process elements over s-process elements in metal-poor stars as a result of a later onset of the s-process has also been documented (see, e.g., [Simmerer et al. 2004](#); [Käppeler et al. 2011](#); [Battistini & Bensby 2016](#)).

Finally, the crossmatch with the SAGA database is also briefly investigated. As a significant portion of the Southern hemisphere is surveyed and 4MOST S2 targets halo stars in particular, most metal-poor stars are expected to be within its target catalogue, and thus a rough first guess would be that almost half of the SAGA stars should be identified in the crossmatch. However, only 1250 out of the 12 094 stars are identified in the target catalogue, due to which it is of interest to figure out for which reason so many stars in SAGA were not found in the crossmatch. In order to be able to apply the selection criteria of 4MOST S2 to SAGA, the photometric information has to be transformed into Gaia photometry, which is done using the photometric relations provided by the Gaia team ([Gaia \(E\)DR3 Documentation, Chapter 5.5.1](#); [Jordi 2018](#)). In particular,  $G - V = f(B - V)$  and  $G - V = f(V - I)$  are used to derive  $G$  for objects where either  $B - V$  and  $V$  or  $V$  and  $I$  are available, and the Gaia colour is derived via  $G_{\text{BP}} - G_{\text{RP}} = f(B - V)$  and  $G_{\text{BP}} - G_{\text{RP}} = f(V - I)$  if either  $B - V$  or  $V$  and  $I$  are available in SAGA.

Since the SAGA database is built up from many different sources, not all parameters necessary to apply all criteria are available for all stars, but the stepwise application of criteria to the catalogue is only mean-

ingful if all stars have all information, since otherwise a big drop in numbers after applying a criterium could be either due to an actual observational reason, or simply due to unavailable data. Thus, only the subset of stars for which all necessary information is available is considered, which after the photometric transformations is around 5600 or 46%. After applying the survey area conditions,  $\sim 3900$  stars are left, and after applying the apparent magnitude range of  $12.0 \leq G \leq 17.0$ , only  $\sim 700$  stars remain. After the application of the colour criterium,  $\sim 650$  stars are left, and the final metallicity cut  $[\text{Fe}/\text{H}] < -0.4$  leaves 623 stars in total, which is about  $p \approx 11\%$  of the starting number. Simply extrapolating this value that is based on only the subset with enough available parameters to the full SAGA database of 12 094 stars yields an expectation of  $p * 12\,094 = 1349$  stars to be identified in the 4MOST S2 target catalogue, which is only slightly more than the 1250 that have successfully been crossmatched. It should be noted that the colours are not dereddened and the absolute magnitude criterium getting rid of the lower section of the main sequence has not been taken into account, which might alter the final number of 1349 slightly, and as was the case for GALAH, the metallicity estimates given in SAGA and by the Pristine method show some deviations, making the metallicity cut another source of uncertainty in the final guess.

From the numbers in the previous paragraph it becomes clear that the large majority of stars in SAGA are not in the crossmatch due to the apparent magnitude criterium. This is perhaps counter-intuitive, as the majority of well-studied metal-poor stars that are expected to be collected by a database such as SAGA should be in the brighter magnitude ranges also covered by 4MOST S2. However, [Suda et al. \(2017\)](#) have extended the SAGA database by about 6000 stars situated within objects of our local group, most of which are at large distances with apparent magnitudes well below  $G = 17$  mag and will thus not appear in the 4MOST S2 survey, explaining the considerable drop in numbers. Overall, the results obtained in this section agree very well with the expectations, indicating a consistent selection process without undesired selection effects affecting the 4MOST S2 target catalogue.

#### 4.4 Spatial distribution: target catalogue vs. Besançon Galaxy model

The Besançon Galaxy model is a population synthesis model that has been built in order to aid in developing scenarios of galaxy formation and evolution, and investigate particular topics such as star formation rates and the initial mass function. It provides star counts based on observational constraints from large datasets such as Hipparcos or Tycho-2, where the Galactic potential is calculated taking into account dynamical self-consistency ([Robin et al. 2003; Czekaj et al. 2014](#)). However, for the Gaia mission in particular, the Besançon Galaxy model star counts are based on the Gaia Universe model snapshot (GUMS; [Robin et al. 2012a](#)), which is a simulation carried out in order to analyse the expected objects and characteristics to be observed during the Gaia mission, and makes use of the Gaia Object Generator (GOG; [Luri et al. 2014](#)). Since [Robin et al. \(2003\)](#), a number of significant updates have been made to the models of the bar population ([Robin et al. 2012b](#)), the thick disk and halo populations ([Robin et al. 2014](#)), the thin disk population ([Czekaj et al. 2014](#)), and the kinematics ([Robin et al. 2017](#)), amongst other changes.

In order to get an overview of the spatial distribution of the 4MOST S2 target catalogue, and to compare this distribution to model predictions, the distribution of stars in the  $z$ - $R$ -plane is employed for the cartesian galactocentric coordinates  $R = \sqrt{x^2 + y^2}$  and  $z$ . This is done for the 4MOST S2 target catalogue and the Besançon model after applying the 4MOST S2 selection criteria to it to simulate the survey. For the 4MOST S2 catalogue, only stars at  $\delta < 5^\circ$  will be retained in the analysis from this point on, which leaves just about 3.8 million objects. To be able to transform the coordinates to the galactocentric frame, it is necessary to have three-dimensional information  $(\alpha, \delta, d)$  on the stars, so the distance  $d$  is determined from the zero-point-corrected parallax similarly to Section 4.1. The transformation  $(\alpha, \delta, d) \rightarrow (x, y, z)_{\text{GC}}$  is then carried out using the ASTROPY.COORDINATES library of the Python package ASTROPY ([Astropy Collaboration et al. 2013, 2018, 2022](#)), with the frame parameters of the conversion being set to  $R_\odot = 8122$  pc ([GRAVITY Collaboration et al. 2018](#)) and  $z_\odot = 20.8$  ([Bennett & Bovy 2019](#)).

To simulate the expected star counts of the 4MOST S2 survey, the input forms on the Besançon Galaxy model website<sup>16</sup> are used. Here, the galactic latitude range  $|b| > 20^\circ$  and the apparent magnitude range of  $12.0 \leq G \leq 17.0$  are implemented, and the distance is set to be in the range between 0 and 20 kpc, yielding a total of around 24 million stars. As the Besançon model website only calculates up to 2 million objects at once, the survey area was divided up into 16 sections in order not to exceed that limit in any section. As the output does not provide the Gaia colour  $G_{\text{BP}} - G_{\text{RP}}$ , it is estimated using the photometric relation for  $G_{\text{BP}} - G_{\text{RP}} = f(J - K)$  from the [Gaia \(E\)DR3 Documentation, Chapter 5.5.1](#), as  $J - K$  is provided. The absolute  $G$  band magnitude is computed via  $M_G = G + 5 - 5 \log d$ , where  $d$  is the distance provided in the output (converted into pc).

The metallicity is also given, albeit in the form of  $[\text{M}/\text{H}]$ , and thus all the remaining cuts  $-80^\circ \leq \delta \leq 5^\circ$ ,  $0.15 \leq G_{\text{BP}} - G_{\text{RP}} \leq 1.10$ ,  $1.10 < G_{\text{BP}} - G_{\text{RP}} \leq 1.60$  and  $M_G < 3.5$ , as well as  $[\text{M}/\text{H}] < -0.4$  can be applied now. In order to account for the fact that the  $G$  values given in the model's output are not considering extinction effects, the  $G$  band extinction  $A_G$  is calculated for every galactic coordinate pair  $(l, b)$  the same way as in Section 4.1 using the Total Galactic Extinction map in Gaia DR3 ([Delchambre et al. 2023](#)). However, as the colour  $G_{\text{BP}} - G_{\text{RP}}$  is not directly provided, the parameter  $T_{\text{eff}}$  is instead used to derive  $k_G$  and thus  $A_G = k_G A_0$ . Then, all objects for which  $G + A_G > 17$  mag are discarded to finally yield the Besançon Galaxy model prediction of the 4MOST S2 survey, which contains about 5.1 million objects. Since the cartesian galactocentric coordinates  $(x, y, z)$  are contained in the model output,  $R$  and  $z$  are immediately available.

The distribution of stars in the galactocentric  $z-R$ -plane is plotted for the full sample and each of the three sub-surveys, with the 4MOST S2 target catalogue distributions being shown side by side with the Besançon model plots (Figure 33). The appearance of the plots on the left and right is overall similar, but particularly in the top and bottom panels, it seems that the Besançon model has many more stars in the direction of  $R = 0$  and less at  $R > 10$  kpc. Also, it should be noted that the colourbar scales are different, with the plots of the Besançon model containing quite a few more objects, which is reflected by the total numbers of 3.8 million in the 4MOST S2 target catalogue vs. 5.1 million in the Besançon simulation thereof. While ideally, the immediate region around the Sun should not be too crowded, as the survey targets the halo, it can be seen that a majority of the objects is still rather close to the Sun, as the object density there greatly exceeds the density in the other parts of the plot. The same plots are repeated in Figure 34, but only for stars at  $[\text{Fe}/\text{H}] < -1.0$  (or  $[\text{M}/\text{H}] < -1.0$  for Besançon). The agreement between 4MOST S2 and the Besançon model appears to be better here, in particular concerning the absolute numbers, with the colourbars on the left and right panels being almost identical in each row. However, it still seems as though the Besançon model includes more stars towards  $R = 0$  in the bottom panel, and less at higher  $R > 10$  kpc in the top and bottom panels.

The distribution in  $z$  of the stars in the range  $7 \text{ kpc} < R < 9 \text{ kpc}$  is plotted for 4MOST S2 and the Besançon model as well (Figure 35). The plots show that while the form of the distribution looks similar above and below  $z$ , there are less stars at positive  $z$ , and while the total number of stars is increased by a lot for the Besançon model, the difference between the two peaks is similar in both panels. In particular, the distributions look identical in the tail towards large  $|z|$ . This is also reflected in the rough numbers: At  $|z| > 3$  kpc, the numbers are  $\sim 162\,000$  stars in the 4MOST S2 catalogue vs.  $\sim 194\,000$  in the Besançon model, which is a much better agreement than the complete sample's 3.8 million vs. 5.1 million. One likely reason why the total number of stars is so much higher for the Besançon model is the uncertainty associated with the metallicity cut. As was seen in Figure 22, a difference in even 0.1 dex in the range around  $[\text{Fe}/\text{H}] = -0.4$  dex means the inclusion or removal of at least a few hundred thousand stars. For instance, setting the cut at  $[\text{M}/\text{H}] = -0.5$  instead for the Besançon model drops the total number from 5.1 million to 4.3 million. As the number at  $|z| > 3$  kpc changes from  $\sim 194\,000$  to  $\sim 186\,000$  – which is a much less dramatic change even when considering the different total sizes – this change is clearly

---

<sup>16</sup><https://model.obs-besancon.fr/>, retrieved 22.10.2023

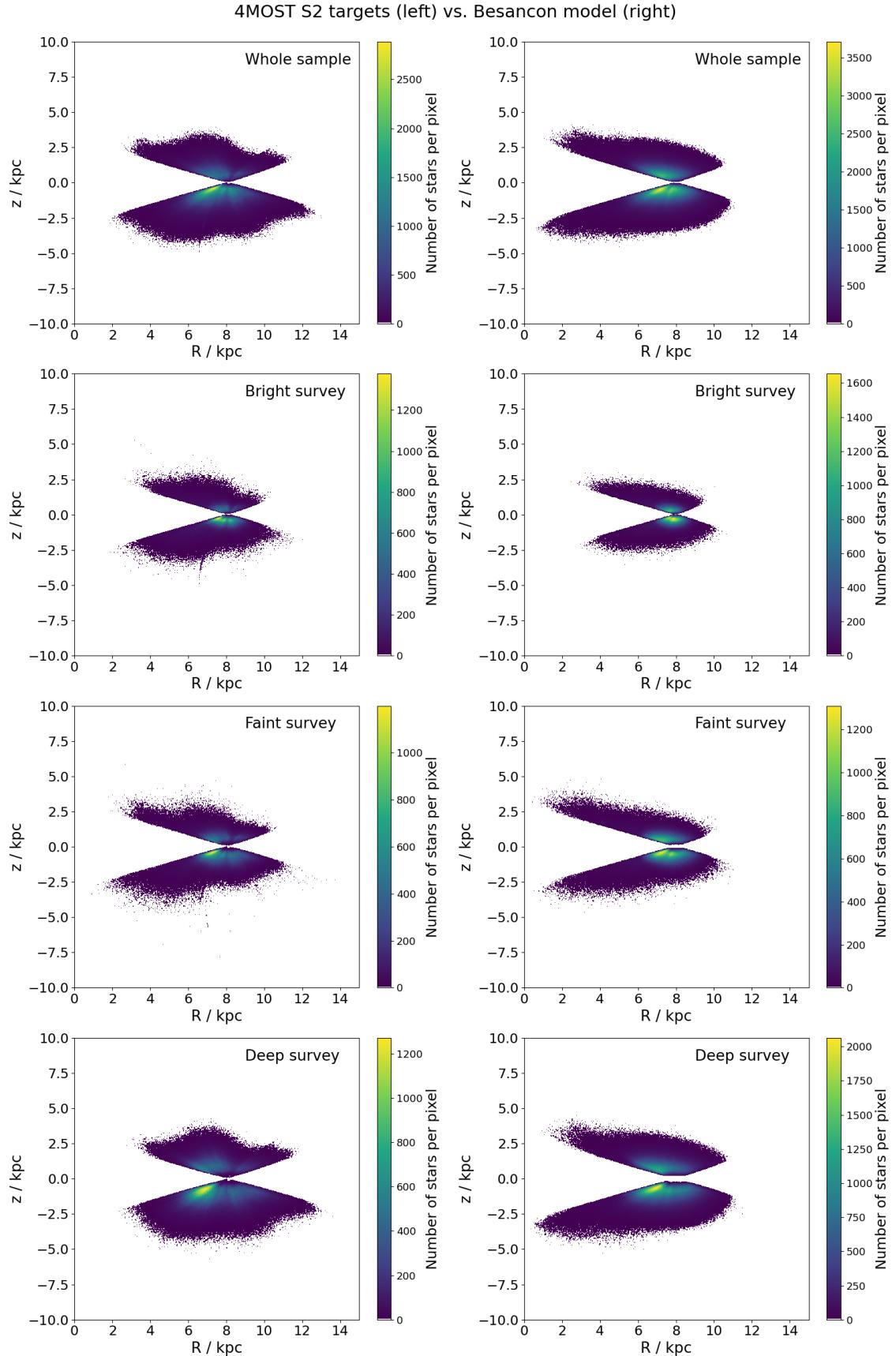


Figure 33: The spatial distribution in galactocentric coordinates  $z$  vs.  $R$  of the whole sample and the three subsurveys for the 4MOST S2 target catalogue (left) and the Besançon model simulation of the catalogue (right). Note the higher values on the colourbars on the right, due to the higher total number of stars in the Besançon model.

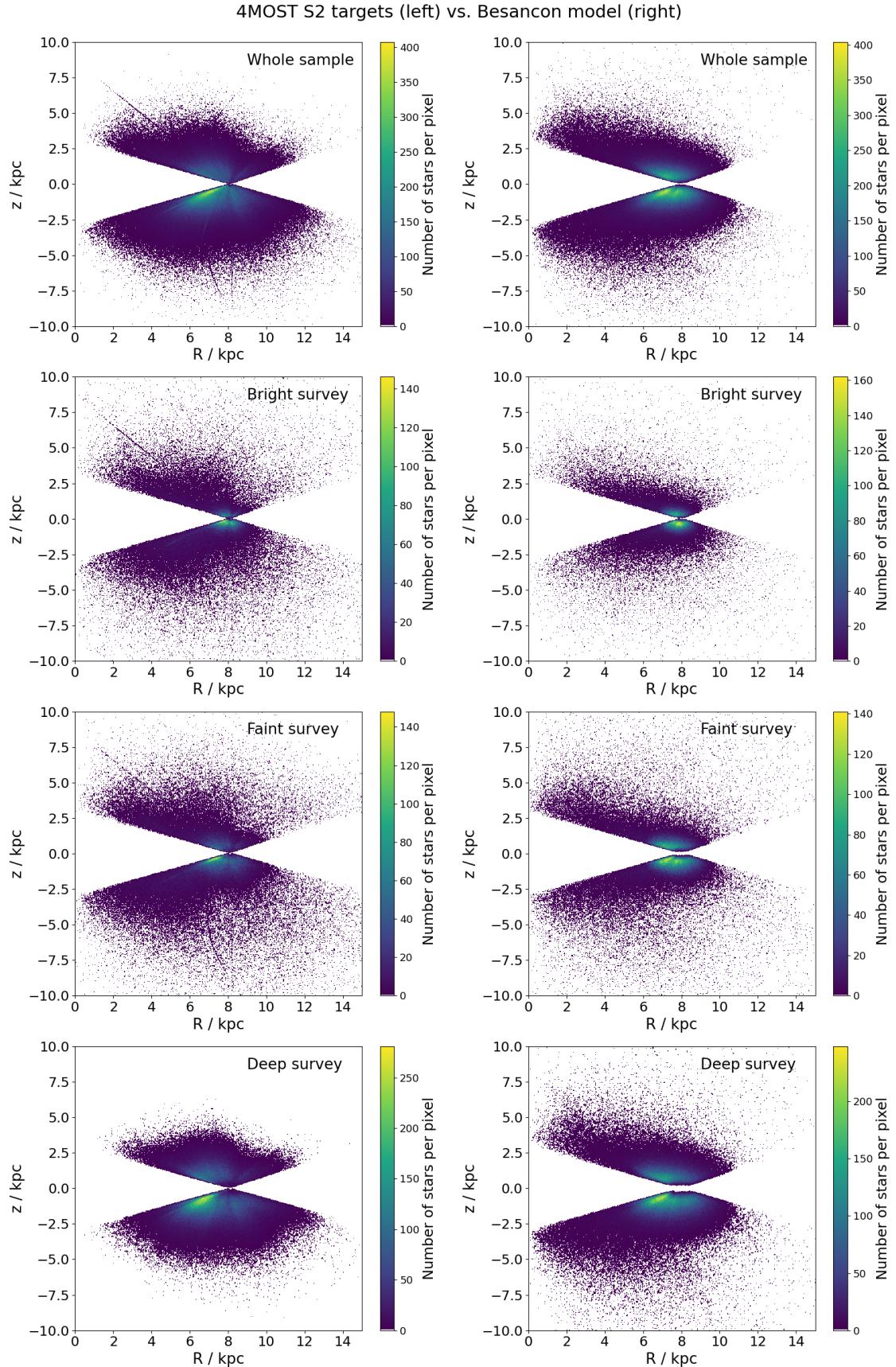


Figure 34: The spatial distribution in galactocentric coordinates  $z$  vs.  $R$  at  $[\text{Fe}/\text{H}] < -1.0$  (or  $[\text{M}/\text{H}] < -1.0$ ) for the 4MOST S2 target catalogue (left) and the Besançon model simulation of the catalogue (right). The colourbars are chosen at the same scale in each row.

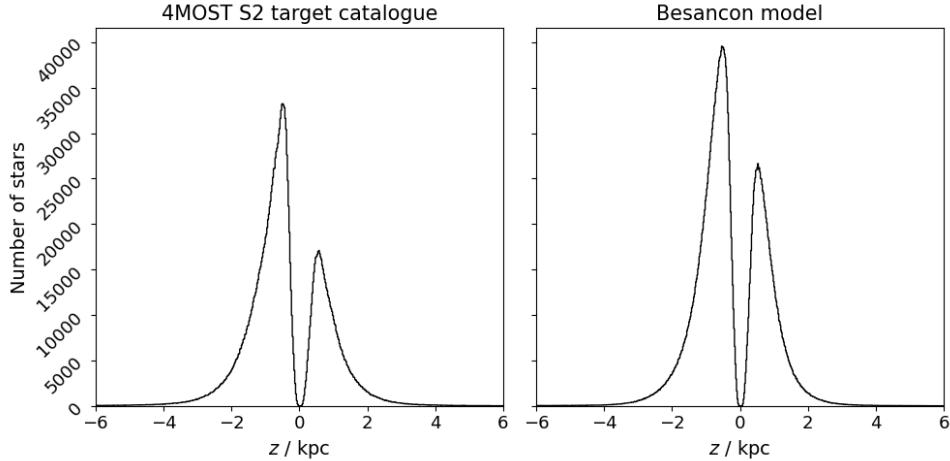


Figure 35: Distribution in the galactocentric  $z$ -coordinate for the slice  $7 \text{ kpc} < R < 9 \text{ kpc}$  for the 4MOST S2 target catalogue (left) and the Besançon model (right).

primarily happening at the inner parts of the  $z$ - $R$ -diagram close to the Sun. This is also in line with Figure 34 (which includes only objects at  $[\text{Fe}/\text{H}] < -1.0$ ) showing better agreement between 4MOST S2 and Besançon compared to the full samples, indicating that the deviation in numbers stems mainly from the higher-metallicity stars, which tend to be disk stars at small  $z$ . In particular, it should also be noted that [Martin et al. \(2023\)](#) stress that they placed the main effort into optimising the metallicities of stars with low-metallicities, and that stars at  $[\text{Fe}/\text{H}] > -1.0$  may be affected by unexpected biases.

The explanation for the generally increased number of stars at low  $R$  for the Besançon model, particularly visible in some panels in Figure 33, might be that in the direction towards the Galactic centre, the extinction due to increased amounts of dust and stellar density can become quite high. Perhaps, this extinction was underestimated in those regions, and many objects that were kept in the Besançon model reference are actually outside the apparent magnitude requirement of 4MOST S2 when their proper extinction is considered. Other sources of error for the comparison include the colour transformation that was done to derive  $G_{\text{BP}} - G_{\text{RP}}$  for the Besançon model, which has a statistical uncertainty of  $\sigma = 0.1 \text{ mag}$  associated with it, and the fact that the colour criterium was then applied to  $G_{\text{BP}} - G_{\text{RP}}$  directly instead of  $(G_{\text{BP}} - G_{\text{RP}})_0$ . Furthermore, the Besançon model provides a smooth galaxy model which does not account for any inhomogeneities. The Besançon model coordinates are given in epoch J2000.0, whereas the Gaia DR3 coordinates are J2016.0, but this should not make a difference to the analysis here, as only the large-scale distribution of stars is of interest.

To make the overall comparison of the  $z$ - $R$ -diagrams more quantitative, the  $z$ - $R$ -plane is divided up into small cells of size  $0.5 \text{ kpc}$  by  $0.5 \text{ kpc}$  for  $-20 < z < 20 \text{ kpc}$  and  $0 < R < 20 \text{ kpc}$ . The number of stars contained in a cell is counted for the 4MOST S2 target catalogue and for the Besançon model, and then the fraction  $\frac{\#\text{stars Besançon}}{\#\text{stars 4MOST S2}}$  is calculated for that cell. This is done for all cells covering the aforementioned range, and the result is plotted in Figure 36, with the colour-coding indicating the values of the fractions of star counts in each cell on a logarithmic scale. A value close to 1 means good agreement, whereas a large value means that the Besançon model predicts more values, and a value between 0 and 1 means that there are more stars of 4MOST S2 in that cell. In case there are no 4MOST S2 stars in a cell, it is left empty, as there would be a division by 0 otherwise. The plot is shown for both the full samples, and including only stars at  $[\text{Fe}/\text{H}] < -1.0$  (or  $[\text{M}/\text{H}] < -1.0$  for Besançon).

In Figure 36a, the colour-coding clearly indicates the comparably increased number of stars in the Besançon model at low  $R$ , with the numbers showing from a few times to more than 10 times more stars in the Besançon model than 4MOST S2 there. There is also an increased number of bright cells at

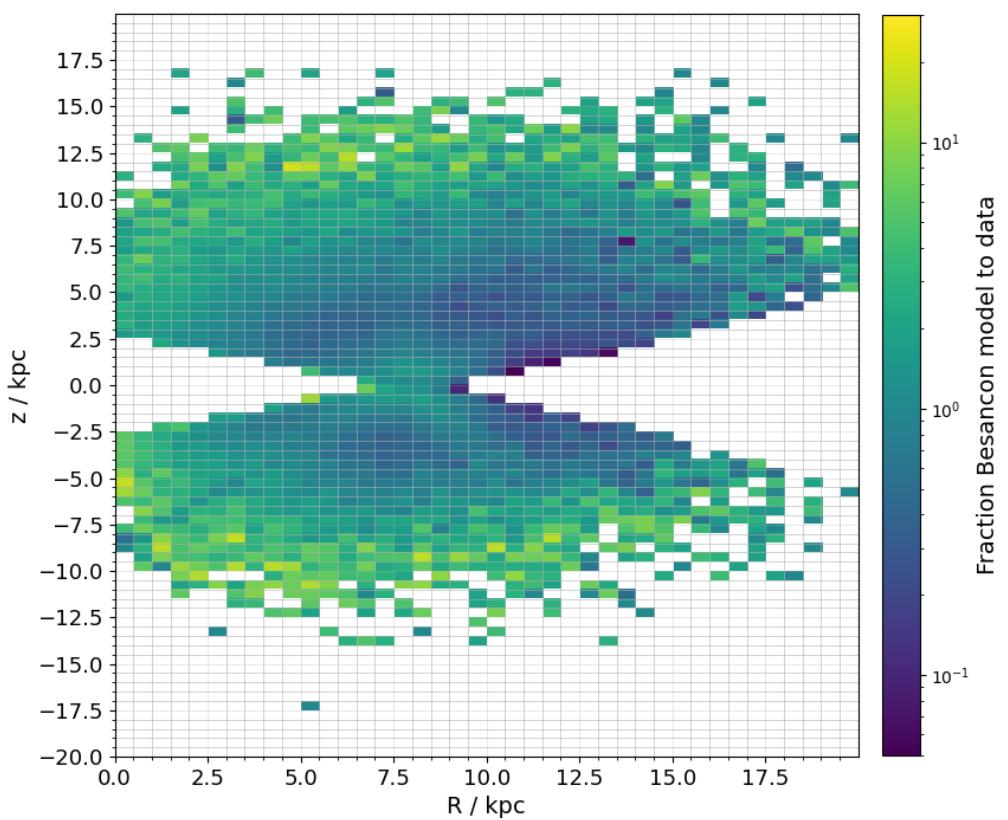
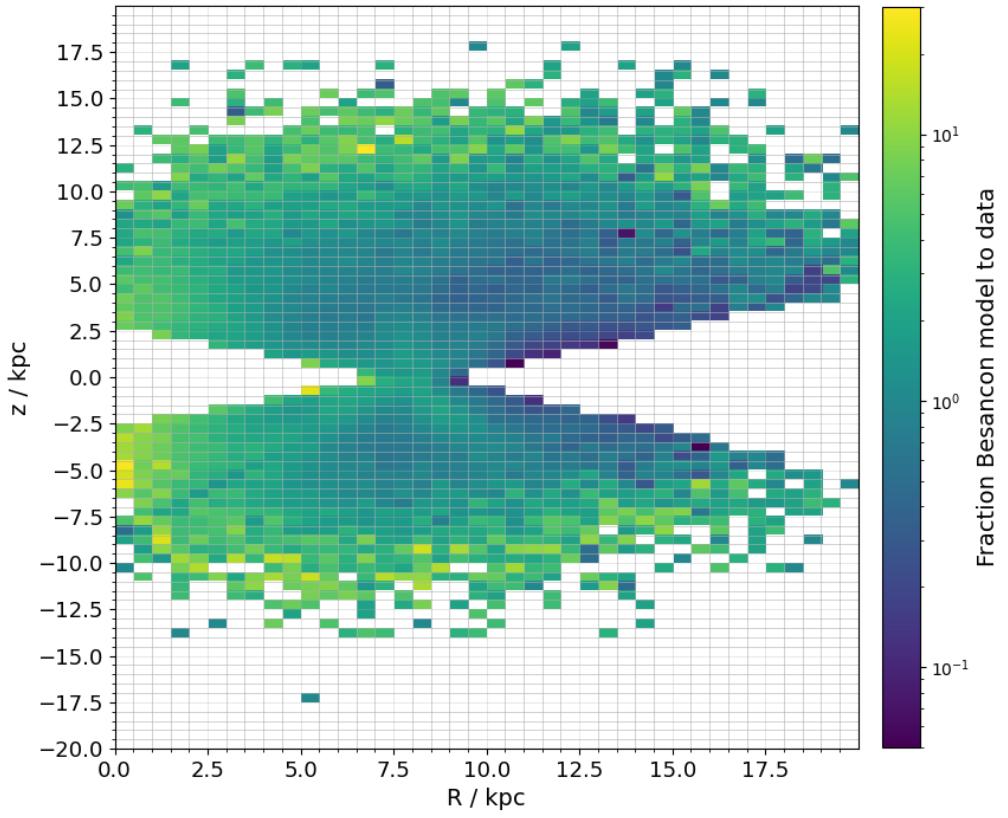


Figure 36: Star counts in the Besançon model devided by star counts in 4MOST S2 for each cell. The colour scale is logarithmic in order to give the same weight to deviations in both directions from 1.

the highest  $|z|$ -values, although the plot becomes less meaningful towards the edges due to the very small numbers of stars leading to random effects becoming more dominant. However, this still implies that at very large heliocentric distances of  $\sim 9$  kpc and beyond, the Besançon model predicts slightly more stars than are in the 4MOST S2 sample. Since the extinction is expected to be more significant towards the Galactic centre and at large distances, these findings also support the explanation that one of the reasons for the lower number of stars in 4MOST S2 is related to extinction effects, with the extinctions applied to the Besançon model data above perhaps being not entirely sufficient. Finally, the Besançon model also seems to be more populated at the centre of the distribution around the Sun.

As is to be expected from the visual appearance of Figure 34 already, the colour-coding in Figure 36b (which includes only objects at  $[Fe/H] < -1.0$ ) indicates a much better agreement at low  $R$  than Figure 36a, although the trend that at very large distances there are less stars in the 4MOST S2 catalogue than the Besançon model is also present, here from heliocentric distances of  $\sim 7.5$  kpc onwards. As before, the star counts per cell are very small at these distances though, so the comparison is less reliable. The increased fraction at the centre is also reduced here. It is notable that in the ranges of  $|z| \sim 2.5$  to 5 kpc at metallicities below  $[Fe/H] = -1.0$ , there are actually more stars in the 4MOST S2 target catalogue than in the Besançon model, and the total numbers below  $-1.0$  are  $\sim 960\,000$  for 4MOST S2 vs.  $\sim 718\,000$  for the Besançon model, which could be a further sign that the metallicity scales may not be entirely consistent with each other, leading to a more uncertain comparison overall. In summary, the agreement in distribution between 4MOST S2 and the Besançon model is satisfactory outside of small  $R$  and the Solar vicinity, and the overall agreement improves when including only lower-metallicity stars, and also at higher  $z$ -values up to  $8\sim 9$  kpc, from which point onwards the small number of objects makes it difficult to make a meaningful comparison. However, the total numbers are quite different for the full datasets at 3.8 million vs. 5.1 million, the reasons for which are most likely related to metallicity uncertainties and extinction effects. Since 4MOST S2 is particularly interested in metal-poor stars, it is reassuring that the Besançon model prediction star count does not exceed 4MOST S2 at the low-metallicity end, and that the comparison at low metallicities is improved with respect to the full comparison.

## 5 Chemical and kinematic analysis

### 5.1 The 6D-Gaia-data subset of 4MOST S2

Gaia DR3 contains the radial velocities of 33 million stars measured by the Gaia RVS system, thus making available an extensive 6-dimensional dataset suited to a variety of studies into the kinematics of the Galaxy. For the 4MOST S2 target catalogue of 3.8 million stars, 909 496 have radial velocity measurements by Gaia, which is about 24%. This 6D data subset is used in the following analysis. As seen in Figure 37 depicting the distribution of radial velocities by sub-survey, the radial velocities are distributed symmetrically around  $\sim 0$  km/s and are almost entirely within  $\pm 200$  km/s. Notably, most of this 6D subset is part of the bright sub-survey, with the rest being part of the faint sub-survey, and the deep sub-survey not being represented at all. This is also shown in the right panel, where the distribution of apparent magnitudes  $G$  is plotted. Here, the completeness limit of the Gaia RVS instrument at  $G \sim 14.5$  mag can be seen, with an accelerating rise in the number of stars due to the growing volume until that limit is reached, at which point the rise breaks and the number plummets. Beyond  $G = 15.5$  mag in the deep sub-survey, no stars are left. This means that the following analysis pertains only to the bright end of the 4MOST S2 target catalogue and cannot be used to draw conclusions about the faint end.

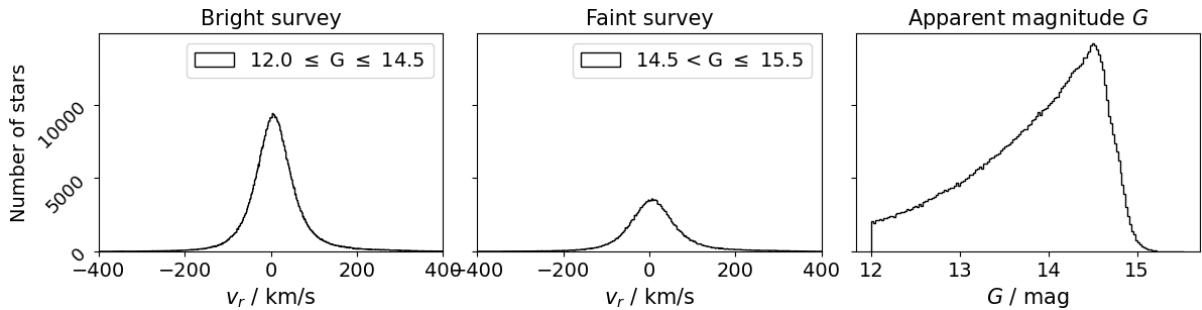


Figure 37: The two panels on the left show the distribution of the radial velocity  $v_r$  in the bright and faint sub-surveys. The panel on the right shows the apparent magnitude distribution of the 6D subset. No stars in the deep sub-survey have radial velocity measurements by the Gaia RVS instrument.

The radial velocity uncertainties in Gaia DR3 have been determined via two methods depending on the star's brightness ([Gaia DR3 Documentation, Chapter 20.1.1](#)): For bright stars, they are simply calculated from the uncertainty on the median of the epoch radial velocities, including the calibration floor contribution. For faint stars, the cross-correlation function and its second derivative are used to determine the radial velocity uncertainty following the approach by [Zucker \(2003\)](#). In the right panel of Figure 38, a histogram of the radial velocity uncertainties is shown, and with 95% of stars having a radial velocity uncertainty of less than 10.1 km/s, the majority of the 6D subset has moderately accurate radial velocity measurements. The two left panels depict the distribution of radial velocity uncertainties across magnitudes  $G$  and  $G_{RP}$ . The plot shows similar behaviour to Figure 4 in [Babusiaux et al. \(2023\)](#) of all Gaia DR3 radial velocities, with the bulk of the stars having a very small uncertainty at bright magnitudes which is slowly rising to around 5 km/s at the completeness limit of  $\sim 14.5$  mag. There are notable outliers at  $12\sim 13$  mag and at  $14\sim 15$  mag going up to very high uncertainties, which are much more pronounced in their figure, visible in the colour coding as two vertical columns. To remove these outliers and the stars without a good radial velocity measurement in general, only stars with radial velocity uncertainty  $\sigma_{v_r} \leq 10$  km/s are retained for the analysis.

Despite the radial velocity uncertainties being smaller than 10 km/s for the vast majority of the 6D subset, it is still interesting to verify the radial velocities against external reference datasets with more precise measurements. A crossmatch with GALAH DR3 and APOGEE DR17 has already been carried out within the scope of Section 4.2, and a crossmatch to RAVE DR6 ([Steinmetz et al. 2020b,a](#)) is carried out in the same manner as for APOGEE. In the 6D subset, the crossmatches contain 55 043, 14 173, and 15 649 for the three references, respectively. While the median of the radial velocity uncertainty is about

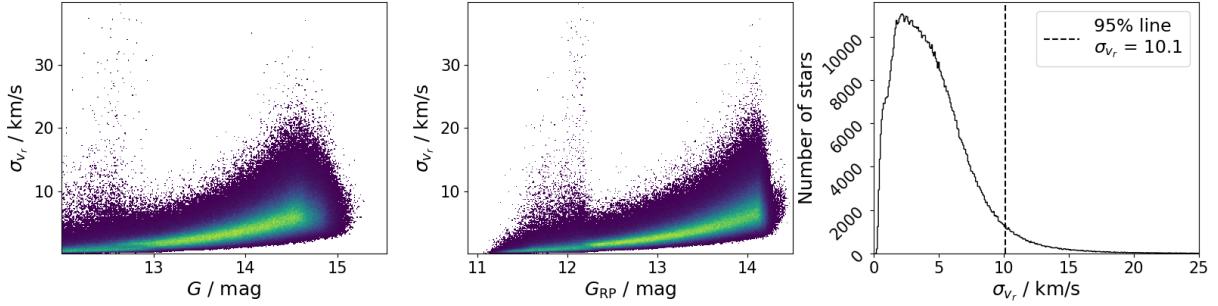


Figure 38: Histogram of the radial velocity uncertainties (right) and the radial velocity distribution across the Gaia magnitudes  $G$  and  $G_{\text{RP}}$ .

4.0 km/s for the Gaia radial velocities in the 6D subset, it is 0.13 km/s for the GALAH radial velocities in the crossmatch, 0.06 km/s for APOGEE, and 1.7 km/s for RAVE, meaning that all three of the reference catalogues are largely more precise than the Gaia radial velocities, but they also contain two orders of magnitude less stars.

The first three panels of Figure 39 depict a comparison of the Gaia DR3 radial velocities to the references by plotting the radial velocity estimates of Gaia against the three reference radial velocity estimates. In all three, the agreement is very good, and only around 0 are there an increased number of stars deviating visibly from the identity line. This is because the radial velocity distribution peaks at around 0 (see Figure 37), meaning that this region contains by far the most data point, so increased scatter is expected. The bottom right panel shows histograms of the  $\sigma$ -deviations between the radial velocities of Gaia and the references, where

$$\sigma_{\text{Gaia-ref.}} = \frac{|v_{r,\text{Gaia}} - v_{r,\text{ref.}}|}{\sqrt{\sigma_{v_{r,\text{Gaia}}}^2 + \sigma_{v_{r,\text{ref.}}}^2}}. \quad (3)$$

It confirms the tight correlation seen in the other three panels, with almost all stars having radial velocity estimates from Gaia and the references that are within  $2\sigma$  of each other, although it should be noted that the comparatively large uncertainties of Gaia will typically dominate in the formula above, as can be seen from the median uncertainties listed in the previous paragraph.

## 5.2 Kinematic parameter determination and selection of halo

In order to determine the parameters relevant to the kinematics of the stars, in particular the integrals of motion such as angular momentum or energy, it is necessary to convert the coordinates to galactocentric cartesian coordinates,  $(\alpha, \delta, d, \mu_\alpha, \mu_\delta, v_r) \rightarrow (x, y, z, v_x, v_y, v_z)_{\text{GC}}$ , where  $d$  is a distance estimate. In order to have reliable coordinates, the analysis is limited to stars which have radial velocity uncertainties  $\sigma_{v_r} < 10$  km/s and a relative parallax uncertainty of  $\frac{\pi}{\sigma_\pi} \geq 3$ , in which case the distances  $d$  are determined via the inverse parallax. Doing this removes around 8% of the 6D subset.

The transformation of the spatial coordinates is carried out using the ASTROPY transformation to the Galactocentric Coordinate Frame<sup>17</sup>. This transforms from ICRS to Galactocentric coordinates by first expressing the ICRS coordinates as cartesian instead of spherical, then rotating the  $x$ -axis such that it points towards the centre of the Galaxy, rotating around  $x$  by an extra angle to have the normal of the Galactic plane point to the North pole, then subtracting the distance from the Galactic centre in  $x$ -direction, and finally rotating around the final  $y$ -axis such that the height of the Sun above the Galactic midplane is reproduced. The default parameters of this transformation can be set manually. In this analysis, the galactocentric distance of the Sun  $R_\odot = (8122 \pm 31)$  pc by GRAVITY Collaboration et al.

<sup>17</sup><https://docs.astropy.org/en/stable/coordinates/galactocentric.html#coordinates-galactocentric>, retrieved 22.10.2023

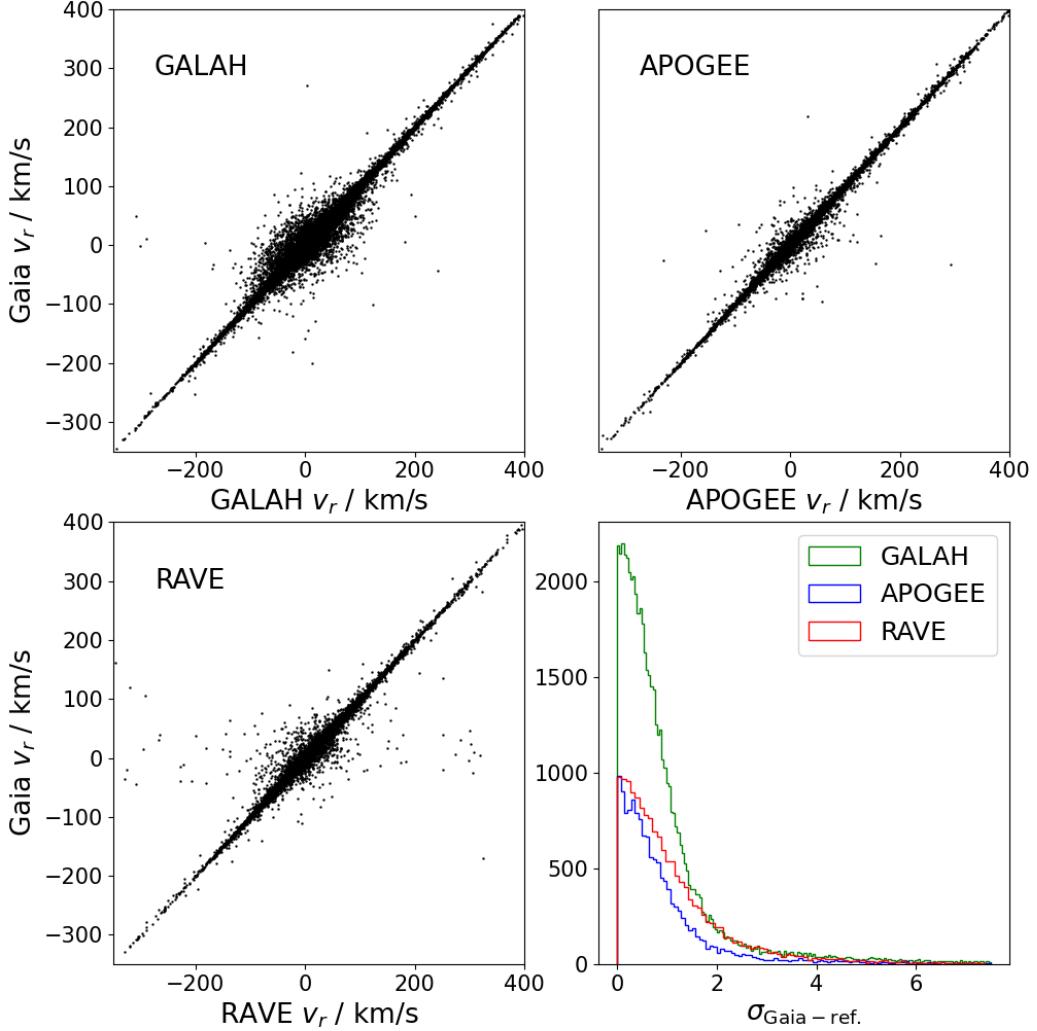


Figure 39: Gaia DR3 radial velocity plotted against reference (GALAH, APOGEE, RAVE) radial velocity (first three panels). The bottom right panel shows histograms of the  $\sigma$ -deviations between Gaia data and the references.

(2018) and the height above the midplane of the Sun  $z_\odot = (20.8 \pm 0.3)$  pc by Bennett & Bovy (2019) are used.

For the velocities, the transformations presented by Johnson & Soderblom (1987) are used:

$$\begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = \mathbf{T} \cdot \mathbf{A} \cdot \begin{pmatrix} v_r \\ k\mu_{\alpha*}d \\ k\mu_\delta d \end{pmatrix}, \text{ where} \quad (4)$$

$$\mathbf{T} = \begin{pmatrix} \cos l_C & \sin l_C & 0 \\ \sin l_C & -\cos l_C & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -\sin \delta_G & 0 & \cos \delta_G \\ 0 & -1 & 0 \\ \cos \delta_G & 0 & \sin \delta_G \end{pmatrix} \begin{pmatrix} \cos \alpha_G & \sin \alpha_G & 0 \\ \sin \alpha_G & -\cos \alpha_G & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (5)$$

$$\mathbf{A} = \begin{pmatrix} \cos \alpha \cos \delta & -\sin \alpha & -\cos \alpha \sin \delta \\ \sin \alpha \cos \delta & \cos \alpha & -\sin \alpha \sin \delta \\ \sin \delta & 0 & \cos \delta \end{pmatrix}. \quad (6)$$

Here,  $\mu_{\alpha*} = \mu_\alpha \cos \delta$  denotes the  $\delta$ -corrected proper motion in right ascension,  $k = 4.74057$  is the velocity in one astronomical unit per tropical year corresponding to km/s,  $\alpha_G = 192.85948^\circ$  and  $\delta_G = 27.12825^\circ$  are the right ascension and declination of the North Galactic Pole, and  $l_C = 122.93192^\circ$  is the Galactic longitude of the North Celestial Pole (Perryman & ESA 1997). For the peculiar motion of

the Sun, the values by Schönrich et al. (2010) of  $(U, V, W)_\odot = (11.1^{+0.69}_{-0.75}, 12.24^{+0.47}_{-0.47}, 7.25^{+0.37}_{-0.36})$  km/s are corrected for, whereas the rotation of the Sun is taken from McMillan (2017) to be 232.8 km/s in  $y$ -direction.

When viewed from the North Galactic Pole, the Milky Way rotates clockwise. Due to this, kinematic analyses are usually carried out in a left-handed galactocentric coordinate system, as the unit vector in  $z$ -direction  $\hat{e}_z$  would have to point towards the South Galactic Pole otherwise. However, the frame described above and ASTROPY frames in general are right-handed coordinate systems. Thus, the  $x$ -axis after the transformation is flipped in the following analysis by setting  $x \rightarrow -x$  and  $v_x \rightarrow -v_x$  to make the system left-handed, ensuring that the plots are in line with similar plots in the literature.

From the 6D galactocentric cartesian coordinates  $(x, y, z, v_x, v_y, v_z)$ , a number of relevant parameters can be calculated directly, including the total radial distance from the centre  $r = \sqrt{x^2 + y^2 + z^2}$ , the cylindrical distance  $R = \sqrt{x^2 + y^2}$ , the absolute space velocities  $v = \sqrt{v_x^2 + v_y^2 + v_z^2}$ , and finally the specific angular momentum  $\mathbf{L} = \mathbf{r} \times \mathbf{v}$ , where  $\mathbf{r} = (x, y, z)^\top$  and  $\mathbf{v} = (v_x, v_y, v_z)^\top$ . The specific orbital energy is given by

$$E = \Phi + v^2/2, \quad (7)$$

where  $\Phi$  is the gravitational potential. While the Gaia mission supplies precise measurements of the position, parallax and proper motion for stars in our Galaxy, it cannot measure any acceleration, meaning that in order to know the stars orbits, the Galaxy's gravitational potential must be established.

McMillan (2017) constructs the Milky Way's gravitational potential from mass models which are based on both observations and theoretical predictions. Due to their differing structure, the Galaxy is split into its several constituents, namely the bulge, dark matter halo, thin stellar disk, thick stellar disk, HI gas disk, and molecular gas disk, for each of which axial symmetry is assumed. For the bulge, the density profile follows Bissantz & Gerhard (2002) but with an axisymmetric approximation, while both the thin and thick stellar disks follow an exponential density profile. The molecular and HI disks' density profiles also fall off exponentially, but there is a low-density hole in the centre, and for the dark-matter halo, McMillan (2017) assumes a Navarro-Frenk-White profile (Navarro et al. 1996). The parameters of the models are determined by fits using Bayesian statistics, with a Markov Chain Monte Carlo method being employed due to the high number of parameters. It should be noted that in the inner parts of the Galaxy, the model is less reliable due to deviations from axisymmetry such as the bar. The Python package GALPOT<sup>18</sup> (McMillan 2017; Dehnen & Binney 1998) is employed to calculate the potentials  $\Phi(R, z)$  in Equation 7 for each star.

An alternative potential is also calculated in order to get a rough idea how the choice of potential may influence the subsequent data and the appearance of the integrals-of-motion space. Using the GALA Python package<sup>19</sup> (Price-Whelan 2017; Price-Whelan et al. 2020), the potential is calculated using their MilkyWayPotential, which assumes a Hernquist potential for a spherical bulge and nucleus (Hernquist 1990), a Miyamoto-Nagai potential for the disk (Bovy 2015; Miyamoto & Nagai 1975), and a Navarro-Frenk-White potential for a spherical dark matter halo. The model parameters for the nucleus ( $m = 1.7 \cdot 10^9 M_\odot$ ,  $c = 1$  kpc), bulge ( $m = 5.0 \cdot 10^9 M_\odot$ ,  $c = 1$  kpc), disk ( $m = 6.8 \cdot 10^{10} M_\odot$ ,  $a = 3$  kpc,  $b = 0.28$  kpc), and halo ( $m = 5.4 \cdot 10^{11} M_\odot$ ,  $a = 15.62$  kpc) are adopted from Naidu et al. (2020). The resulting potentials are plotted against the potentials calculated using the McMillan (2017) potential, as depicted in Figure 40. There is an extremely tight correlation between the two potentials, and the comparison to the identity line shows that the difference between the two is basically just an offset, although this offset grows slightly towards higher potentials. As Helmi et al. (2017) point out, the potential is mostly a zero-point offset from the kinetic energy of the stars, and as such the particular form

<sup>18</sup><https://github.com/PaulMcMillan-Astro/GalPot>, retrieved 22.10.2023

<sup>19</sup><https://gala.adrian.pw/en/latest/index.html>, retrieved 22.10.2023

that it takes is not central here as long as it describes the conditions within the volume covered by the sample reasonably well, which was also shown by [Naidu et al. \(2020\)](#). The [McMillan \(2017\)](#) potential is used in the remainder of this work.

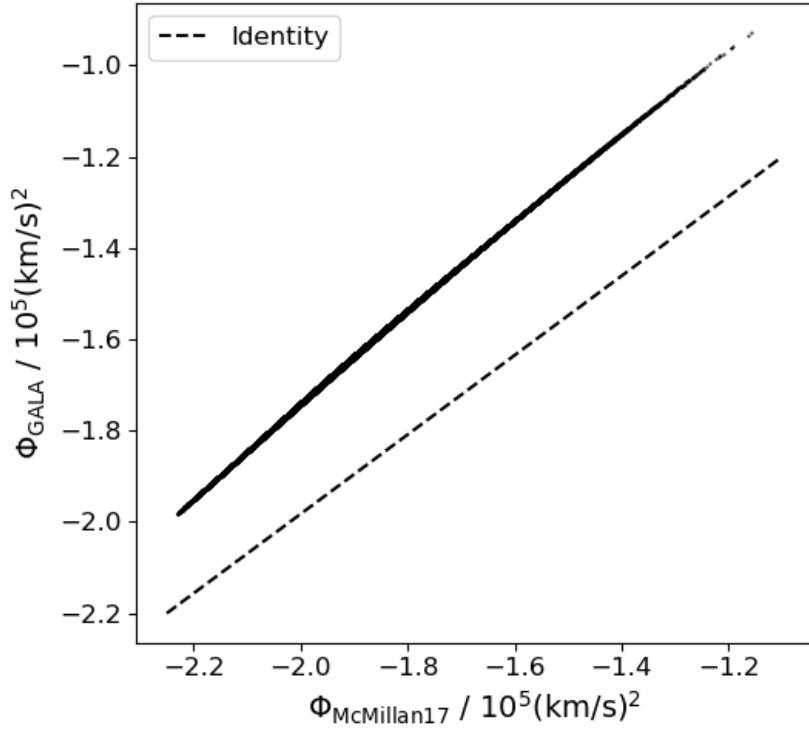


Figure 40: Comparison between the potential calculated via the default GALA potential `MilkyWayPotential` and the [McMillan \(2017\)](#) potential. The dashed line represents identity.

As it was decided to include no kinematic selection criteria in the creation of the 4MOST S2 target catalogue, it is expected that a majority of the targets contained are in fact still part of the disk. A commonly used simple kinematic criterium to separate the halo from the rest of the Galaxy is making use of the high velocity dispersion of halo stars via the condition  $|v - v_{\text{LSR}}| > 210 \text{ km/s}$  (see, e.g., [Helmi et al. 2018](#); [Ruiz-Lara et al. 2022](#); [Koppelman et al. 2018](#)), where  $v_{\text{LSR}} = 232 \text{ km/s}$  for the local standard of rest ([McMillan 2017](#)). This criterium is known to become increasingly less effective towards the inner parts of the Galaxy, as the velocity dispersion of disk stars becomes higher there. The cut at 210 km/s can be relaxed to a lower value, or chosen more conservatively at a higher value depending on the science case. Using this condition, the number of halo stars in the 6D subset is  $\sim 78\,300$  or about 9.4% of the stars for which kinematic parameters were determined.

A more detailed decomposition of the data into the morphological Galactic components is based on the orbit circularity  $\lambda_z$ . This parameter is defined as

$$\lambda_z = \frac{L_z}{L_{\text{circ},z}}, \quad (8)$$

where  $L_z$  is the  $z$ -component of the specific angular momentum and  $L_{\text{circ},z}$  is the  $z$ -component of the specific angular momentum that a star at the same radius would have if its orbit was circular. It is given by  $L_{\text{circ},z} = rv_{\text{circ}}$ , where  $v_{\text{circ}}$  is the Galaxy's circular velocity at radius  $r$ . It can be determined via

$$v_c^2 = r|\mathbf{F}| = r \frac{d\Phi}{dr} = \frac{GM_{\text{enc}}}{r}, \quad (9)$$

where  $|\mathbf{F}|$  is the gravitational attraction,  $\Phi$  is the gravitational potential, and  $M_{\text{enc}}$  is the mass enclosed within the radius  $r$  ([Binney & Tremaine 2008](#), p. 62). As `GALPOT` provides routines to calculate  $v_c$ ,

$M_{\text{enc}}$ , the forces  $F_R$  and  $F_z$ , the potential derivatives, and  $L_{\text{circ},z}$  directly, the circularity  $\lambda_z$  is calculated five times via these different parameters. The five results for  $\lambda_z$  differ from each other slightly, likely due to numerical reasons in the calculation routines in GALPOT, so the mean of these five values is adopted as the estimate of the circularity  $\lambda_z$ .

The sample is then decomposed into cold disk, warm disk, bulge, and stellar halo according to the criteria picked by [Sotillo-Ramos et al. \(2023\)](#) and [Chen et al. \(2023\)](#), who follow the approach laid out by [Zhu et al. \(2022\)](#). These criteria are

- cold disk:  $\lambda_z > 0.7$
- warm disk:  $0.5 < \lambda_z < 0.7$  and  $r_{\text{cut}} < r < r_{\text{disk}}$
- bulge:  $\lambda_z < 0.7$  and  $r < r_{\text{cut}}$
- stellar halo:  $(\lambda_z > 0.5 \text{ and } r_{\text{disk}} < r < r_{\text{halo}}) \text{ or } (\lambda_z < 0.5 \text{ and } r_{\text{cut}} < r < r_{\text{halo}})$ ,

where the bulge cutoff is  $r_{\text{cut}} = 3.5 \text{ kpc}$  ([Zhu et al. 2022](#); [Sotillo-Ramos et al. 2023](#)), the maximum extent of the halo is  $r_{\text{halo}} = 300 \text{ kpc}$  ([Sotillo-Ramos et al. 2023](#); [Chen et al. 2023](#)), and the exponential scale-length of the stellar disk is  $r_{\text{disk}} = 6 \cdot r_d = 6 \cdot 2500 \text{ kpc}$  ([Sotillo-Ramos et al. 2022](#); [Binney & Tremaine 2008](#), p. 18). This decomposition of the sample yields about 635 000 cold disk stars, 85 900 warm disk stars, 1730 bulge stars, and 115 600 halo stars. As expected, the majority of the 6D subset for which kinematic parameters were calculated is made up of cold disk stars. This method of decomposition assigns halo membership to more stars compared to the simple criterium based only on the stars' velocities and the local standard of rest, at 13.8% instead of only 9.4%. This is a notable difference, which is why in the following sections both halo selection criteria will be used, and in some cases, the results arising from the different halo selection criteria will be compared. A verification of these two halo selections will be carried out in Section 5.3. Finally, a third selection is to only include stars with a distance of  $|z| > 4 \text{ kpc}$  from the Galactic midplane, which is the case for about 25 600 stars. This halo sample is of course highly incomplete and biased, and is only selected to have a halo sample that is as clean as possible in the sense of being made up only of halo stars.

Figure 41 showcases the effect of the selections on the spatial distribution of the sample in galactocentric coordinates. Whereas the concentration in the region close to the Sun is so strong in the top left panel of the full sample that the colour-coding leads to only the inner parts being visible, the distribution in  $z$ - $R$ -space is much more spread out due to the removal of the disk contribution in the other three panels depicting the sample after the halo selection. In particular, the top right panel shows the bulge cutoff radius  $r_{\text{cut}} = 3.5 \text{ kpc}$ , whereas the bottom right panel simply visualises how only high- $z$  stars were selected. The halo selection based on the circularity parameter leaves a higher concentration of stars at negative  $z$  just below the Sun's position compared to the velocity-based selection, which is likely where much of the higher numbers in the former case comes from.

Figure 41 also suggests that the heliocentric distances should on average be much higher for the halo selections. To verify this, histograms of the heliocentric distances are plotted in Figure 42. As expected, this is the case, with the full sample's distance distribution peaking at a mere  $\sim 750 \text{ pc}$ , whereas the halo selections based on  $\lambda_z$  and  $v$  peak at around 5000 parsec before dropping off in a power-law decrease. Notably, they still have a second small peak at around 1000 pc, which is due to the remains of the (thick) disk that have not been removed by the selection criteria. This peak is slightly more prominent in the circularity-based selection, but otherwise the two distributions in the top right and bottom left take the same form, with only the numbers in the top right panel being higher. The distributions show that 4MOST S2 will sample the halo at least from about 2.5 kpc to 5 kpc.

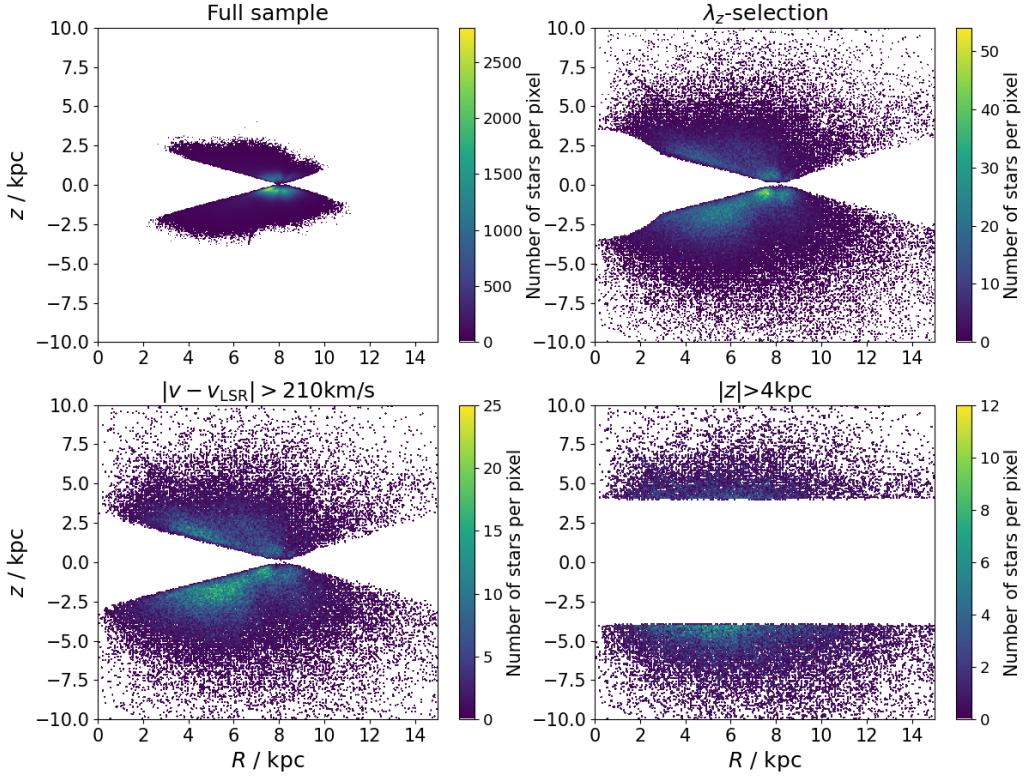


Figure 41: The spatial distribution in  $z$ - $R$ -space of the full 6D subset and the three halo selections. Most of the top left panel remains white due to the high number of stars close to the Sun dominating.

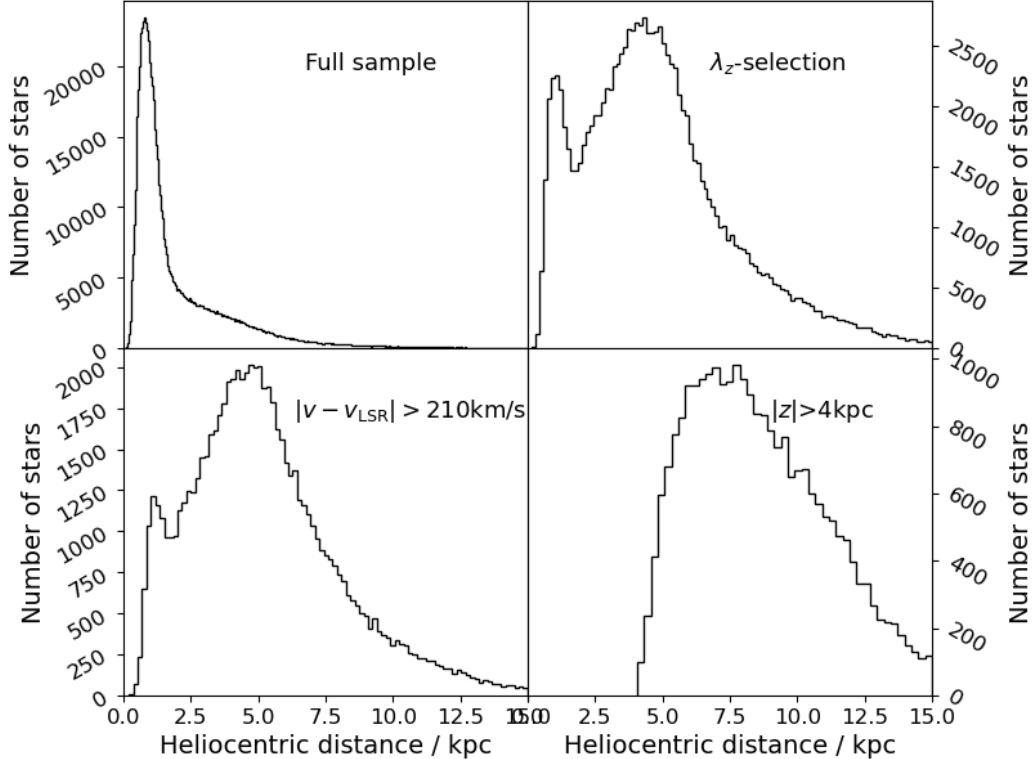


Figure 42: Histograms of the heliocentric distances for the full 6D subset and the three halo selections.

Finally, Figure 43 depicts the distribution of absolute space velocities in the sample and the halo selection. The full sample has a narrow peak at about 230 km/s, which signifies the ordered rotation of disk stars and again showcases how the majority of the sample consists of disk stars. The halo selections peak at less than 200 km/s, but have a broader distribution and include comparably more stars with very small

and very large space velocities. Notably, the peak of the circularity-based halo selection is very pointed, whereas the distributions in the bottom two panels look very similar aside from the absolute numbers on the  $y$ -axis. On one hand, this might mean that the velocity-based selection of the halo might yield a more realistic representation of the halo, but on the other hand, both the bottom panels may suffer from a similar bias in removing stars of the inner halo with lower absolute space velocities. As a side note, there are 1693 high-velocity stars at  $v > 500$  km/s in the sample, the closer investigation of which may also be interesting, but is beyond the scope of this thesis.

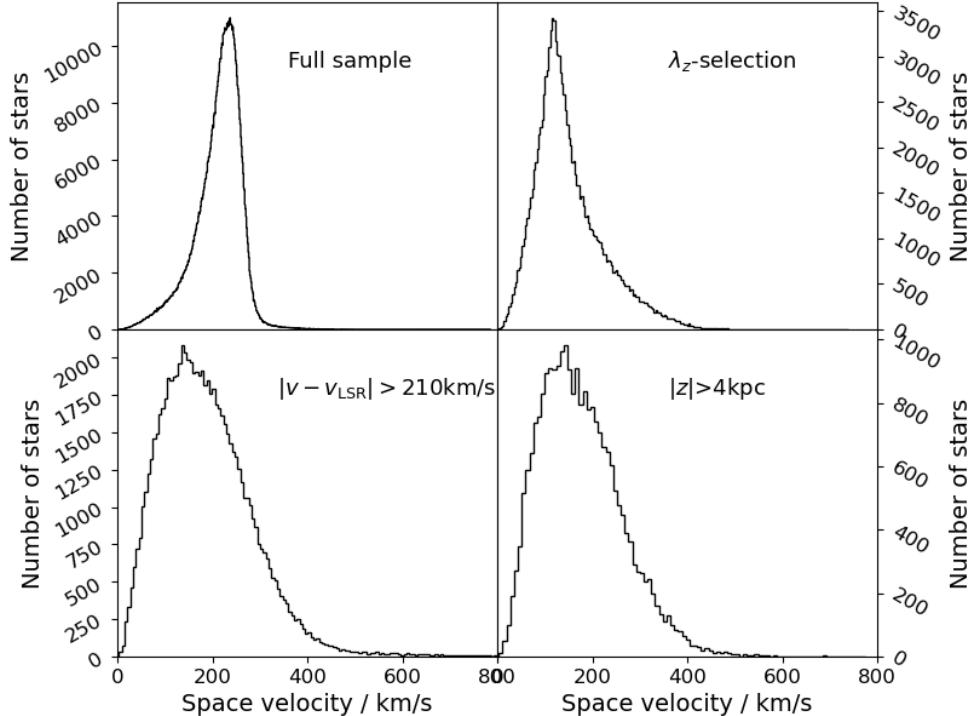
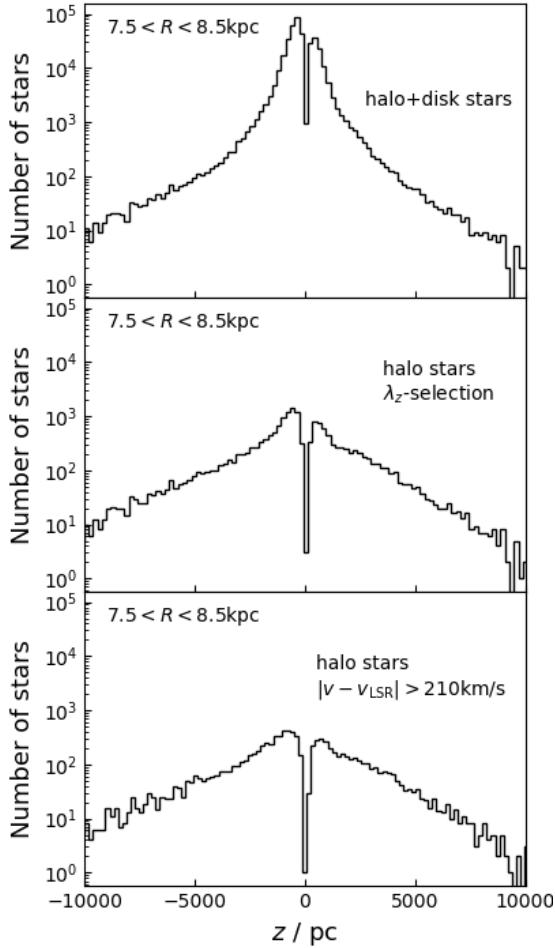


Figure 43: Histograms of the absolute space velocities for the full 6D subset and the three halo selections.

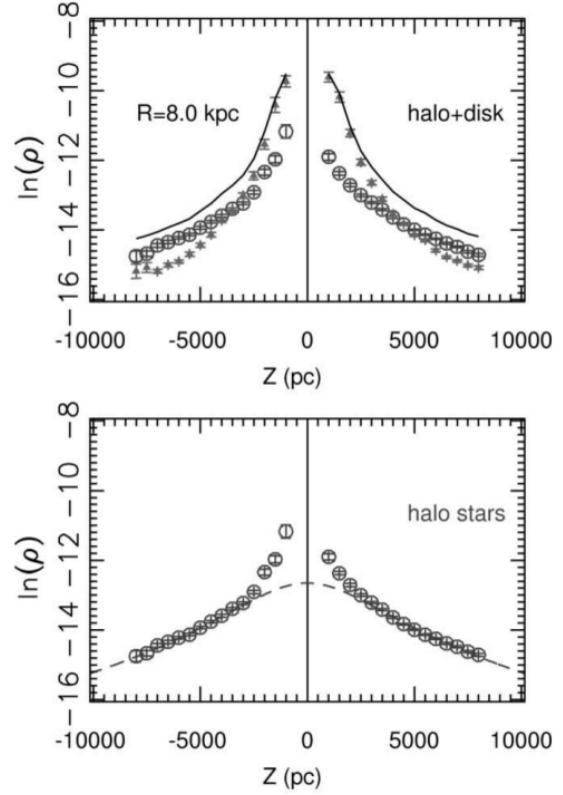
### 5.3 Verification of the halo selection

One property setting the stellar halo apart from the disk is the number density distribution of stars as a function of the position in the Galaxy, in particular how it changes with the distance from the Galactic midplane and with the radial distance from the Galactic centre. Although it is not possible to tell if components are physically distinct systems from the number densities alone without the inclusion of kinematic and/or chemical abundance data, it does turn out that the density profiles of the disk and stellar halo are very different. Using data from the SDSS to create stellar number density maps for the investigation of Galactic structure, Juric et al. (2008) determine that the disk profile is significantly more steep than the stellar halo profile and dominates at small distances from the Galactic midplane  $|z| < 3$  kpc, whereas the halo becomes dominant far away from the midplane. In Figure 44, their results for the distribution of the stellar number density as a function of the distance from the Galactic midplane (galactocentric  $z$ -coordinate) at  $R = 8$  kpc is shown next to a similar distribution for the two halo selections of the data in this work. However, the quantity plotted to emulate the distribution in Juric et al. (2008) is the total number of stars in the slice  $7.5 \text{ kpc} < R < 8.5 \text{ kpc}$  rather than the number density, which is equivalent in this case, as each bin in the histogram represents the number of stars in a spatial volume of the same size,  $4\pi \cdot (1 \text{ kpc})^2 \cdot \text{binwidth} [\text{kpc}]$  to be specific, so having the number density on the  $y$ -axis instead would merely result in a constant offset.

As can be seen, the visual agreement is quite good, as the top panel in both cases shows the narrow



(a) Stellar number distribution across  $z$  in the slice  $7.5 \text{ kpc} < R < 8.5 \text{ kpc}$  for the full sample (top panel,  $\sim$ disk+halo) and just the halo (bottom two panels). Both selection criteria from Section 5.2 are shown.



(b) Stellar counts in SDSS across  $z$  at  $R = 8 \text{ kpc}$  for the halo and disk (top) and only the halo (bottom). The circles represent the halo counts, the triangles represent the disk counts, and the line in the top panel is the sum of these two counts (Juric et al. 2008, Fig. 33).

Figure 44: Stellar counts as a function of the distance from the Galactic midplane.

peak of the low- $z$  region being dominated by the disk, whereas the halo profile is much more flat in the regions close to  $z \sim 0$ . The difference between the height of the peaks of the profile for the halo+disk sample and just the halo sample is around two orders of magnitude in both plots. In particular, it seems that the halo selected via the circularity parameter  $\lambda_z$  criterium is most similar to the halo distribution by Juric et al. (2008), as their data also has the halo profile rising more steeply for negative  $z$  going towards zero in a shape very similar to the middle panel on the right, whereas the profile in the bottom panel remains rather flat even there. Finally, it should also be noted that the halo selection in their work is based on a simple metallicity-based colour cut, as opposed to this work’s kinematic selection, and is thus only approximate as there will be strong contaminations, particularly of disk stars in the halo sample at the bright end, as there are significantly more disk stars there.

To have a more quantitative check against previous works researching the shape of the Milky Way’s stellar halo profile, a single power-law and a broken power-law model fit to the radial galactocentric distances are carried out to compare the fit parameters of the models to the literature. Assuming an oblate density distribution of fixed  $q = 0.6$ , the flattening-corrected galactocentric radius is  $r_q = \sqrt{x^2 + y^2 + z^2/q^2}$ . As for the models, the single power-law takes the form

$$N = N_0 \cdot r_q^\alpha, \quad (10)$$

where  $N_0$  is the normalisation (which is not relevant here) and  $\alpha$  is the radial slope. The broken power-law modelling a two-component halo is given by

$$N = \begin{cases} N_0 \cdot \left( \frac{r_{\text{break}}}{r_q} \right)^\alpha & \text{if } r_q < r_{\text{break}} \\ N_0 \cdot \left( \frac{r_{\text{break}}}{r_q} \right)^\beta & \text{if } r_q > r_{\text{break}} \end{cases}, \quad (11)$$

where  $N_0$  is again the irrelevant normalisation,  $\alpha$  is the inner component's slope,  $\beta$  is the outer component's slope, and  $r_{\text{break}}$  is the radius at which the outer component starts to dominate. The stellar counts across the flattening-corrected coordinate  $r_q$  are then fitted to these two models for both the halo selections from 5.2, as shown in Figure 45. The two dashed lines show that the range of the fit was restricted to a rather small distance range of  $7.5 \text{ kpc} < r_q < 20 \text{ kpc}$ , as the number of objects beyond 20 kpc is on the order of less than 10, meaning that no meaningful fit can be carried out there, and because the stellar halo's slope of decreasing stellar numbers starts only at 7.5 kpc. In fact, even before 20 kpc, the fluctuations of the histogram relative to the total number of stars becomes pretty significant, meaning that this range is a major source of uncertainty for the fit. In general, the fits seem to be successful, with the coefficients of determination being 0.98 for both models in the top panel, and 0.97 for both models in the bottom panel.

Table 4 lists the fit parameters found for the two models for both halo selections, and also lists parameters found in the literature for comparison. The values for  $\alpha$  (and  $\beta$ ) are generally slightly higher than those commonly reported in the literature, i.e. the number of objects is dropping faster in comparison. It should be noted that the references typically have data up to much higher radii, but even so there appears to be an incompleteness at the far end of the sample, as the number of objects is already insignificant starting at around 20 kpc. This incompleteness at the far end beyond 20 kpc is consistent with the findings of Section 4.4, where a lack of faraway stars was indicated in comparison to the Besançon model. The steepness of the slope and incompleteness of the sample are likely a result of the apparent magnitude limit, which is 17 mag for 4MOST S2 in general, but is actually only around 15 mag for the 6D subset that is the basis of this kinematic analysis, as the completeness limit seen in Figure 37 is at  $\sim 14.5$  mag. As for the two different halo selection criteria, both selections appear to be similarly consistent with the literature for the single power-law model, whereas for the broken power-law model, the criterium based on the velocity seems more consistent.

Table 4: The fit parameters for the single power-law model (middle) and the broken power-law model (right) for the two halo selections carried out in Section 5.2, and some literature values for comparison.

Source	$\alpha$	$q$	$\alpha$	$\beta$	$r_{\text{break}} / \text{kpc}$	$q$
This work ( $\lambda_z$ -criterium)	$4.36 \pm 0.03$	0.6	$4.31 \pm 0.04$	$4.80 \pm 0.35$	$16.9 \pm 1.3$	0.6
This work ( $v$ -criterium)	$4.37 \pm 0.04$	0.6	$3.89 \pm 0.09$	$5.01 \pm 0.11$	$13.4 \pm 0.5$	0.6
Xue et al. (2015)	$3.6 \pm 0.1$	0.68	$2.1 \pm 0.3$	$3.8 \pm 0.1$	$18 \pm 1$	0.7
Xue et al. (2015)	$3.4 \pm 0.1$	0.74	$2.8 \pm 0.1$	$4.3 \pm 0.1$	$29 \pm 2$	0.77
Pila-Díez et al. (2015)	$4.31 \pm 0.09$	0.79	$2.4 \pm 0.05$	$4.8 \pm 0.05$	$19.0 \pm 0.5$	0.77
Hernitschek et al. (2018)	$4.4^{+0.05}_{-0.04}$	0.92	$4.97^{+0.02}_{-0.05}$	$3.93^{+0.05}_{-0.04}$	$38.7^{+0.69}_{-0.58}$	0.91
Deason et al. (2011)	$2.9^{+0.04}_{-0.06}$	0.53	$2.3 \pm 0.1$	$4.6^{+0.2}_{-0.1}$	$27 \pm 1$	0.59
Faccioli et al. (2014)	n.a.	n.a.	$2.7 \pm 0.6$	$3.6 \pm 0.4$	$26.5 \pm 8.9$	0.59
Faccioli et al. (2014)	n.a.	n.a.	$2.3 \pm 0.5$	$4.8 \pm 0.5$	$28.5 \pm 5.6$	0.7

In any case, the parameters found in the literature generally vary widely between different publications, as can also be seen in Table 6 of Bland-Hawthorn & Gerhard (2016), because many factors such as the inclusion or exclusion of substructure, the dataset which is used as well as the selection function relating to this, the choice of  $q$  as fixed or free, and in the former case the choice of the value of  $q$ , all

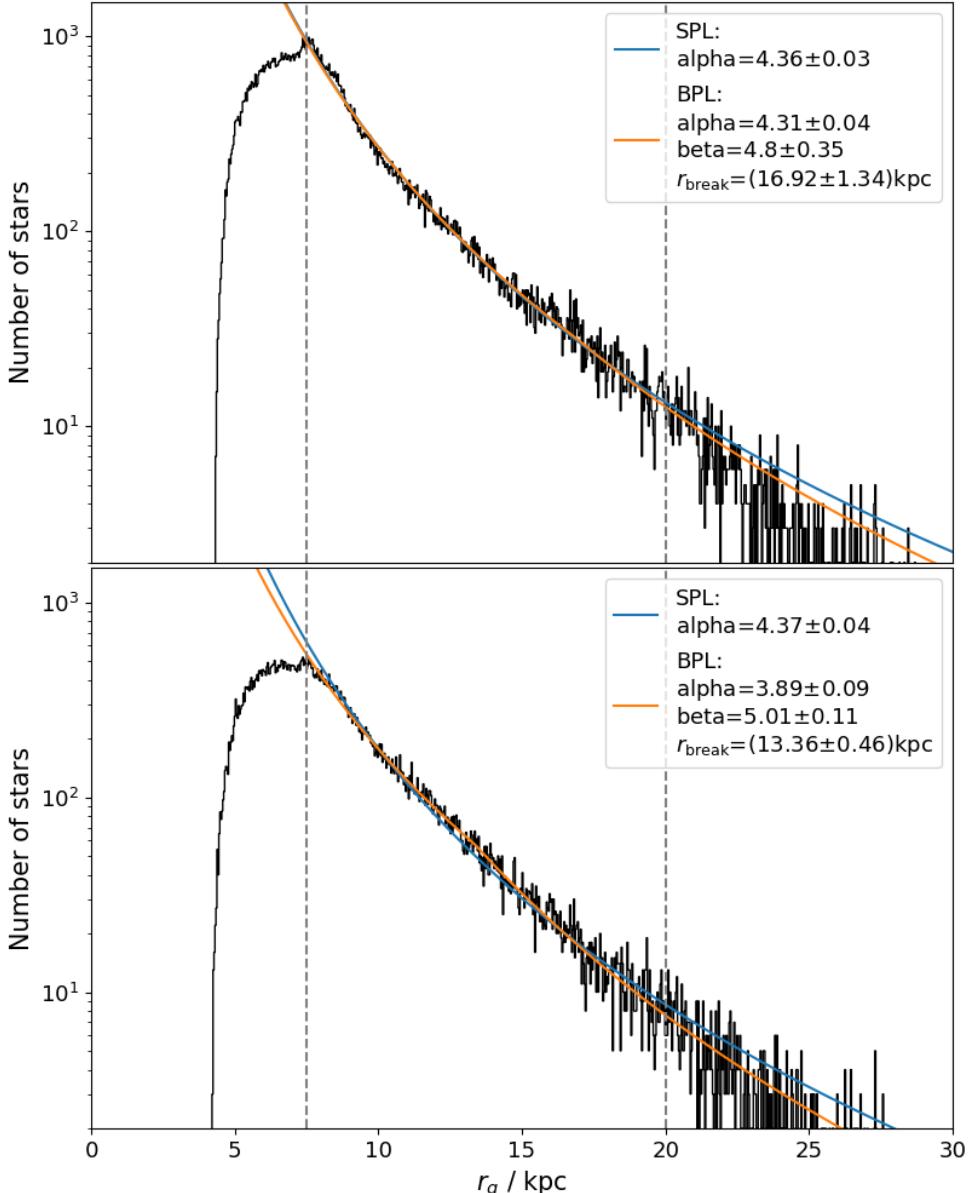


Figure 45: Histograms of the flattening-corrected galactocentric radius  $r_q$  including a simple power-law and broken power-law fit. Top panel:  $\lambda_z$ -based selection; bottom panel: velocity-based selection.

play an important role in shaping the output of the fit to the models. Finally, the major caveat of the fit in this work is that, as opposed to the references in Table 4, the fits to the models were carried out using the absolute stellar counts rather than the number density. This means that the slopes found are likely even larger, as the count of each bin in  $r_q$  represents the number of stars in an ellipsoidal shell around the Galactic centre, the volume of which increases with  $r_q$ . Here, the analysis is limited to simply using the absolute numbers to obtain a rough estimate, as the intricacies of properly taking into account the selection function for the dataset to derive accurate stellar densities are beyond the scope of this thesis.

Next, colour-magnitude diagrams are plotted similarly to Section 4.1, but this time for the different halo selection cuts. This is shown in Figure 46, where the top left panel includes all stars in the 6D subset, whereas the other three panels are the three selections of the halo. It is immediately apparent that the number of giants, in particular horizontal branch stars, relative to all stars is much higher for the halo selections than for the full sample, which is expected as most of the halo is farther away from the Sun while the Solar neighbourhood is dominated by disk stars. Conversely, the ratio of main sequence stars

is much higher in the full sample, and there are more close stars which have a lower luminosity and are mainly at the main sequence turnoff, so by applying the kinematic criterium, different populations are sampled. Since no more selection effects from the 4MOST S2 target selection are visible in the CMDs of only the halo, it does not seem like adjusting the selection criteria in colour and absolute magnitude will help to increase the ratio of halo stars in the target catalogue.

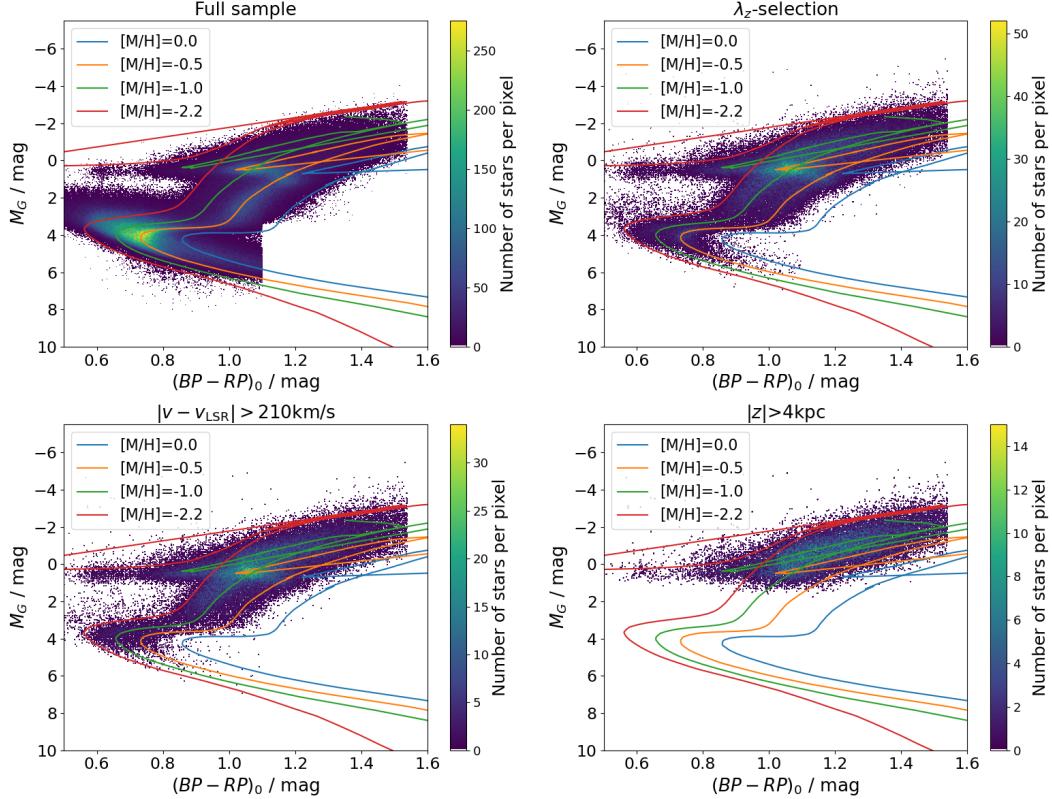


Figure 46: Colour-magnitude diagrams of the full 6D subset and the three halo selections, including a set of PARSEC isochrones.

A notable difference between the halo selection using the circularity parameter and the halo selection using the velocity is that there seem to be relatively more main sequence turnoff stars left in the former case, which are possibly disk stars not removed by the criterium. The bottom right panel obviously only contains the brightest stars in the upper part of the CMD, as only those are bright enough to be observed from that distance above the Galactic midplane, considering the apparent magnitude limitation. Finally, the different metallicity regimes of the halo and disk are visible from the positions of the isochrones: The full sample is most well-represented by the isochrone at  $[M/H] = -0.5$ , as the colour-coding depicting the density of objects indicates that most objects follow that isochrone. On the other hand, the data in the other three panels traces the isochrone at  $[M/H] = -1.0$  best, with a tendency towards even lower metallicities in the red giant branch. This is further verification for the halo selection, as the halo consists largely of metal-poor stars.

Making use of the spatial density profiles, stellar orbits, and stellar metallicities, [Carollo et al. \(2007\)](#) show that there is a division of the stellar halo into an inner halo and an outer halo, with the former having net prograde rotation and the latter having net retrograde rotation as well as a much lower metallicity. They also conclude that the processes underlying the formation of the two components is probably fundamentally different, as their properties are distinctly different. Following their approach (see their Figure 2), the metallicity distribution function is plotted for a large number of cuts in velocity in the rotation direction of the Galaxy, which is  $v_y$  in this case (Figure 47). Here, the halo was selected according to  $|v - v_{\text{LSR}}| > 210 \text{ km/s}$ , and a histogram of  $[\text{Fe}/\text{H}]$  is shown for increasingly more retrograde selections

of the velocity component  $v_y$ , starting from largely prograde stars  $v_y \geq -10$  km/s in the top left panel to  $v_y < -450$  km/s. The bottom right panel shows the metallicity distribution function for the whole halo sample.

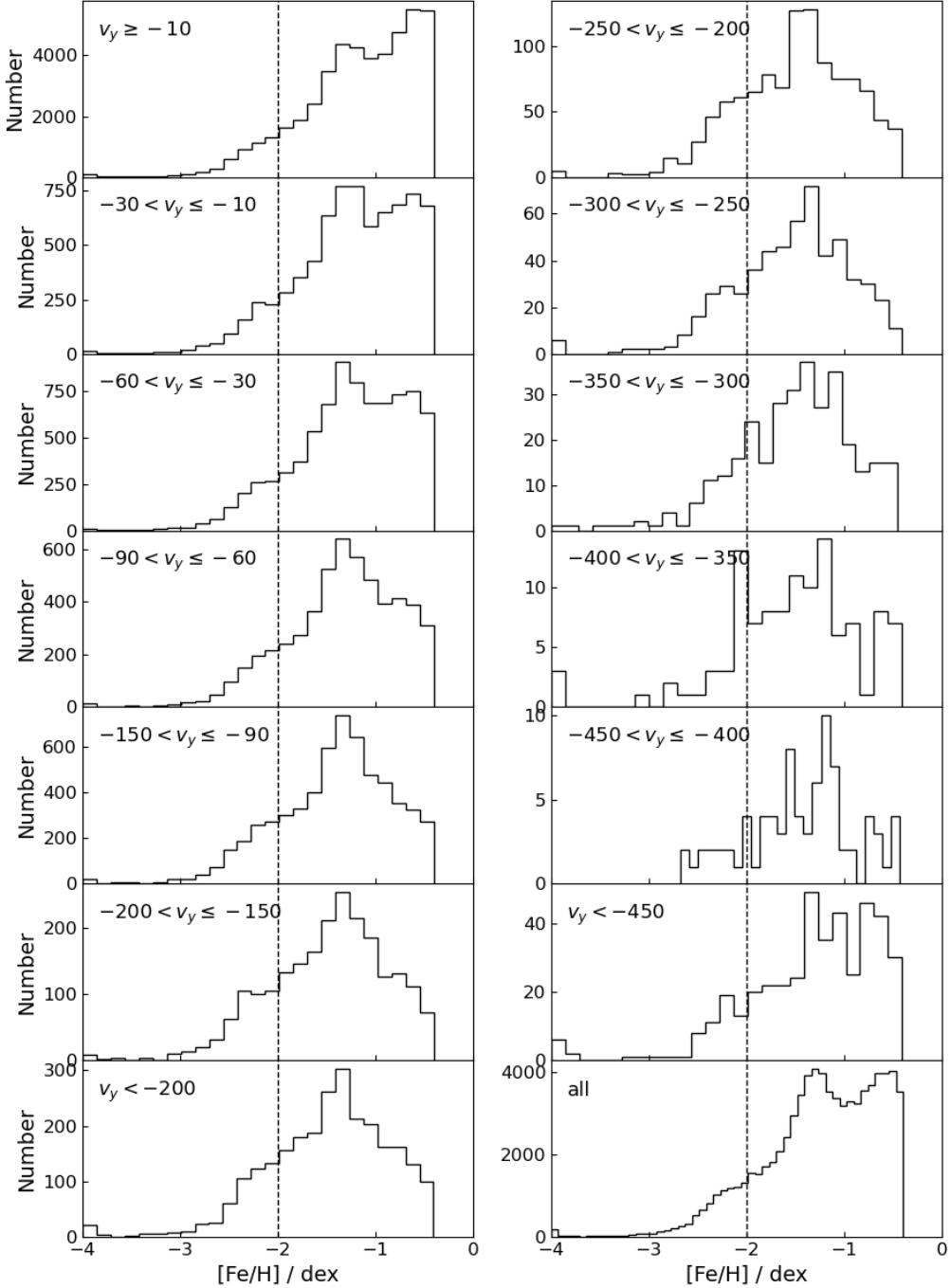


Figure 47: The metallicity distribution function in different velocity cuts, where a more negative  $v_y$  indicates a more retrograde sample. The dashed line represents a metallicity of  $[Fe/H] = -2.0$  throughout the figure.

The first thing that is apparent is that the halo becomes more metal-poor for more retrograde selections. For the top left panel containing mostly prograde stars, the most significant peak is at the highest-metallicity end of the sample at about  $[Fe/H] \sim -0.5$ , although a second peak at about  $[Fe/H] \approx -1.4$  is also visible. The former peak is likely due to contamination of thick disk stars, and as it is also very prominent in the bottom right panel containing the whole halo sample, it seems that the amount of (thick) disk stars left after the purely kinematic selection  $|v - v_{LSR}| > 210$  km/s is still notable. In any case, this

peak drops quickly as the velocity cuts become more retrograde, and is not really present anymore from  $v_y < -90$  km/s at the latest. The peak at about  $-1.4$  remains dominant throughout the rest of the cuts, but as the cuts become more retrograde, there is a hint of a second peak emerging at  $[\text{Fe}/\text{H}] \approx -2.0$ , particularly in the right column of the figure. Unfortunately, the present sample does not contain a significant amount of stars in the range  $-450 < v_y < -350$  km/s, but the upper panels in the right column still contain enough stars to make out the emerging second peak. Even in the panel depicting the full halo sample, the distribution does not simply keep dropping towards low metallicities, but has a section between  $[\text{Fe}/\text{H}] = -2.0$  to  $-2.3$  where the slope of the histogram is basically flat, indicating that there is another peak hidden in the tail of the first peak here. As opposed to the data of [Carollo et al. \(2007\)](#), however, the peak at about  $-1.4$  remains stronger until the end.

The emerging second peak at about  $-2.0$  corresponds to the outer halo component, whereas the more prominent peak at about  $-1.4$  corresponds to the inner halo component identified by [Carollo et al. \(2007\)](#). Clearly, the inner halo dominates in this sample, as the outer halo peak never emerges strongly. Different ideas about the nature of the halo duality have been put forward, with some authors explaining the more metal-poor part as belonging to Gaia-Enceladus, whereas the more metal-rich part belongs to the tail of the thick disk, and some authors even suggesting that the Milky Way's halo is entirely accreted, and any in-situ halo stars originate from the heated thick disk following heating from Galaxy interactions such as accretion events ([Helmi 2020](#)). For instance, [Di Matteo et al. \(2019\)](#) find that metal-rich thick disk stars make up about half of their kinematically selected halo stars in the Solar neighbourhood. Keeping this in mind, the peak at  $[\text{Fe}/\text{H}] = -2.0$  might represent the contribution of Gaia-Enceladus in the data, but it appears as though the presence of metal-poor accreted stars constituting the outer halo is overshadowed by a dominant inner halo down to the very retrograde end, and additionally by the contamination from disk stars at the more prograde end of the selection, although the analysis is limited by a lack of very retrograde stars. On the other hand, the results by [Ruiz-Lara et al. \(2022\)](#) suggest that Gaia-Enceladus is contained within the peak at  $[\text{Fe}/\text{H}] \approx -1.4$ , as their overall halo metallicity distribution function looks very similar to the bottom right panel in Figure 47 and their selection of Gaia-Enceladus closely traces the peak at about  $-1.4$ . This would also support Gaia-Enceladus constituting a significant fraction of the inner halo, as discussed in Section 1.2.

## 5.4 The search for kinematic substructure in the halo

Galaxy mergers and accretion events are vital to the understanding of both the formation and the evolution of galaxies. Previous studies concentrated on integral-of-motion spaces (for example angular momentum or energy), as objects that were accreted from the same origin stay close together in this space even if they are spatially indistinguishable from other objects (see, e.g., [Helmi et al. 2017](#); [Ruiz-Lara et al. 2022](#)). As a result of such studies, a lot of debris from a number of different accretion events was discovered in the Milky Way's stellar halo, the largest and most striking being Gaia-Enceladus, Sequoia, Thamnos I and II, as well as the Helmi Streams, and many more smaller ones have been identified as well (e.g. [Koppelman et al. 2019a](#); [Naidu et al. 2020](#); [Helmi et al. 2018](#)). Using the parameters and halo selections determined in Section 5.2, the three halo selections and the full 6D sample are shown in the integral-of-motion space involving the angular momentum and energy of stars (Figure 49, 50), as well as in velocity space (Figure 48).

In all three plots of the full sample, the dominance of the disk kinematics is visible in the form of an extremely high concentration of stars in one region of the plot. In the plot showing  $\sqrt{v_x^2 + v_z^2}$  vs.  $v_y$ , this is at around  $v_y \sim 240$  km/s at low  $\sqrt{v_x^2 + v_z^2}$ , in the  $L_\perp = \sqrt{L_x^2 + L_y^2}$  vs.  $L_z$  space, it is just under  $L_z \approx 2000$  kpc·km/s at low  $L_\perp$ , while in the  $E_{\text{orb}}$  vs.  $L_z$  plot, it is around  $(L_z, E_{\text{orb}}) \sim (2000 \text{ kpc}\cdot\text{km/s}, -150\,000 (\text{km/s})^2)$ . In each of the three plots, the application of the halo selections clearly removes the absolute majority of the disk kinematics, as the distributions go from being heavily centered on the disk kinematics to being much more evenly distributed around  $L_z = 0$ .

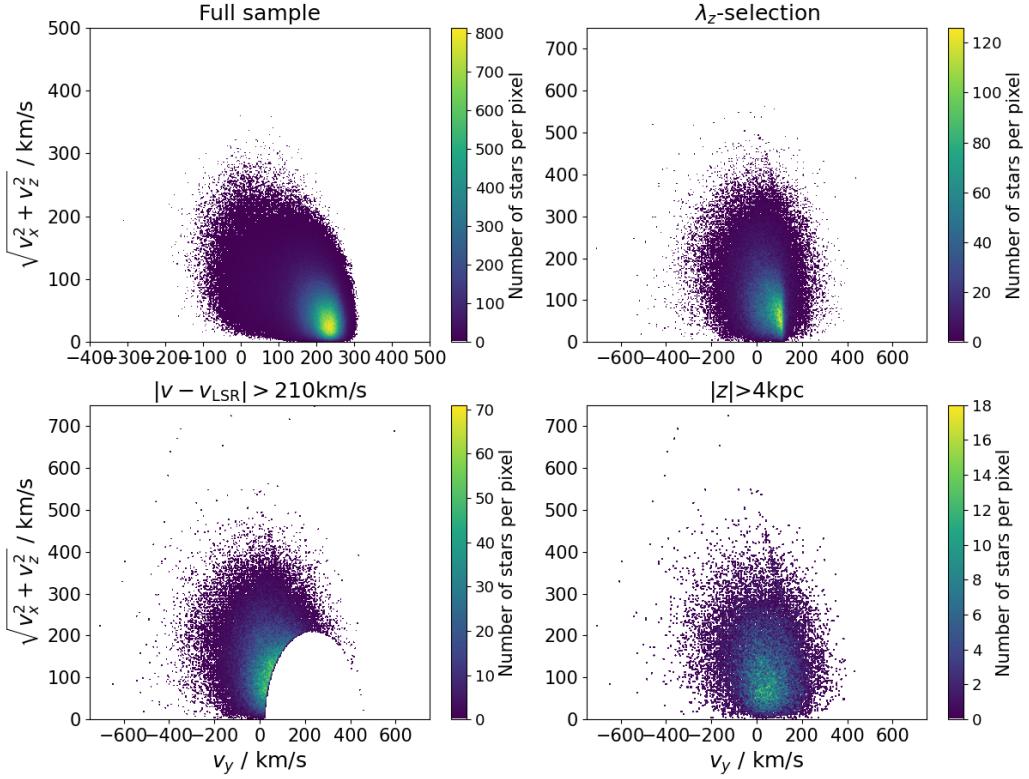


Figure 48: Perpendicular velocity component  $v_{\perp} = \sqrt{v_x^2 + v_z^2}$  vs.  $v_y$  for the full 6D subset and the three halo selections. In the bottom left, the selection cut is clearly visible.

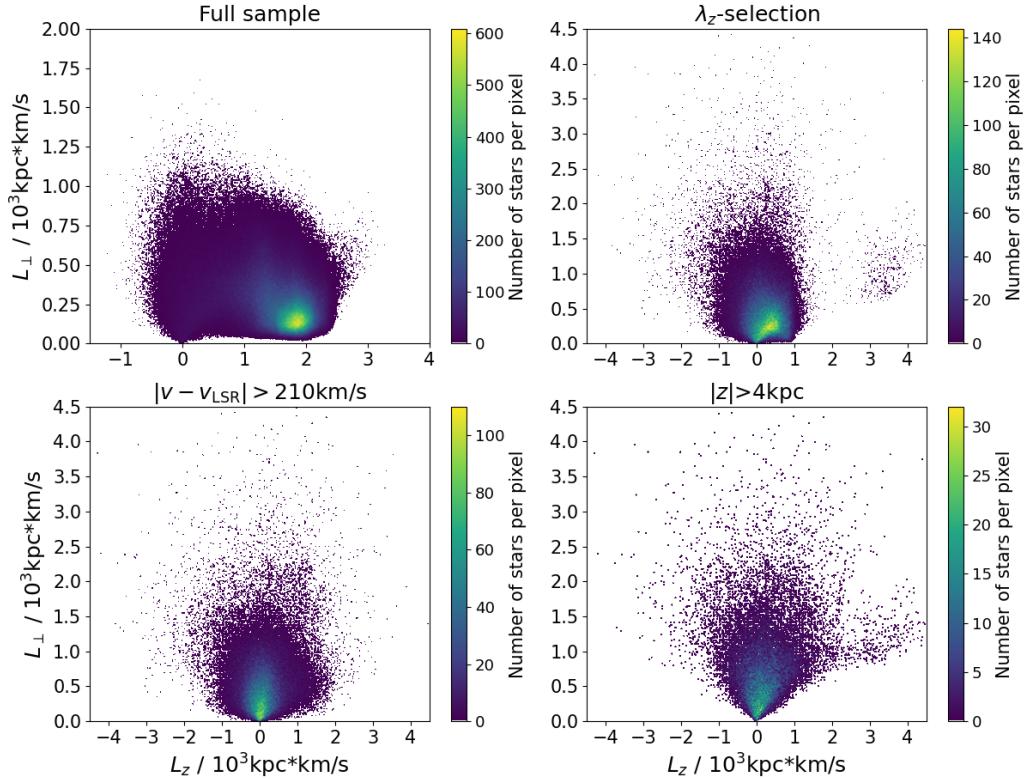


Figure 49: Perpendicular specific angular momentum component  $L_{\perp} = \sqrt{L_x^2 + L_y^2}$  vs.  $L_z$  for the full 6D subset and the three halo selections. The objects on the right of the top right panel are separated from the bulk of the sample due to the selection cut.

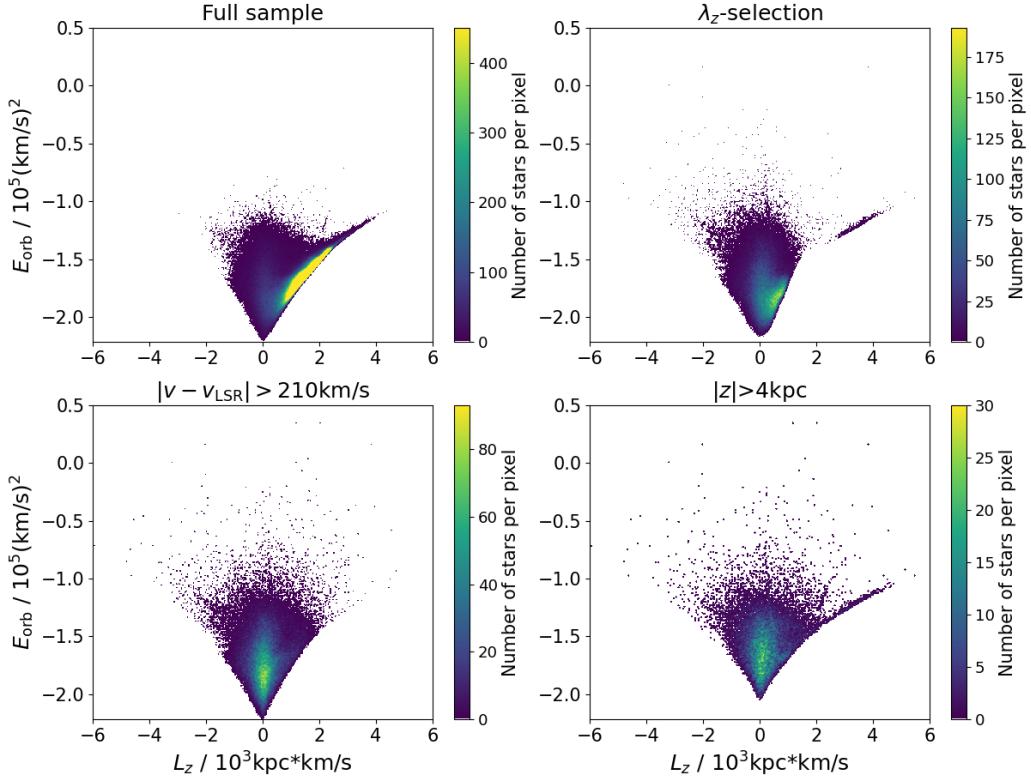


Figure 50: Orbital energy  $E_{\text{orb}}$  vs. specific angular momentum component  $L_z$ . As in the previous figure, the objects on the right of the top right panel are separated from the bulk of the sample due to the selection cut.

This is also reflected in the numbers, as only about 9.4% in the case of the velocity-based selection and 13.8% in the case of the  $\lambda_Z$ -based selection remain. The  $\lambda_z$ -based selection appears to have left in a larger number of stars between  $L_z \sim 0$  and  $L_z \sim 1 \text{kpc}\cdot\text{km/s}$ , giving the distribution a less symmetrical look compared to the traditional selection  $|v - v_{\text{LSR}}| > 210 \text{ km/s}$ , which can also be seen in Figure 48. These stars have more disk-like kinematics compared to the majority of the sample, and alongside hints in Section 5.3, this suggests that this halo selection indeed still includes a significant fraction of disk stars, meaning that perhaps the selection criteria need some more fine-tuning. The effect of the halo selection  $|v - v_{\text{LSR}}| > 210 \text{ km/s}$  can be observed in the bottom left panel of Figure 48. Due to the aforementioned reasons, only the selection  $|v - v_{\text{LSR}}| > 210 \text{ km/s}$  will be used from here on when referring to the halo sample.

Although these figures showcase the distribution in the kinematic parameter space and its differences between disk stars and halo stars, as well as depicting selection cut effects, it is difficult to tell if there are any halo substructures present. Only upon close inspection, one can see in the top right and bottom left panel of Figure 50 two distinct structures, one extending slightly into the direction where the disk stars were removed and the other one extending vertically up to high  $E_{\text{orb}}$  around  $L_z \sim 0$ . This former structure is what is commonly identified as Gaia-Enceladus (Helmi et al. 2018; Koppelman et al. 2018, 2019a) and makes up a significant part of the stellar halo. In order to detect further known substructures and to make the described characteristic shape in  $E_{\text{orb}}-L_z$ -space more visible, further analysis is carried out.

First, the distribution in  $E_{\text{orb}}-L_z$ -space is smoothed and the smoothed distribution subtracted from the data in order to remove the background. For this, the range  $-4000 \text{kpc}\cdot\text{km/s} < L_z < 4000 \text{kpc}\cdot\text{km/s}$  and  $-2.2 \cdot 10^5 (\text{km/s})^2 < E_{\text{orb}} < -1.0 \cdot 10^5 (\text{km/s})^2$  is divided up into 250 by 250 cells, and for each cell the number of stars within it is counted. The resulting plot is shown in the left panel of Figure 51 and represents the density distribution of the stars in this space. This distribution is smoothed using a

Gaussian kernel, as shown in the middle panel. The right panel shows the first minus the second panel, in an attempt to remove the background and make structures deviating from the general homogeneous structure more visible. However, only the central structure is highly pronounced, with the part to the right of it not being very distinct. Even so, it is slightly more visible than it was in either the left panel or Figure 50, and confirms the shape suspected from that figure to be in line with other works aiming to isolate substructure in integrals-of-motion space (Naidu et al. 2020; Lövdal et al. 2022a; Yuan et al. 2020).

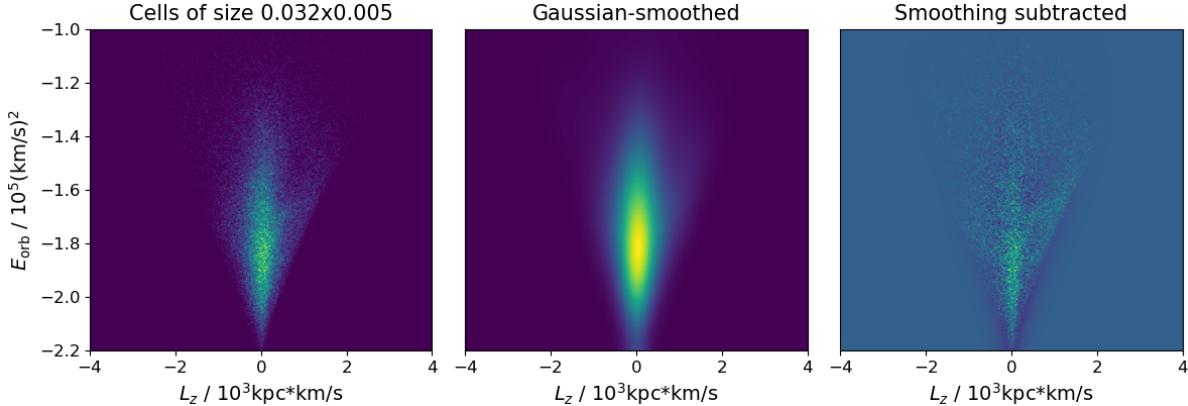


Figure 51: Left panel: Number of stars within each cell on a grid of  $250 \times 250$  cells indicated by colour-coding. Middle panel: Smoothed via a Gaussian kernel of width 20 in  $L_z$  and 5 in  $E_{\text{orb}}$ . Right panel: Subtraction of smoothed version from left panel.

Gaia-Enceladus can be roughly made out by eye from the previous figures, but to recover some other prominent substructures identified in the literature, a clustering algorithm grouping stars based on their position in relative overdensities is necessary. For this, the hierarchical density-based clustering algorithm HDBSCAN, which is part of the Python package SKLEARN (Pedregosa et al. 2011), is used, as it can cluster across different density scales and does not need an input on the distance a data point should have from the core of a sample in order to be considered part of it, but instead goes over all different values and finding the most stable results. As the aim of applying a clustering algorithm to the data is just to check if some known substructures can be made out within the data – rather than a thorough census of exact substructure memberships and definite results concerning even faint halo substructures, for which a more custom-tailored and sophisticated algorithm would be required – HDBSCAN is suitable due to its simplicity in usage and independent mode of operation. The work of Ruiz-Lara et al. (2022) in particular will serve as a reference due to the amount of detail given by the authors for each substructure. A validation of the reliability of the method with which the kinematic parameters were determined for this work is provided in Appendix A.1 in the form of a reproduction of results from previous works using the same approach for the parameter determination as was used in Section 5.2.

HDBSCAN finds 14 clusters in total in the  $E_{\text{orb}}-L_z-L_{\perp}$ -space, where a minimum cluster size of 20 is chosen to not allow the identification of clusters that are too insignificant (Figure 52). One of them is the big blob in the middle where the density is at its maximum, and a few of the others are distributed within and around the big blob where its dominance starts to taper off, indicating that these are mostly not actual physical structures but instead arise randomly as edge effects of the massive central blob. However, there are some interesting ones which do not fall under this category, in particular clusters 0 to 5. Cluster 0 sits at the top left corner of the  $E_{\text{orb}}-L_z$ -distribution and corresponds to the substructure known as Sequoia, as the positions not only match the literature in  $E_{\text{orb}}-L_z$ -space, but also in  $L_{\perp}-L_z$ -space (Figure 53) and  $v_{\perp}-v_y$ -space (Figure 54), albeit with more scatter in the former case when compared to, e.g., the results of Figure B.3 by Ruiz-Lara et al. (2022), with some of the stars also tracing the characteristic arc at  $v_y \sim -300$  visible in their Figure B.2 and other works, e.g. Figure 2 of (Koppelman et al. 2019a). Some authors suggest that Sequoia is part of the metal-poor tail of Gaia-Enceladus (e.g., Koppelman

et al. 2019a).

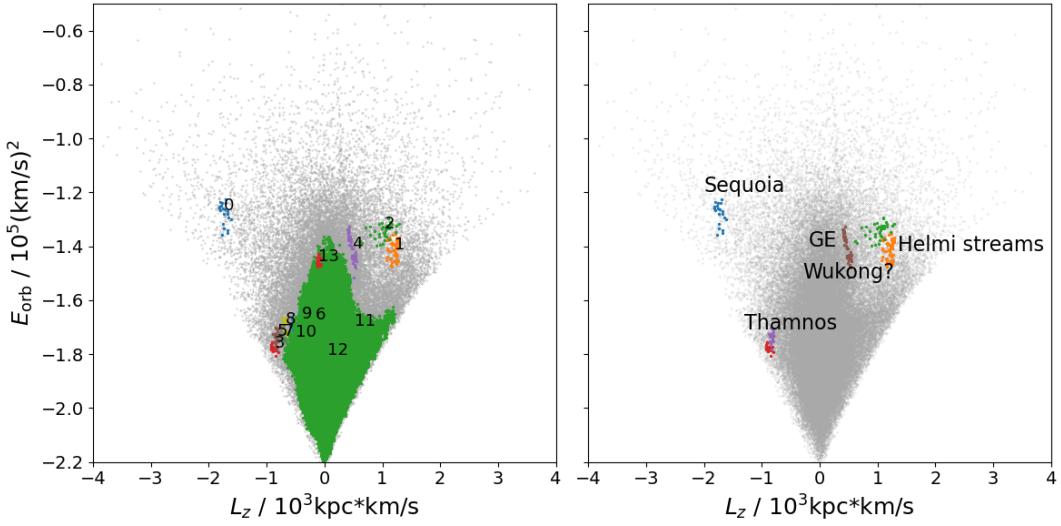


Figure 52: Left panel: All 14 substructures identified by HDBSCAN shown in  $E_{\text{orb}} - L_z$ -space. Right panel: The few selected substructures corresponding to known substructures. The positions of the substructure in this space compare favourably to Figure 4, which was taken from Ruiz-Lara et al. (2022).

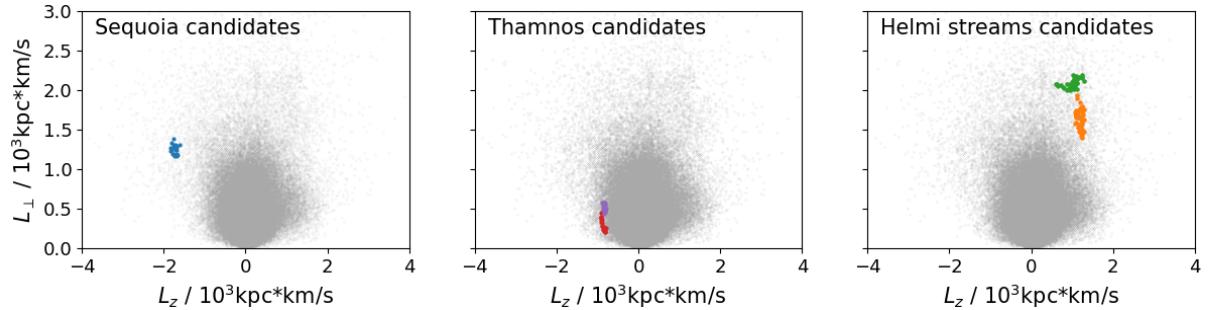


Figure 53: Distribution of the substructures chosen in Figure 52 in  $L_{\perp} - L_z$ -space.

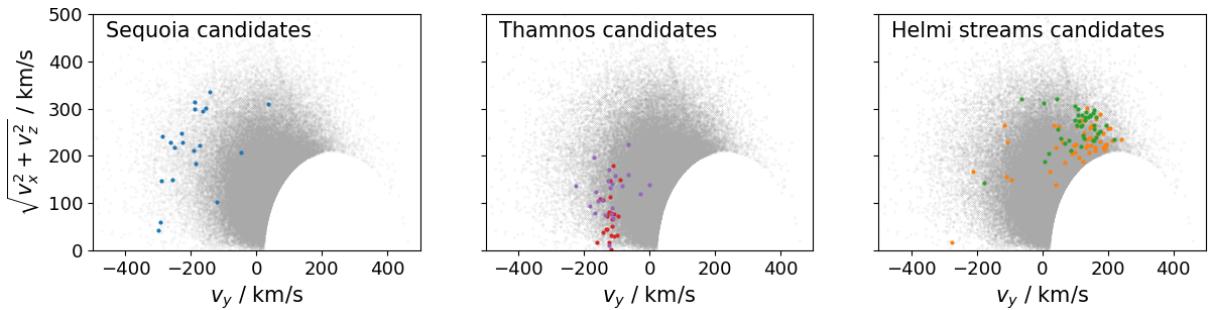


Figure 54: Distribution of the substructures chosen in Figure 52 in  $v_{\perp} - v_y$ -space.

Cluster 1 and 2 correspond to the Helmi streams, with their characteristic high  $L_{\perp}$  being visible in Figure 53, and the positions in the velocity space being consistent with the literature as well, albeit again with more scatter than visible in, e.g., Figure B.1 in Ruiz-Lara et al. (2022). Cluster 3 and 5 are Thamnos candidates, as can be seen via comparison to previous works identifying this substructure, and cluster 3 in particular can be seen to have very good agreement with Figure B.5 and B.6 in Ruiz-Lara et al. (2022) in  $v_{\perp} - v_y$ -space, whereas both are in similarly good agreement in  $L_{\perp} - L_z$ -space. Finally, the position and shape of cluster 4 in Figure 52 corresponds to that reported by Naidu et al. (2020) as a substructure named Wukong, but the existence of this structure has been called into question by other authors, e.g. Ruiz-Lara et al. (2022). When comparing the results presented in Figure 52 to the literature, it should

Table 5: The median properties of the members of each substructure identified in the halo by HDBSCAN. The gravitational potential used is [McMillan \(2017\)](#).

#	Substructure	$N$	$E_{\text{orb}} / (\text{km/s})^2$	$L_z / \text{kpc}\cdot\text{km/s}$	$L_{\perp} / \text{kpc}\cdot\text{km/s}$	[Fe/H] / dex	$\sigma_{[\text{Fe}/\text{H}]} / \text{dex}$
0	Sequoia	20	-127 000	-1740	1260	-1.28	0.53
1	Helmi str. 1	52	-141 000	1210	1650	-1.42	0.60
2	Helmi str. 2	41	-133 000	1040	2090	-1.56	0.43
3	Thamnos 1	23	-178 000	-869	275	-1.63	0.43
4	Wukong	41	-141 000	472	875	-1.25	0.44
5	Thamnos 2	22	-173 000	-826	508	-1.40	0.66

be noted that the choice of potential plays a role for the specific location in  $E_{\text{orb}}$ -direction, as the offset between the [McMillan \(2017\)](#) potential used in this work and, for example, the alternative potential calculated in Section 5.2 is about  $2\sim3 \cdot 10^4$  (km/s)<sup>2</sup>. As it turns out, this alternative potential is actually very similar to the one used by [Ruiz-Lara et al. \(2022\)](#) (see Appendix A.1), meaning that for the sake of the comparison of the substructures' positions in  $E_{\text{orb}}-L_z$ -space to their work, it might have been more useful to use this potential. However, this would merely affect the concrete value given for  $E_{\text{orb}}$  and would otherwise not really change the appearance of the data in  $E_{\text{orb}}-L_z$ -space.

In Table 5, the six selected structures are listed alongside the median properties of its members. Gaia-Enceladus is not listed here, as it is very difficult to clearly separate from the rest of the sample. The median metallicities have very large scatter, which might be due to the small numbers  $N$  in each cluster, but this also suggests that the samples might be contaminated by stars not part of the same substructure. Due to the large spread in metallicities and small number of members, it is difficult to make a meaningful comparison to the literature. However, Sequoia is notably not very metal-poor, which goes counter to the expectation that such a retrograde structure should be more metal-poor than many of the less retrograde structures. In any case, the identification of cluster membership in kinematic space without using additional chemical information can be a challenging matter, due to smearing of some structures in kinematic space, as well as overlap of different structures (e.g., [Naidu et al. 2020](#); [Koppelman et al. 2019a](#)). This further underlines the importance of large-scale spectroscopic surveys such as 4MOST in order to supply extensive datasets of chemistry that can be used for chemical tagging and incorporated into the identification and characterisation of substructures, in particular in the stellar halo.

Another way to make the characteristic Gaia-Enceladus structure visible and separate it further from the thick disk is to restrict the sample to the nearby halo, more specifically only stars at  $\pi > 0.4$ , which corresponds to stars within 2.5 kpc of the Sun. Figure 55 shows kinematic plots for this selection, where the presence of Gaia-Enceladus is very clearly visible in both the density plot and the normal plot, and the plot compares very well to examples from the literature such as [Helmi et al. \(2018\)](#); [Lövdal et al. \(2022a\)](#); [Koppelman et al. \(2018\)](#). The structures identified as Sequoia are also visible in the top left part of the  $E_{\text{orb}}-L_z$  plot, as is a hint of the Helmi streams at  $L_z \sim 1600$  kpc·km/s and  $E_{\text{orb}} \sim -1.35 \cdot 10^5$  (km/s)<sup>2</sup>. Further, the characteristic arc at  $v_y \sim -400$  km/s can be seen in the top left panel, although very sparsely populated. Clearly, this method of visualising Gaia-Enceladus is most effective, but it also severely restricts the number of available stars, with only around 12 000 stars being in this figure, and is obviously only meaningful for the local halo. For further certainty, one may restrict the sample to stricter quality requirements regarding the parallax uncertainty and radial velocity uncertainty limits.

In order to incorporate some chemistry into the analysis of the kinematic space, the distribution of  $E_{\text{orb}}$  vs.  $L_z$  is depicted in different metallicity cuts in steps of -0.25 starting at [Fe/H] = -0.5 and ending at [Fe/H] = -2.0 (Figure 56). For this figure, stricter quality requirements were introduced to ensure the reliability of the data, namely  $\frac{\pi}{\sigma_{\pi}} \geq 5$  and  $\sigma_{v_r} < 4$  km/s, but the proximity condition that only stars closer than 2.5 kpc be included is dropped. The structure identified as Gaia-Enceladus is at its strongest in the range [Fe/H]  $\sim -1.0$  to  $-1.5$ , which matches the range of the peak of its metallicity distribution

according to Ruiz-Lara et al. (2022). Other than that, it seems that the distribution in the energy–angular-momentum-space becomes more homogeneous at the low metallicity end  $[\text{Fe}/\text{H}] < -2.0$ .

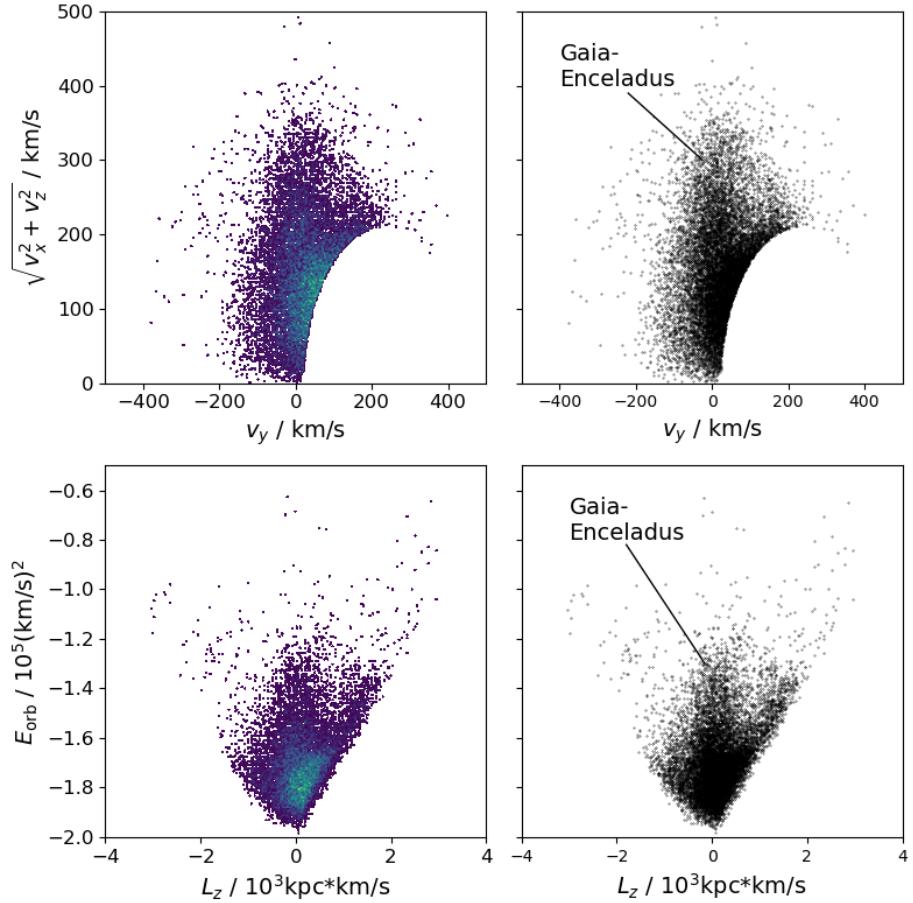


Figure 55:  $v_{\perp}$  vs.  $v_y$  and  $E_{\text{orb}}$  vs.  $L_z$  for the subsample of stars closer than 2.5 kpc from the Sun. Gaia-Enceladus is clearly visible at relatively high orbital energies around  $L_z \sim 0$ , although most authors report it to be, on average, slightly retrograde (e.g., Helmi et al. 2018). For Gaia-Enceladus in the top panels, compare also to Figure 3a.

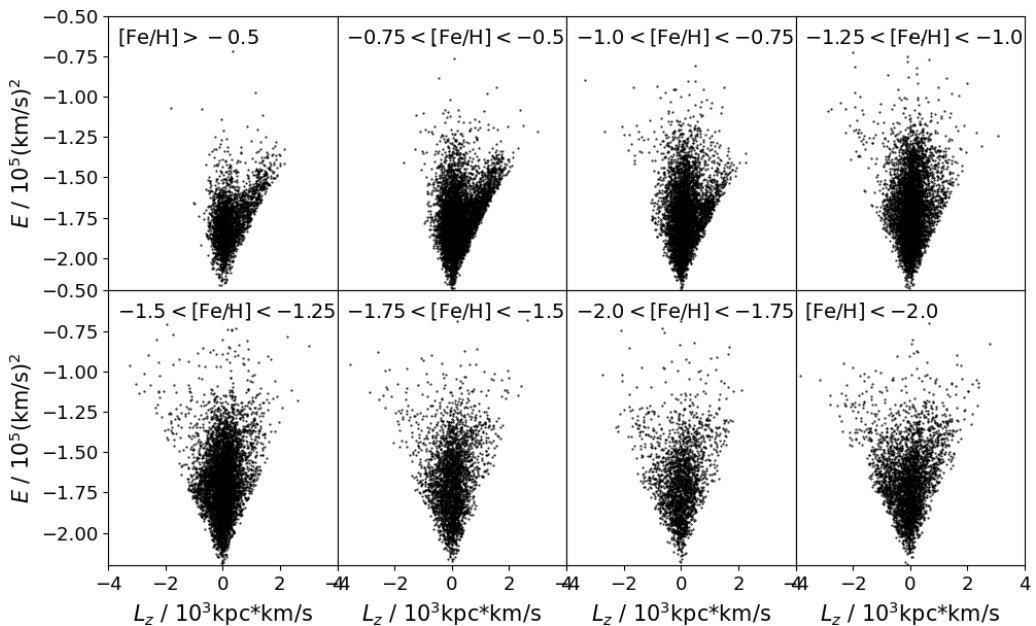


Figure 56:  $E_{\text{orb}}$  vs.  $L_z$  of the halo for different metallicity cuts in steps of 0.25.

The low number of crossmatched stars with GALAH DR3 that are also part of the selected halo prevents a meaningful similar analysis of the kinematics in different cuts with GALAH metallicities. However, the  $\alpha$ -abundances are used in order to make a comparison between the behaviour of the prograde ( $L_z > 0$ ) and the retrograde ( $L_z < 0$ ) stars. For this, the GALAH flags are cleaned, requiring `flag_sp = 0` and `flag_alpha_fe = 0`, which removes even more stars, and in the end, less than 3000 stars are available. Nevertheless, this sample is split into a prograde and retrograde subset, and for each of the two, a histogram is plotted (Figure 57). In order to quantify the general trend of the samples, a single and a two-component Gauss are fitted to the distributions. As many of the retrograde stars have been accreted from dwarf galaxies and these have a lower star formation rate, it is expected that retrograde stars have an  $\alpha$ -underabundance. However, the single Gauss reports almost identical peaks at  $0.22 \pm 0.10$  for prograde and  $0.21 \pm 0.11$  for retrograde, respectively. The two-component Gauss did not identify a second significant peak and mostly follows the single Gauss. This unexpected behaviour of the retrograde stars not having lower  $\alpha$ -abundances might be the result of the small numbers of stars, or may also be related to uncertainties in the GALAH  $\alpha$ -abundances. In addition to GALAH, GSP-Spec abundances were also considered for usage in this work, but later abandoned due to the small number of reliable stars when the flags are appropriately cleaned. More regarding this topic is discussed in Appendix A.2.

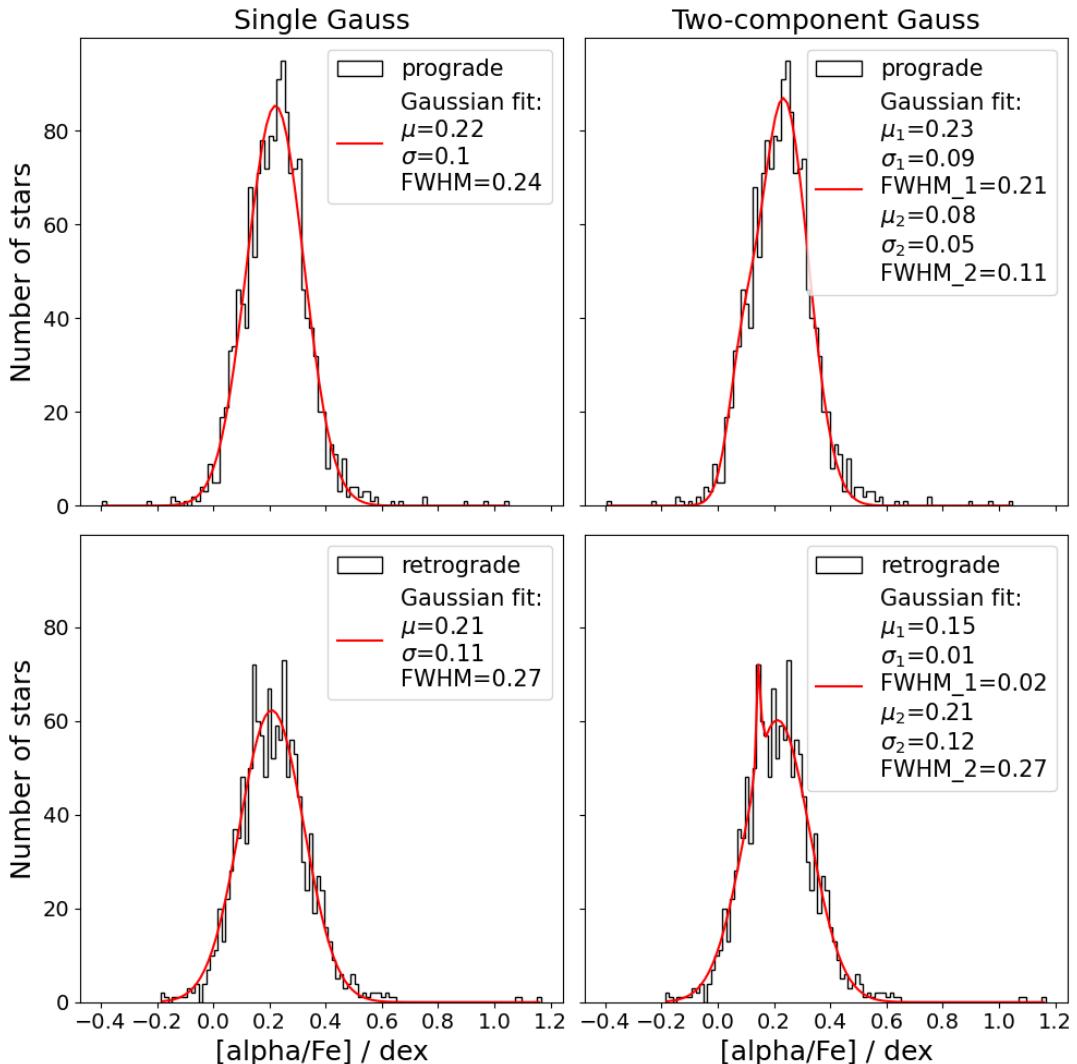


Figure 57:  $\alpha$ -abundance-distribution for prograde and retrograde halo stars including a single Gauss and two-component Gauss fit. Counter to expectations, the distribution of prograde and retrograde stars is very similar.

## 5.5 Identifying candidate n-capture process enriched halo stars

As discussed extensively in Section 1.1, the synthesis of elements beyond the iron peak of stability occurs via neutron capture, with the two general types being the r-process and s-process. However, detailed chemical analyses have led to the discovery that there is likely another component to the r-process. Honda et al. (2006) carry out a high-resolution, high S/N observation of the very metal-poor star HD 122563 in order to pin down the typical characteristics of a weak-r-process-enhanced star. In order to achieve this, they derive abundances for more than 20 mostly neutron-capture elements using synthetic spectra assuming local thermodynamic equilibrium. From their value of  $[Ba/Eu] = -0.5$  dex, they conclude that the heavy neutron-capture elements stem from the r-process, and the s-process plays a neglectable role. Comparing the abundance pattern to the solar-system r-process component, it becomes apparent that for light elements ( $38 \leq Z \leq 47$ ), the abundances in HD 122563 are higher, as are those in the range  $56 \leq Z \leq 59$ , and that even the heavy elements show deviations from the solar system r-process component, albeit to a lesser extent. The authors find that this can only be explained by a separate component of the main r-process, the so-called weak r-process.

In order to identify candidate stars in the 6D kinematic subsample of the 4MOST S2 target catalogue which were enriched by the r-, s-, or weak r-process, the crossmatch to GALAH DR3 is used. In particular, GALAH DR3 provides measurements of the neutron-capture elements Sr, Y, Zr, Ba, La, Ce, Nd, Sm, and Eu for some of its stars, meaning that it should in principle be possible to determine by which of the processes a star was likely enriched based on an element-by-element comparison to the literature estimates of the patterns. For the r- and s-process, the solar r- and s-process components determined by Prantzos et al. (2020, Table 4) as fractions of the Solar elementary abundances provided by Lodders et al. (2009) are used. For the weak r-process, the pattern of the star HD 122563 determined by Honda et al. (2006) is used as a reference. Since the elementary abundances in GALAH DR3 outside of  $[Fe/H]$  and  $[\alpha/Fe]$  can come with significant caveats, the flagging scheme is used to only take into account the abundance of element X if  $\text{flag\_X\_fe} = 0$  meaning that in reality, a lot less than the number of stars in the crossmatch with GALAH DR3 will actually be used. Furthermore, the comparison is only carried out for stars which have measurements of at least 6 elements in GALAH DR3 after the flags are cleaned.

As the solar-system r- and s-process components of element X by Prantzos et al. (2020) are given in the form of fractions  $F_{X,r/s}$  of the Solar abundance of element X  $\log \epsilon(X)_\odot = \log(N_X/N_H)_\odot + 12$ , where  $N_X$  and  $N_H$  are the number densities of element X and hydrogen, respectively, the abundance has to be calculated via

$$\begin{aligned}\log \epsilon(X)_{\odot,r/s} &= \log(F_{X,r/s} \cdot N_{\odot,X}/N_{\odot,H}) + 12 \\ &= \log F_{X,r/s} + \log(N_X/N_H)_\odot + 12 \\ &= \log F_{X,r/s} + \log \epsilon(X)_\odot,\end{aligned}\quad (12)$$

where  $F_{X,r/s}$  is the fraction of an element X being produced by the r/s process.  $\log \epsilon(X)$  abundances for HD 122563, substituting for the weak r-process, are given directly in Honda et al. (2006), but for GALAH DR3, they also have to be derived, as only  $[X/Fe]$  is given. Using the definition of a star's abundance ratio  $[X/H]_\star = \log \epsilon(X)_\star - \log \epsilon(X)_\odot$ , the abundance can be derived as

$$\log \epsilon(X)_\star = [X/H]_\star + \log \epsilon(X)_\odot = [X/Fe]_\star + [Fe/H]_\star + \log \epsilon(X)_\odot, \quad (13)$$

where  $[X/Fe]$  and  $[Fe/H]$  are taken from GALAH DR3, and  $\log \epsilon(X)_\odot$  is taken from the values given in Prantzos et al. (2020), which are based on Lodders et al. (2009). Accordingly, error propagation demands that

$$\sigma_{\log \epsilon(X)} = \sqrt{\sigma_{[X/Fe]}^2 + \sigma_{[Fe/H]}^2}, \quad (14)$$

with the uncertainty in Solar abundances being negligibly small compared to the GALAH uncertainties. Finally, the solar-system r- and s-process component abundance patterns as well as the pattern of HD

122563 need to be scaled to the GALAH data at an appropriate element. For the r-process, Eu is chosen, and Ba is picked for the s-process, but where for the weak r-process Sr would be the usual choice, this is slightly problematic here as less than 20% of stars in GALAH DR3 have reliable Sr measurements, compared to much higher numbers for the other elements used in this analysis. This would mean that the sample from which to pick candidates for the weak r-process is greatly reduced, which is why for the scaling of the HD 122563 pattern, Zr was instead used as an alternative representative of the light neutron-capture elements. After the abundances are scaled, the GALAH DR3 abundances can be compared to the references. Due to being limited to the GALAH crossmatch, the cleaning of flags, requiring at least 6 available elements for the analysis, and some stars not having a measurement of the element used for the scaling, only about 8200 stars remain for the r-process, about 10 100 for the s-process, and about 8000 for the weak r-process analysis.

The approach to quantify the deviation in a star's neutron-capture elemental pattern is chosen to be based on how many standard deviations an abundance reported in GALAH DR3 is away from the scaled reference's abundance, and to take the mean of this deviation across all available elements, i.e.

$$\sigma = \frac{1}{n} \sum_{\text{X}} \frac{|\log \epsilon(\text{X})_{\text{GALAH}} - \log \epsilon(\text{X})_{\text{ref}}|}{\sqrt{\sigma_{\log \epsilon(\text{X})_{\text{GALAH}}}^2 + \sigma_{\log \epsilon(\text{X})_{\text{ref}}}^2}} \quad (15)$$

Here,  $n$  refers to the number of available elements X over which is summed. In the case of the comparison to the solar-system r- and s-process components, the given uncertainties are much smaller than the uncertainties in GALAH DR3, and the second term in the root in Equation 15 is neglected. Figure 58 shows histograms of the computed mean deviations for the three reference patterns. As can be seen, the distribution looks pretty similar in all cases, except for the s-process histogram having comparably slightly more objects in the range  $\sigma < 2$ , indicating that more stars show agreement with the pattern of the solar-system s-process component. It also shows that generally, most stars fall outside the  $2\sigma$ -range, with the peak of the distributions being at around  $3\sigma$ .

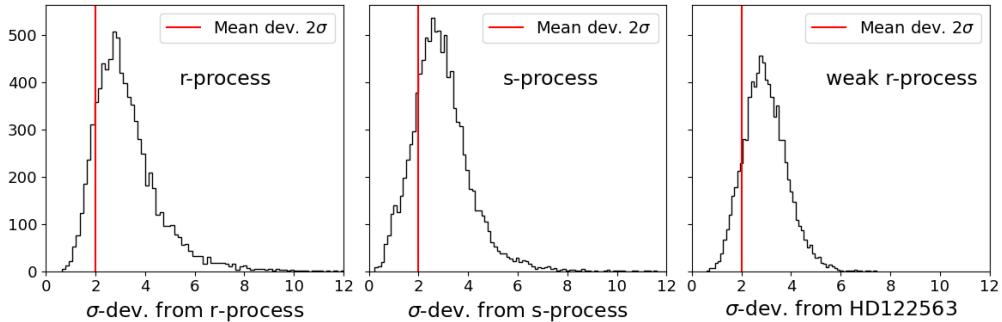


Figure 58: The distribution of the mean  $\sigma$ -deviations between the GALAH pattern and the reference pattern. The red line indicates the  $2\sigma$ -range.

In order to get a better overview of the deviation behaviour between the stars and the references, the deviations from the different patterns are depicted in a two-dimensional plot where the deviation from one reference pattern is on the  $x$ -axis, and the deviation from another one is on the  $y$ -axis (Figure 59). In addition, HD 122563 as a representative of weak-r-process-enriched stars and CS 22892-052 as a representative of r-process-enriched stars are also included in the figures after their deviations from the reference patterns were computed as described above (for the abundances used for CS 22892-052, see Sneden et al. 2003). Their positions are generally consistent with the assumptions, as HD 122563 shows deviations of more than  $5\sigma$  from both the solar-system r- and s-process component's pattern, whereas CS 22892-052 has a significant deviation of around  $7\sigma$  from the s-process, a little bit less at  $5\sigma$  from the weak r-process, and a rather small deviation from the r-process of only  $2\sigma$ . Also, the two top panels in the figure depict the same space, but in the right panel, the rough division into r-process-enriched,

s-process-enriched, and r/s-process-enriched stars according to the definitions by Beers & Christlieb (2005) is colour-coded in red, blue, and green colour, respectively. This illustrates nicely the transition between different enrichment processes in this diagram, and also shows that towards the origin of the coordinate system, the three overlap, as the patterns here are similar to both the r- and s-process patterns.

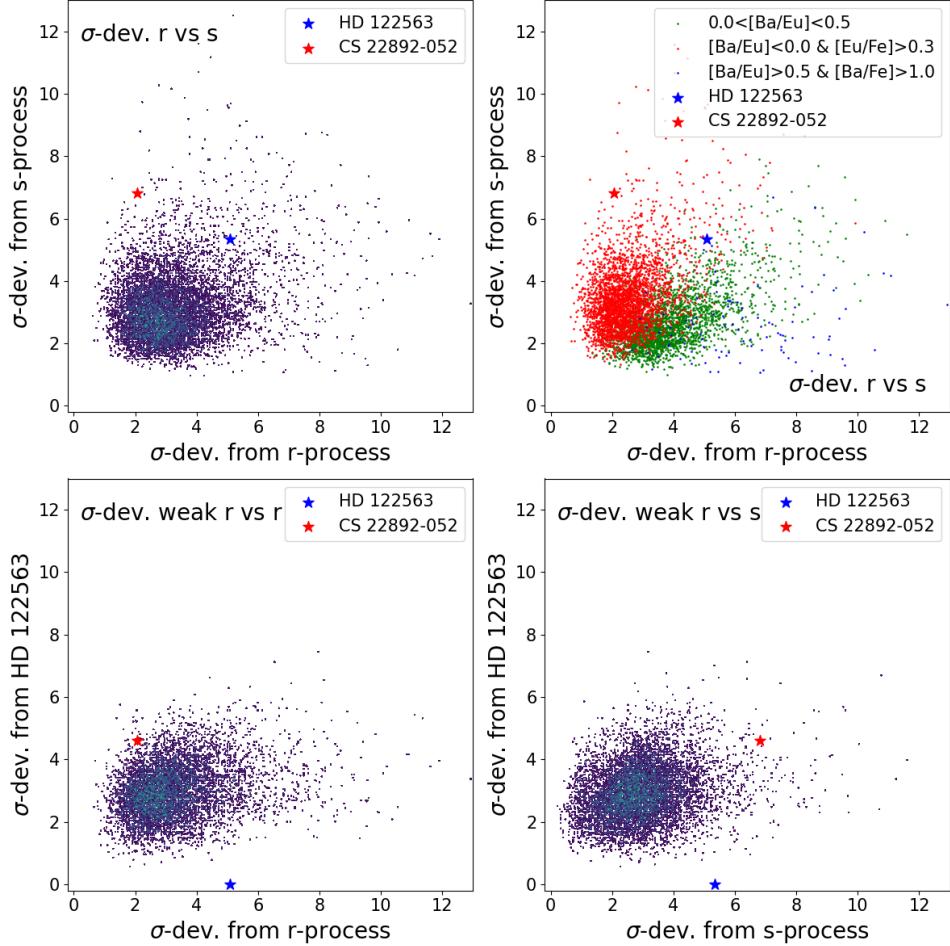


Figure 59: The deviations of the stars' elemental patterns from the three references are shown in two dimensions.

From this data, candidates for each of the three enrichment processes are selected. At this time, the selection is still rather arbitrarily chosen at

- $(\sigma_r < 2.5)$  and  $(\sigma_s > 3)$  and  $(\sigma_{\text{HD}} > 3)$  for the r-process candidates
- $(\sigma_s < 2.5)$  and  $(\sigma_r > 3)$  and  $(\sigma_{\text{HD}} > 3)$  for the s-process candidates
- $(\sigma_{\text{HD}} < 2.5)$  and  $(\sigma_r > 3)$  and  $(\sigma_s > 3)$  for the r-process candidates.

This selection was implemented to ensure that the pattern of a star is not only similar to the reference it is a candidate for, but also at least somewhat different from the patterns of the other references. It also only selects a small minority of the full sample as candidates in each case, but still enough to be able to take a look at the statistical properties of the member stars. Via this selection, three samples of 319 r-process-enriched, 253 s-process-enriched, and 283 weak-r-process-enriched candidate stars, respectively, are identified. In Figure 60, several properties of these candidate samples are depicted, such as their distribution in the sky, their brightness, their colour, as well as Kiel diagrams based on the effective temperatures and surface gravities provided in GALAH DR3. The apparent magnitude histograms show that most stars are at the brighter end of GALAH's observations, which is because those at the faint end

are more likely to not meet the requirements regarding clean flags and minimum number of elements. However, this makes them attractive targets for follow-up observations with high-resolution spectroscopy to confirm the candidate selection.

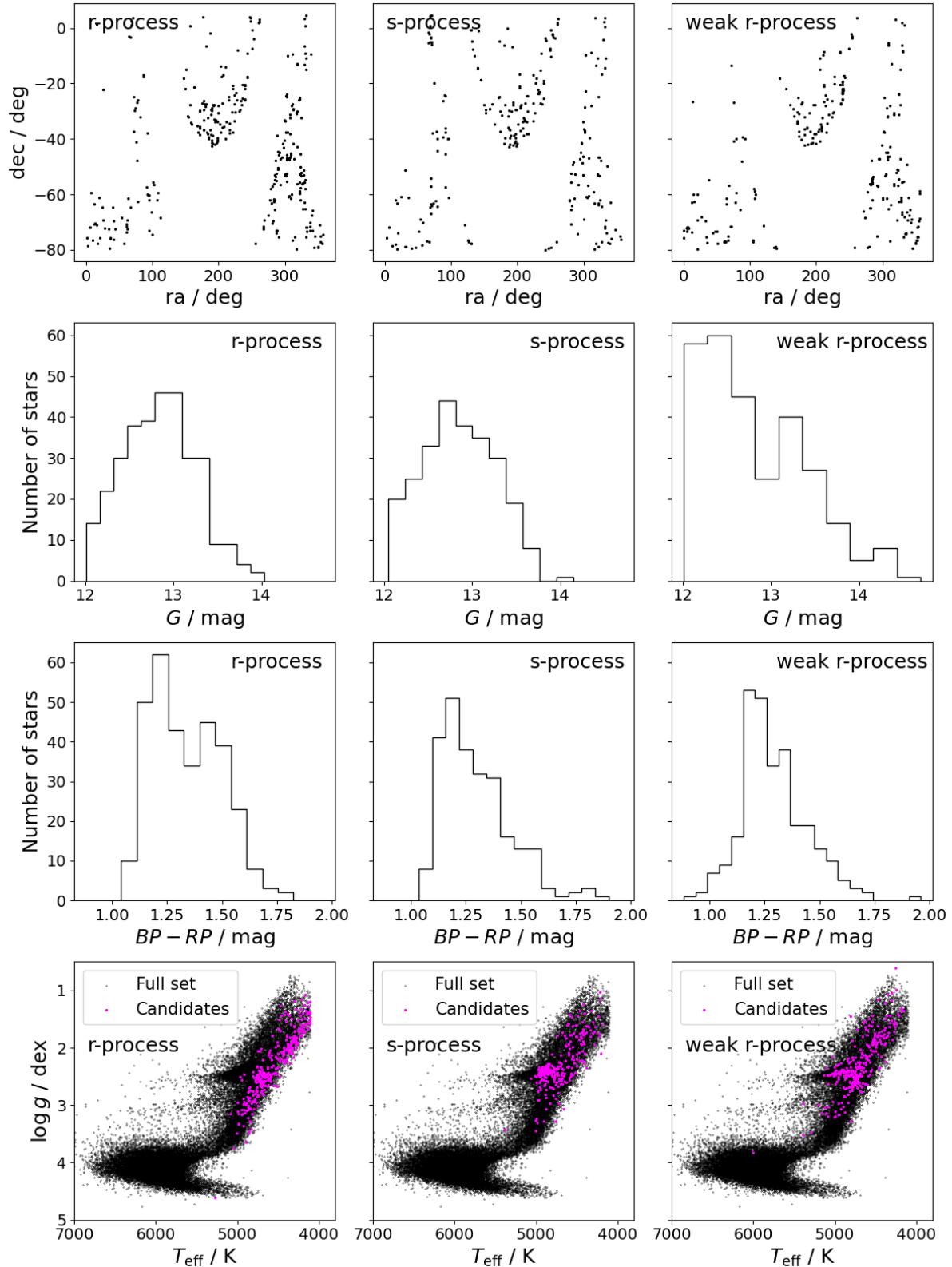


Figure 60: Some statistical properties of the selected candidates. The stellar parameters in the plots in the lower panels were cleaned using `flag_sp = 0`.

The colour histograms put basically all candidates above  $G_{\text{BP}} - G_{\text{RP}} = 1.0$  into a relatively red range, and it can also be seen in the Kiel diagrams that for all three candidate samples the temperatures are rather cool. Also, all three candidate samples appear to consist exclusively of giant stars, but differences in the distributions exist: While the r-process-enriched candidates sit on the cool end of the full sample and are mostly evenly distributed across a surface gravity range from  $\log g \sim 3.0$  to 1.5, the s-process-enriched candidates mostly clump at about  $(T_{\text{eff}}, \log g) = (5000 \text{ K}, 2.5 \text{ dex})$  in what is likely the red clump. The weak-r-process-enriched candidates are again a bit more evenly distributed, but also clearly have a higher number of stars near the red clump and extend slightly into the horizontal branch. It should be noted that it is possible that effects arising from GALAH DR3 quality variations across stellar parameters are at play here, as the vast majority of stars that have enough available elements for the procedure described above turn out to be giant stars in the red range.

In Figure 61, the elemental pattern of an example star picked out of each of the candidate samples is plotted against the pattern of the reference it is a candidate for in order to visualise how well the patterns agree with each other. They showcase such close agreement with the reference that this is unlikely to be random, and show that this method indeed appears to select viable candidates, although the very low corresponding  $\sigma$ -deviations of these three example stars from the references shown in the figure reveal these to be among the more certain candidates, and a detailed verification of the selection and extension of the analysis is necessary.

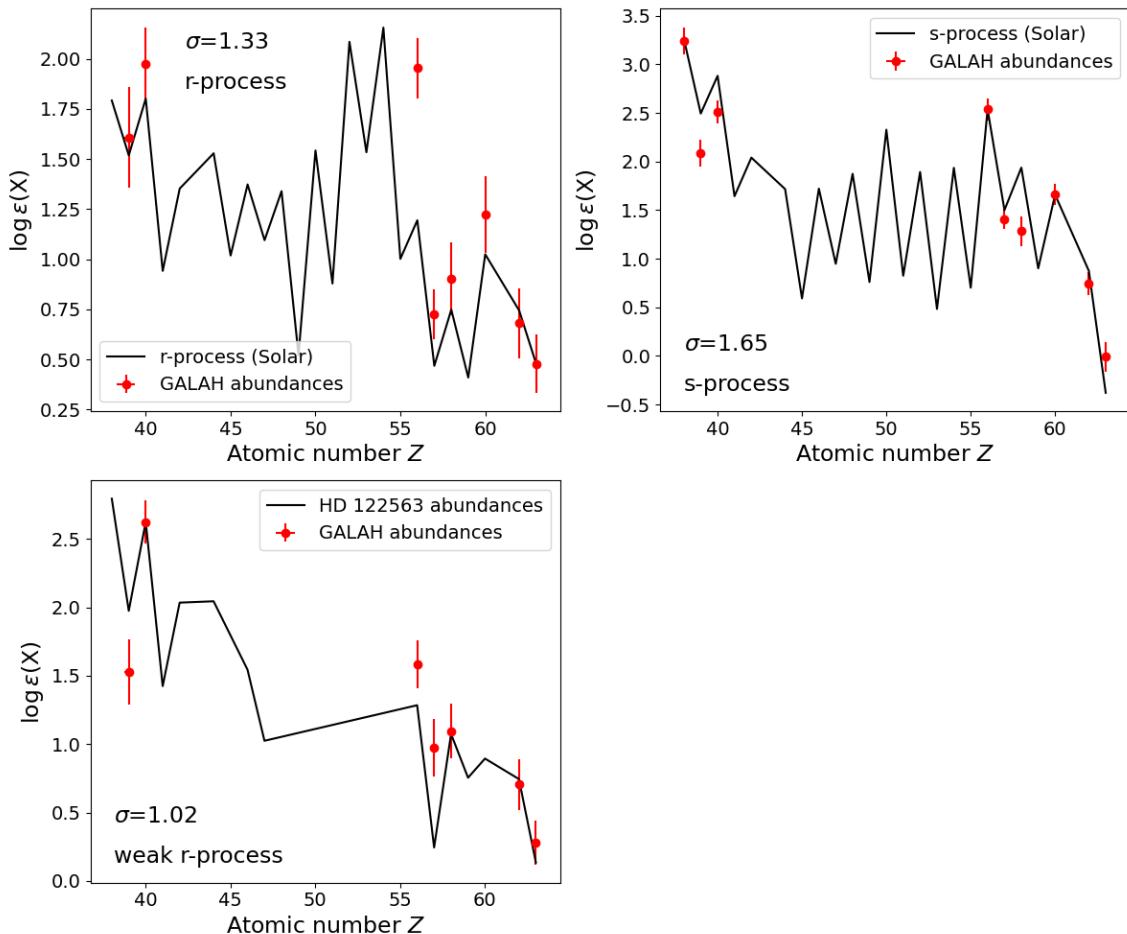


Figure 61: Direct comparison between the patterns of a selected star and the reference it is a candidate for. The examples were hand-picked and not all candidates show such impressive agreement, as the small corresponding  $\sigma$ -deviations given in the bottom left corners show, but the ones seen during the example selection process all followed the patterns at least decently well.

There are a few caveats regarding the procedure in this section that have to be properly addressed in the future. First of all, some kind of reliability scheme or flagging system for the candidate selections based on various factors should be developed. For example, between a star with small uncertainties and a star with much larger uncertainties in the abundance estimates that both have the same small  $\sigma$ -deviation from the reference, the former is much more reliably associated with that enrichment process, as stars with larger uncertainties reach low  $\sigma$ -deviations more easily due to the nature of Equation 15, although any selection incorporating the uncertainties of the measurements will have to address this matter. Another aspect to be included in this is the number of available abundance measurements upon which the mean  $\sigma$ -deviation is based, as more elements should yield more reliable results. For example, the number of available elements for each star may be colour-coded into the 2D plots (Figure 59) to check whether there is a systematic effect of stars with less available elements tending towards smaller or higher  $\sigma$ -deviations. Furthermore, simply taking the mean leaves the calculation of the deviation between patterns vulnerable to strong influence by outliers, but the median is also not a safe choice for such small numbers.

Furthermore, the 2D plots of  $\sigma$ -deviations contain only two reference stars as of now, a number which can and should be increased in order to get a better idea of where known stars fall in this 2D deviation space, and thus learn more about this space's properties. It will also provide an estimate on how much scatter is expected within stars enriched by the same process, and which amount of deviation from the reference is still realistic for candidates. Individual well-studied stars or stellar databases such as SAGA can be used to achieve this. Another matter to be taken note of is that Sr abundances are not available for many of the stars, which leaves only Y and Zr as representatives of the light neutron-capture process elements, and for the comparison with the weak r-process pattern, Zr is already fixed due to scaling the reference there. This leaves in most cases only Y to contribute to the  $\sigma$ -deviation, which can lead to an increased unreliability in the selection of weak r-process candidates. In fact, this analysis is almost entirely reliant on the quality of the GALAH abundances, which also has the potential to influence results significantly even after the flags are cleaned.

The analysis described here is restricted to the crossmatch of the 6D subset of the 4MOST S2 target catalogue to GALAH DR3, as this catalogue is the subject of this thesis. However, it can be extended to the full GALAH DR3 catalogue, and the corresponding figures that were presented in this section are shown for the extension to the full catalogue in Appendix A.3. The higher number of stars reveals some structure to the distribution of stars in the 2D plot, which is a topic that should be researched more quantitatively, for instance using appropriate clustering algorithms to identify different regions such as characteristic parameter ranges for each process, and connect their presence to physical reasons. In particular, a scheme to identify stars enriched by both the r and the s process (r/s process stars) would also be desirable, as these stars appear to make up a significant portion of the sample judging from the rough division colour-coded into the upper right panel of the 2D plots. Finally, the insights gained from this study can be used to check for the neutron-capture enhancement patterns of halo substructure, for instance to confirm the r-process enhancement reported for several halo substructures in the literature (Matsuno et al. 2021; Aguado et al. 2021; Roederer et al. 2018). Follow-up high-resolution spectroscopic observations of some selected targets from the candidate samples should be carried out in order to confirm or disprove the usefulness of this approach.

## 6 Results and conclusions

In this section, the most important results are summarised and overall conclusions are drawn concerning the analysis. In Section 4.1, colour-magnitude diagrams were produced for the 4MOST S2 target catalogue including several PARSEC isochrones (Bressan et al. 2012) representing different metallicities (Figure 20). For the absolute magnitude, the parallax was corrected before inverting it according to Lindegren et al. (2021a) to obtain the distance if  $\pi/\sigma_\pi \geq 3$ , and the extinction in the  $G$  band was determined using the Total Galactic Extinction map by Gaia DR3 (Delchambre et al. 2023). The colour-magnitude diagrams shows that for the full catalogue, the main sequence turnoff heavily dominates, and particularly in the deep survey, there are no more giant stars, which is attributed to the low stellar density of the halo at faraway distances. The characteristic metallicity according to the isochrones was shown to be around  $[\text{Fe}/\text{H}] = -0.5$  and indicates that the sample is heavily dominated by disk stars, with Figure 22 also stating how the number of objects drops significantly for lower metallicity cuts, as more of the disk is removed. Finally, the large scatter present in the colour-magnitude diagrams results from cosmic scatter, uncertainties in the reddening estimates, and the large spread in metallicities.

In the next step, the metallicities derived by Martin et al. (2023) and used in the creation of the target catalogue were verified, as the metallicity cut is an important ingredient for the appropriate selection of the catalogue (Section 4.2). This was done via various reference catalogues including GALAH DR3 (Buder et al. 2021), SAGA (Suda et al. 2008), data-driven metallicity estimates by Andrae et al. (2023b), and APOGEE DR17 (Majewski et al. 2017). For this, crossmatches and cross-identifications were carried out, and the crossmatch to GALAH DR3 in particular was used at several points in the remainder of the analysis. The verification of the Martin et al. (2023) metallicities with reference catalogues shows good agreement overall, and the offset between their metallicities is around  $-0.1$  for GALAH, the reliable subset of the data-driven metallicities, and APOGEE, meaning the references report higher metallicities (Figure 24, 27, 28). Only for SAGA, the offset is positive at  $0.11$  (Figure 25), but the overall trend of the comparisons indicates that the Pristine metallicities generally underestimate the metallicities of stars.

The comparisons display mostly uniform behaviour down to about  $[\text{Fe}/\text{H}] \sim -2.5$  to  $-3.0$ , but there is a lack of a significant number of reference objects below  $-3.0$ . Less reliable estimates on the faint end were found, but are within expectations, as Martin et al. (2023) warn that their results are most reliable at  $G_{\text{BP}} < 16$  mag. Furthermore, the comparison to the two data-driven samples by Andrae et al. (2023b) showcases that the Pristine method remains more suitable in the case of 4MOST S2. Finally, the effect that carbon enhancement has on the Pristine-Gaia-Synthetic metallicity estimates was briefly explored, and a comparison of carbon-enhanced and carbon-normal stars according to SAGA yielded a difference of  $0.36$  dex in the offset between the Martin et al. (2023) and the SAGA metallicities, suggesting that carbon-enhanced stars have an overestimated metallicity (Figure 29). In particular, the offset between carbon-normal stars and SAGA is  $0$ , suggesting that the overall offset reported as  $0.11$  is driven by the systematic effect of overestimating the metallicity of carbon-enhanced stars. Even at  $[\text{Fe}/\text{H}] < -1.0$ , the fraction of carbon-enhanced stars was found to be only about  $0.45\%$  using the catalogue by Lucey et al. (2023), which is still about  $2.5$  times as much as for the full 4MOST S2 catalogue, indicating the growing fraction of CEMP stars at lower metallicities.

In order to verify whether the selection process shows any unexpected behaviour, the selection criteria were applied to the GALAH DR3 catalogue step by step, and furthermore, a brief comparison between the selected stars and the full catalogue was carried out in Section 4.3. The stepwise application did not yield any unexplained effects, with most stars being removed by the survey area (Figure 30) and metallicity conditions, and the final number of  $\sim 65\,000$  stars that is expected to be contained within the 4MOST S2 target catalogue agrees well with the  $63\,000$  stars found in the crossmatch. However, Section 4.2 showed a systematic offset in metallicities between GALAH and 4MOST S2, making the final cut not necessarily an accurate representation of the metallicity cut applied to select the 4MOST S2 catalogue. The application of the selection criteria to SAGA yielded an expectation of about  $1350$ , which is also

consistent with the 1250 stars recovered in the crossmatch. Comparing the  $[\alpha/\text{Fe}]$  vs.  $[\text{Fe}/\text{H}]$  distribution for the crossmatched GALAH sample and the full GALAH DR3 catalogue showcased the increased prominence of what is likely the  $\alpha$ -knee in the crossmatched sample (Figure 31), which is consistent with expectations for a sample with more metal-poor stars. The majority of stars in the crossmatched catalogue seem to have been enriched by the r-process (Figure 32), which might be related to the later onset of the s-process (Simmerer et al. 2004; Käppeler et al. 2011) and the reported r-process enhancement in halo substructures (Matsuno et al. 2021; Roederer et al. 2018), although the halo makes up only the minority of the 4MOST S2 catalogue.

As the final part of the verification, the survey criteria were applied to the Besançon Galaxy model (Robin et al. 2003; Czekaj et al. 2014) in order to simulate how the distribution of stars in 4MOST S2 should look according to models (Section 4.4). For this, a transformation into three-dimensional cartesian galactocentric coordinates was carried out using ASTROPY (Astropy Collaboration et al. 2013). The comparison of the  $R-z$ -distribution of the 4MOST S2 sample with the Besançon model showcases overall similarity, but the Besançon model contains significantly more stars at 5.1 million vs. the 4MOST S2 sample's 3.8 million (Figure 33). The deviations that were found were mostly a larger amount of stars in the Besançon model at small  $R$  and close to the Sun, and for both cases, this is much reduced when including only stars of metallicity less than  $-1.0$  (Figure 34). The main explanations for the deviations include a difference in the metallicity scale, as it was also shown that small differences in metallicities can make a significant difference for the number of included objects, and an underestimate of the extinctions used, as the extinction becomes increasingly significant in the direction of the Galactic centre. Both explanations are consistent with a more quantitative approach comparing the stellar counts in individual cells of size 0.5 kpc (Figure 36), and Martin et al. (2023) also stress that the main effort of the Pristine method was put into yielding reliable estimates for the more metal-poor stars. In any case, the improved agreement between 4MOST S2 and the model at lower metallicities is reassuring, as 4MOST S2 targets metal-poor halo stars in particular. Overall, the verification carried out in Section 4 indicates that the selection was successful in building up the intended target catalogue for the 4MOST S2 survey.

As Section 5 provides a kinematic study of the 4MOST S2 target catalogue, only the subset of the catalogue which includes radial velocity measurements was used, which is about 24% or 909 000 stars. In Section 5.1, the radial velocities and their uncertainties were briefly characterised, and found to agree very well with references despite a larger median uncertainty of about 4 km/s (Figure 39). As only 5% of radial velocities were found to have uncertainties of more than 10 km/s (Figure 38), these were excluded from the rest of the analysis to obtain more reliable kinematics. Transformations to the galactocentric cartesian coordinates, in which various kinematic parameters were determined, were carried out using ASTROPY (Astropy Collaboration et al. 2013) for the positions, and the transformations by Johnson & Soderblom (1987) for the velocities (Section 5.2). For the orbital energy, the gravitational potential of McMillan (2017) was used. The separation of the halo from the disk was carried out using the traditional velocity condition  $|v - v_{\text{LSR}}| > 210$  km/s, as well as a second approach based on the circularity parameter (Sotillo-Ramos et al. 2023; Zhu et al. 2022). The two different halo selections were found to vary widely in their size, with the former leaving about 78 000 stars and the latter about 116 000. A first look at the heliocentric distance distribution of the two selections (Figure 42) shows that the circularity-based selection left in comparably more closeby stars assumed to be part of the disk.

In Section 5.3, the halo was characterised regarding its spatial, evolutionary, and chemical properties. Comparisons of the distribution of distances from the Galactic midplane for the two halo selections to Juric et al. (2008) show good agreement for both (Figure 44). The model fits carried out to fit a single and a broken power-law model of the halo's stellar number density distribution yielded  $\alpha \approx 4.4$  for the single power-law in both selections, while for the broken power law the results were  $\alpha = 4.31$  and  $\beta = 4.8$  in the case of the circularity-based selection, and  $\alpha = 3.89$  and  $\beta = 5.01$  in the case of the velocity-based selection (Figure 45). These fit parameters are a bit higher than most reported in the literature (Table

4), indicating that the numbers are dropping a bit too fast at growing radius, which hints at an incompleteness that might be the result of the apparent magnitude limitation and is consistent with results from Section 4.4. Colour-magnitude diagrams of the halo selections show that the selection effects that are clearly visible for the full sample are no longer visible for the halo selections (Figure 46), indicating that the selection criteria for 4MOST S2 are successful in not removing any halo stars in photometric space. From the isochrones, the characteristic metallicity was found to be  $-1.0$  or even slightly less, which is a significant shift from the full sample's  $-0.5$  and further verifies the halo selections. Finally, the metallicity distribution was plotted for various velocity cuts following the approach of Carollo et al. (2007), and hints of a second peak at about  $[Fe/H] = -2.0$  were found in the more retrograde selections (Figure 47), which is also reported by Carollo et al. (2007) to be a sign of the dual nature of the halo. The main peak remains dominant throughout, however, and shows that the apparent-magnitude-limited 4MOST S2 catalogue mainly samples the inner halo.

In order to identify substructure in the halo (Section 5.4), the 6D subset of the 4MOST S2 target catalogue was depicted in various kinematic spaces in Figure 48, 49, and 50, where the removal of the disk is clearly reflected. However, the circularity-based selection still contains more stars with disk kinematics, showing that the higher number of stars left in this selection is in significant parts due to disk contamination, which is why only the velocity-based selection was used from there on. In the  $E_{\text{orb}}$  vs.  $L_z$  plot of the full halo selection, only faint hints of Gaia-Enceladus can be seen (Helmi et al. 2018; Koppelman et al. 2018), as the distribution appears mostly homogeneous, even after smoothing the distribution and subtracting it as background (Figure 51). To recover other, smaller substructures quantitatively, the hierarchical clustering algorithm HDBSCAN (Pedregosa et al. 2011) was applied to the data to find clusters in the 3D  $E_{\text{orb}}-L_z-L_{\perp}$  space, and was able to recover the most prominent identified kinematic substructures in the halo besides Gaia-Enceladus, which are Sequoia, Thamnos, and the Helmi streams (Figure 52), the general properties of which were listed in Table 5. Their positions in kinematic space were found to be consistent with existing literature, particularly Ruiz-Lara et al. (2022) and Koppelman et al. (2019a). The discovery of substructures even in just the subset of the target catalogue showcases that 4MOST S2 will be very useful in the detailed chemical characterisation of substructures in the halo. Finally, the limitation to the local halo (Figure 55) and a division of the data into different metallicity cuts (Figure 56) make Gaia-Enceladus much more easily discernible.

In Section 5.5, GALAH DR3 elemental abundance data of neutron-capture elements was used to produce candidate samples of stars that were enriched by the r-process, s-process, and weak r-process. This was achieved by comparing the pattern of a star's abundance on an element-by-element basis to reference patterns representing the three processes, which were chosen to be the solar-system r-process component, solar-system s-process component, and the star HD 122563, respectively (Prantzos et al. 2020; Honda et al. 2006). As a measure of similarity between two patterns, the mean of the  $\sigma$ -deviations between the references and GALAH abundances of each available elemental abundance for a star was taken. A two-dimensional space in  $\sigma$ -deviations from different processes was introduced (Figure 59), and candidates for each process were selected by requiring a small deviation from the process the star is a candidate for as well as large deviations from the other two processes. As a result, samples of 319 r-process-enriched, 253 s-process-enriched, and 283 weak-r-process-enriched candidate stars among the 4MOST S2 targets were produced, and their distribution in the Kiel diagram reveals them to exclusively be giant stars, with s- and weak r-process candidates clustering around the red clump, and r-process candidates being more evenly distributed. Current caveats of this final part of the analysis and its need for more detailed selection and flagging schemes, as well as the general outlook on it, were discussed in detail at the end of Section 5.5. Overall, Section 5 encompasses an overview of the kinematic content of the 4MOST S2 target catalogue, and provides candidate samples for kinematic substructures as well as candidate samples for enrichment via specific neutron-capture processes. Thus, it showcases the broad spectrum of scientific topics that can be addressed once detailed chemo-dynamical data collected by 4MOST S2 becomes available.

## 7 Outlook

Several aspects of the analyses presented in this work can be extended in future investigations. In order to have a more complete sample regarding the removal due to uncertain parallaxes, the distance estimates of  $\sim 1.5$  billion stars by [Bailer-Jones et al. \(2021\)](#) can be used, who use Bayesian statistics with a Galaxy-model-based prior to estimate the distances via geometric and photometric data. Furthermore, as the metallicity verification lacks a decent number of reference stars at the lowest metallicity end  $[\text{Fe}/\text{H}] < -3.0$ , the addition of catalogues and individual objects beyond what is contained in SAGA is necessary to confirm the accuracy of the Pristine calibration even in the range of extremely metal-poor stars. For the comparison between the Besançon model and 4MOST S2, the suspected influence of the metallicity deviations should be studied in more detail, for instance by checking the appearance of the figures if the cutoff for the Besançon model is shifted around slightly, as even small differences have been shown to yield significant changes in numbers. A further subject of verification could be a closer inspection of object densities across the sky, and how this plays into the observing strategy.

Although radial velocities by GALAH, APOGEE, and RAVE were briefly used to validate the Gaia DR3 radial velocities, the calculations within the scope of the kinematic analysis were done using only Gaia DR3 radial velocities. Complementing this sample with radial velocities from other surveys such as GALAH, APOGEE, RAVE, or LAMOST would not only extend the sample size, but also provide alternative estimates with lower uncertainties in some cases. In fact, as the uncertainties of the other surveys tend to be smaller than Gaia DR3 (although the number of stars is also much smaller), this would be particularly beneficial to a high-quality subset of the data, which would have many more stars available for analysis in that case. Furthermore, the halo selection was carried out only based on kinematic parameters, and can possibly be improved by incorporating data on metallicity and photometry, or by carrying out a photometric selection, for example as shown in [An et al. \(2013\)](#) or [Youakim et al. \(2020\)](#). A more sophisticated method of halo separation would be desirable to keep disk contamination at a minimum for properly constraining properties of the halo and its substructures.

The search for kinematic substructure can also be optimised and extended. Additional parameters such as the orbital eccentricity, circularity, or – depending on availability – more detailed chemical information should be added to the data in order to carry out hierarchical clustering in this high-dimensional space. Some authors also use the actions instead of angular momenta to define their intergrals-of-motion space, but their determination can be rather challenging ([McMillan & Binney 2008](#); [Myeong et al. 2018a](#)). Furthermore, alternative algorithms should be employed as well, for instance STREAMFINDER ([Malhan & Ibata 2018](#)), which can find even thin stellar streams and is especially effective for the high-precision Gaia data. Further ideas expanding on the analysis carried out in Section 5.4 include an evaluation of the probabilities of cluster membership given by HDBSCAN, as well as the overall significance of the clusters, and to compare the observed distributions in kinematic space with a random sample (achieved by shuffling velocity components of the sample; see, e.g., [Helmi et al. 2017](#)). This random sample can be subtracted from the actual sample to make structures more visible. Furthermore, Gaia-Enceladus should also be identified in the metallicity distribution functions of different kinematic selections. Finally, the  $\sim 1700$  high-velocity stars found in the sample also merit further inspection. As the outlook on Section 5.5 was already provided there, it shall only be restated that follow-up observations of some candidates are necessary to assess the performance of the approach.

With new data releases of several catalogues used in this thesis being planned within the foreseeable future and other upcoming wide-angle spectroscopic surveys such WEAVE or MOONS, the means available to analyse the Universe from all sides and angles will only increase. In particular, the Gaia mission will continue to increase the accuracy of data from previous releases, and extend the dataset even further in future releases. As the start of the 4MOST survey’s science operations draws nearer, so do the high-resolution spectra of millions of stars in the Southern sky, providing the science community with a wealth of high-quality data to be exploited for years to come.

## References

- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017a, ApJL, 848, L13
- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017b, Phys. Rev. Lett., 119, 161101
- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017c, ApJL, 848, L12
- Abdurro'uf, Accetta, K., Aerts, C., et al. 2022, ApJS, 259, 35
- Aguado, D. S., Belokurov, V., Myeong, G. C., et al. 2021, ApJL, 908, L8
- Akiyama, M., Smedley, S., Gillingham, P., et al. 2008, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7018, Advanced Optical and Mechanical Technologies in Telescopes and Instrumentation, ed. E. Atad-Ettedgui & D. Lemke, 70182V
- Amarsi, A. M., Lind, K., Osorio, Y., et al. 2020, A&A, 642, A62
- An, D., Beers, T. C., Johnson, J. A., et al. 2013, ApJ, 763, 65
- Andrae, R., Fouesneau, M., Sordo, R., et al. 2023a, A&A, 674, A27
- Andrae, R., Rix, H., & Chandra, V. 2023b, ApJS, 267, 8
- Angulo, C., Arnould, M., Rayet, M., et al. 1999, Nucl. Phys. A, 656, 3
- Antoja, T., Helmi, A., Romero-Gomez, M., et al. 2018, Nature, 561, 360
- Arentsen, A., Placco, V., Lee, Y., et al. 2022, MNRAS, 515, 4082
- Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, ApJ, 935, 167
- Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
- Azais, N., Barden, S. C., Smith, G., et al. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI, ed. C. J. Evans, L. Simard, & H. Takami, 990889
- Babusiaux, C., Fabricius, C., Khanna, S., et al. 2023, A&A, 674, A32
- Bailer-Jones, C. 2010, MNRAS, 403, 96
- Bailer-Jones, C. 2011, MNRAS, 411, 435
- Bailer-Jones, C., Rybizki, J., Fouesneau, M., Demleitner, M., & Andrae, R. 2021, AJ, 161, 147
- Bailer-Jones, C. A. L., Andrae, R., Arcay, B., et al. 2013, A&A, 559, A74
- Baiotti, L. & Rezzolla, L. 2017, Rep. Prog. Phys., 80, 096901
- Barnes, J. 1992, ApJ, 393, 484
- Battistini, C. & Bensby, T. 2016, A&A, 586, A49
- Bauer, F. E., Lira, P., Anguita, T., et al. 2023, The Messenger, 190, 34
- Beers, T. & Christlieb, N. 2005, ARA&A, 43, 531
- Beers, T., Preston, G., & Shectman, S. 1985, AJ, 90, 2089
- Belokurov, V., Erkal, D., Evans, N., Koposov, S., & Deason, A. 2018, MNRAS, 478, 611
- Bennett, M. & Bovy, J. 2019, MNRAS, 482, 1417
- Bensby, T. 2014, in Setting the scene for Gaia and LAMOST, ed. S. Feltzing, G. Zhao, N. A. Walton, & P. Whitelock, Vol. 298, 17
- Bensby, T., Bergemann, M., Rybizki, J., et al. 2019, The Messenger, 175, 35
- Bensby, T., Zenn, A., Oey, M., & Feltzing, S. 2007, ApJ, 663, L13
- Bignone, L. A., Helmi, A., & Tissera, P. B. 2019, ApJL, 883, L5
- Binney, J. & Tremaine, S. 2008, *Galactic Dynamics*, 2nd edn. (Princeton: Princeton University Press)
- Bissantz, N. & Gerhard, O. 2002, MNRAS, 330, 591
- Bisterzo, S., Gallino, R. adn Straniero, O., Cristallo, S., & Käppeler, F. 2010, MNRAS, 404, 1529
- Bisterzo, S., Gallino, R. adn Straniero, O., Cristallo, S., & Käppeler, F. 2011, MNRAS, 418, 284
- Bisterzo, S., Travaglio, C., Gallino, R., Wiescher, M., & Käppeler, F. 2014, ApJ, 787, 10
- Bland-Hawthorn, J. & Gerhard, O. 2016, ARA&A, 54, 529
- Bovard, L., Martin, D., Guercilena, F., et al. 2017, Phys. Rev. D., 96, 124005
- Bovy, J. 2015, ApJS, 216, 29
- Bovy, J., Rix, H.-W., & Hogg, D. W. 2012, ApJ, 751, 131
- Bressan, A., Marigo, P., Girardi, L., et al. 2012, MNRAS, 427, 127
- Bromm, V. & Larson, R. 2004, ARA&A, 42, 79
- Brzeski, J., Baker, G., Baker, S., et al. 2018, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10702, Ground-based and Airborne Instrumentation for Astronomy VII, ed. C. J. Evans, L. Simard, & H. Takami, 1070279

- Buder, S., Sharma, S., Kos, J., et al. 2021, MNRAS, 506, [150](#)
- Carollo, D., Beers, T. C., Lee, Y. S., et al. 2007, Nature, 450, [1020](#)
- Carroll, B. W. & Ostlie, D. A. 2017, [An Introduction to Modern Astrophysics](#), 2nd edn. (Cambridge: Cambridge University Press), p. 893
- Casagrande, L., Lin, J., Rains, A. D., et al. 2021, MNRAS, 507, [2684](#)
- Castelli, F. & Kurucz, R. 2003, in Modelling of Stellar Atmospheres, ed. N. Piskunov, W. Weiss, & D. Gray, Vol. 210, [A20](#)
- Chen, L.-H., Pillepich, A., Glover, S. C. O., & Klessen, R. S. 2023, MNRAS, 519, [483](#)
- Chiappini, C., Minchev, I., Starkenburg, E., et al. 2019, The Messenger, 175, [30](#)
- Christlieb, N., Battistini, C., Bonifacio, P., et al. 2019, The Messenger, 176, [26](#)
- Cioni, M. . R. L., Storm, J., Bell, C. P. M., et al. 2019, The Messenger, 175, [54](#)
- Cirasuolo, M., Afonso, J., Carollo, M., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V, ed. S. K. Ramsay, I. S. McLean, & H. Takami, [91470N](#)
- Cirasuolo, M., Fairley, A., Rees, P., et al. 2020, The Messenger, 180, [10](#)
- Clayton, D. D. 1968, [Principles of stellar evolution and nucleosynthesis](#), 1st edn. (Chicago: University of Chicago Press)
- Clementini, G., Ripepi, V., Garofalo, A., et al. 2023, A&A, 674, [A18](#)
- Collett, T. E., Sonnenfeld, A., Frohmaier, C., et al. 2023, The Messenger, 190, [49](#)
- Cowan, J. J., Sneden, C., Lawler, J. E., et al. 2021, Review of Modern Physics, 93, [015002](#)
- Creevey, O. L., Sordo, R., Pailler, F., et al. 2023, A&A, 674, [A26](#)
- Cropper, M., Katz, D., Sartoretti, P., et al. 2018, A&A, 616, [A5](#)
- Cui, X.-Q., Zhao, Y.-H., Chu, Y.-Q., et al. 2012, Res. Astron. Astrophys., 12, [1197](#)
- Cunningham, M. H., Doel, P., Brooks, D., et al. 2023, Journal of Astronomical Telescopes, Instruments, and Systems, 9, [015002](#)
- Czekaj, M. A., Robin, A. C., Figueras, F., Luri, X., & Haywood, M. 2014, A&A, 564, [A102](#)
- Dalton, G., Trager, S., Abrams, D. C., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V, ed. S. K. Ramsay, I. S. McLean, & H. Takami, [91470L](#)
- Dalton, G., Trager, S. C., Abrams, D. C., et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV, ed. I. S. McLean, S. K. Ramsay, & H. Takami, [84460P](#)
- de Jong, R. S., Agertz, O., Berbel, A. A., et al. 2019, The Messenger, 175, [3](#)
- de Jong, R. S., Barden, S., Bellido-Tirado, O., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V, ed. S. K. Ramsay, I. S. McLean, & H. Takami, [91470M](#)
- de Jong, R. S., Barden, S. C., Bellido-Tirado, O., et al. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI, ed. C. J. Evans, L. Simard, & H. Takami, [99081O](#)
- de Jong, R. S., Bellido-Tirado, O., Chiappini, C., et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV, ed. I. S. McLean, S. K. Ramsay, & H. Takami, [84460T](#)
- De Silva, G. M., Freeman, K. C., Bland-Hawthorn, J., et al. 2015, MNRAS, 449, [2604](#)
- Deason, A. J., Belokurov, V., & Evans, N. W. 2011, MNRAS, 416, [2903](#)
- Deason, A. J., Belokurov, V., & Sanders, J. L. 2019, MNRAS, 490, [3426](#)
- Dehnen, W. & Binney, J. 1998, MNRAS, 294, [429](#)
- Delchambre, L., Bailer-Jones, C. A. L., Bellas-Velidis, I., et al. 2023, A&A, 674, [A31](#)
- Deng, L.-C., Newberg, H. J., Liu, C., et al. 2012, Res. Astron. Astrophys., 12, [735](#)
- DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016, [arXiv:1611.00036](#)
- Di Matteo, P., Haywood, M., Lehnert, M. D., et al. 2019, A&A, 632, [A4](#)
- Dierickx, M. I. P. & Loeb, A. 2017, ApJ, 836, [92](#)
- Driver, S. P., Liske, J., Davies, L. J. M., et al. 2019, The Messenger, 175, [46](#)
- Duncan, K., Baker, A., Best, P., et al. 2023, The Messenger, 190, [25](#)
- Dwelly, T., Merloni, A., Walcher, J. C., et al. 2016, in Society of Photo-Optical Instrumentation En-

- gineers (SPIE) Conference Series, Vol. 9910, Observatory Operations: Strategies, Processes, and Systems VI, ed. A. B. Peck, R. L. Seaman, & C. R. Benn, [99101Q](#)
- Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, AJ, 142, [72](#)
- El-Badry, K., Bland-Hawthorn, J., Wetzel, A., et al. 2018, MNRAS, 480, [652](#)
- Emerson, J., McPherson, A., & Sutherland, W. 2006, The Messenger, 126, [41](#)
- Emerson, J. & Sutherland, W. 2002, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4836, Survey and Other Telescope Technologies and Discoveries, ed. J. A. Tyson & S. Wolff, [35](#)
- Emerson, J. P., Sutherland, W. J., McPherson, A. M., et al. 2004, The Messenger, 117, [27](#)
- Euclid Collaboration, Scaramella, R., Amiaux, J., et al. 2022, A&A, 662, [A112](#)
- Evans, D. W., Riello, M., De Angeli, F., et al. 2018, A&A, 616, [A4](#)
- Fabricius, C., Luri, X., Arenou, F., et al. 2021, A&A, 649, [A5](#)
- Faccioli, L., Smith, M. C., Yuan, H., et al. 2014, ApJ, 788, [105](#)
- Farouqi, K., Kratz, K. L., Pfeiffer, B., et al. 2010, ApJ, 712, [1359](#)
- Feltzing, S., Bensby, T., Bergemann, M., et al. 2018, in Rediscovering Our Galaxy, ed. C. Chiappini, I. Minchev, E. Starkenburg, & M. Valentini, Vol. 334, [225](#)
- Fernández-Alvar, E., Kordopatis, G., Hill, V., et al. 2021, MNRAS, 508, [1509](#)
- Finoguenov, A., Merloni, A., Comparat, J., et al. 2019, The Messenger, 175, [39](#)
- Fischer, T., Guo, G., Dzhioev, A. A., et al. 2020, Phys. Rev. C, 101, [025804](#)
- Fitzpatrick, E. L. 1999, PASP, 755, [63](#)
- Fitzpatrick, E. L., Massa, D., Gordon, K. D., Bohlin, R., & Clayton, G. C. 2019, ApJ, 886, [108](#)
- Foucart, F., O'Connor, E., Roberts, L., et al. 2016, Phys. Rev. D, 94, [123016](#)
- Fouesneau, M., Andrae, R., Sordo, R., & Dharmawardena, T. 2022, [dustapprox](#)
- Fouesneau, M., Frémat, Y., Andrae, R., et al. 2023, A&A, 674, [A28](#)
- Frebel, A. & Norris, J. 2013, in Planets, Stars and Stellar Systems. Volume 5: Galactic Structure and Stellar Populations, ed. T. D. Oswalt & G. Gilmore, Vol. 5, [55](#)
- Frebel, A. & Norris, J. 2015, ARA&A, 53, [631](#)
- Freeman, K. & Bland-Hawthorn, J. 2002, ARA&A, 40, [487](#)
- Fujibayashi, S., Kiuchi, K., Wanajo, S., et al. 2023, ApJ, 942, [39](#)
- Fujibayashi, S., Wanajo, S., Kiuchi, K., et al. 2020, ApJ, 901, [122](#)
- Gaia Collaboration, Babusiaux, C., van Leeuwen, F., et al. 2018a, A&A, 616, [A10](#)
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018b, A&A, 616, [A1](#)
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2021, A&A, 649, [A1](#)
- Gaia Collaboration, Montegriffo, P., Bellazzini, M., et al. 2023a, A&A, 674, [A33](#)
- Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, [A1](#)
- Gaia Collaboration, Schultheis, M., Zhao, H., et al. 2023b, A&A, 674, [A40](#)
- Gaia Collaboration, Vallenari, A., Brown, A. G. A., et al. 2023c, A&A, 674, [A1](#)
- Gallart, C., Bernard, E. J., Brook, C. B., et al. 2019, Natur Astronomy, 3, [932](#)
- Geurts, P., Ernst, D., & Wehenkel, L. 2006, Mach. Learn., 63, [3](#)
- Giacomazzo, B., Rezzolla, L., & Baiotti, L. 2009, MNRASL, 399, [L164](#)
- Giacomazzo, B., Zrake, J., Duffell, P. C., MacFadyen, A. I., & Perna, R. 2015, ApJ, 809, [39](#)
- Gilmore, G. & Reid, N. 1983, MNRAS, 202, [1025](#)
- Gomez, F. A., Helmi, A., Cooper, A. P., et al. 2013, MNRAS, 436, [3602](#)
- Gomez, F. A., Minchev, I., Villalobos, A., O'Shea, B. W., & Williams, M. E. K. 2012, MNRAS, 419, [2163](#)
- GRAVITY Collaboration, Abuter, R., Amorim, A., et al. 2018, A&A, 615, [L15](#)
- Green, G. M. 2018, JOSS, 3, [695](#)
- Green, G. M., Schlafly, E., Zucker, C., Speagle, J. S., & Finkbeiner, D. 2019, ApJ, 887, [93](#)
- Gruen, D., McCullough, J., Amon, A., et al. 2023, The Messenger, 190, [28](#)
- Guedes, J., Callegari, S., Madau, P., & Mayer, L. 2011, ApJ, 742, [76](#)
- Guiglion, G., Battistini, C., Bell, C. P. M., et al. 2019, The Messenger, 175, [17](#)
- Guilet, J., Bauswein, A., Just, O., & Janka, H.-T. 2017, MNRAS, 471, [1879](#)
- Gustafsson, B., Edvardsson, B., Eriksson, K., et al. 2008, A&A, 486, [951](#)

- Haines, C., Jaffé, Y., Tejos, N., et al. 2023, The Messenger, 190, 31
- Hayden, M. R., Bovy, J., Holtzman, J. A., et al. 2015, ApJ, 808, 132
- Haynes, D. M., Kelz, A., Barden, S. C., et al. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI, ed. C. J. Evans, L. Simard, & H. Takami, 99088I
- Haywood, M., Di Matteo, P., Snaith, O., & Lehnert, M. 2015, A&A, 579, A5
- Heil, M., Käppeler, F., Uberseder, E., et al. 2008, Phys. Rev. C, 78, 025802
- Helmi, A. 2020, ARA&A, 58, 205
- Helmi, A., Babusiaux, C., Koppelman, H., et al. 2018, Nature, 563, 85
- Helmi, A., Irwin, M., Deason, A., et al. 2019, The Messenger, 175, 23
- Helmi, A., Veljanoski, J., Breddels, M. A., Tian, H., & Sales, L. 2017, A&A, 598, A58
- Helmi, A., White, S. D. M., de Zeeuw, P. T., & Zhao, H. 1999, Nature, 402, 53
- Helmi, A., White, S. D. M., & Springel, V. 2003, MNRAS, 339, 834
- Hernitschek, N., Cohen, J. G., Rix, H.-W., et al. 2018, ApJ, 859, 31
- Hernquist, L. 1990, ApJ, 356, 359
- Hix, W. R. & Thielemann, F.-K. 1999, J. Comput. Appl. Math., 109, 321
- Holl, B. 2012, Characterization and analysis of the astrometric errors in the global astrometric solution for Gaia (Department of Astronomy and Theoretical Physics, Lund University), 6
- Honda, S., Aoki, W., Ishimaru, Y., Wanajo, S., & Ryan, S. G. 2006, ApJ, 643, 1180
- Hopkins, P. F., Kereš, D., Oñorbe, J., et al. 2014, MNRAS, 445, 581
- Ibata, R., Battaglia, G., Bellazzini, M., et al. 2023, The Messenger, 190, 10
- Iovino, A., Mercurio, A., Gallazzi, A. R., et al. 2023, The Messenger, 190, 22
- Johnson, D. R. H. & Soderblom, D. R. 1987, AJ, 93, 864
- Jordi, C. 2018, GAIA-C5-TN-UB-CJ-041
- Jordi, C., Gebran, M., Carrasco, J. M., et al. 2010, A&A, 523, A48
- Juric, M., Ivezić, Z., Brooks, A., et al. 2008, ApJ, 673, 864
- Just, O., Bauswein, A., Ardevol Pulpillo, R., Goriely, S., & Janka, H. T. 2015, MNRAS, 448, 541
- Kawata, D. & Chiappini, C. 2016, Astronomische Nachrichten, 337, 976
- Klessen, R. & Glover, S. 2023, ARA&A, 61, 65
- Koppelman, H., Helmi, A., & Veljanoski, J. 2018, ApJL, 860, L11
- Koppelman, H. H., Helmi, A., Massari, D., Price-Whelan, A. M., & Starkenburg, T. K. 2019a, A&A, 631, L9
- Koppelman, H. H., Helmi, A., Massari, D., Roelenga, S., & Bastian, U. 2019b, A&A, 625, A5
- Krogager, J. K., Leighly, K. M., Fynbo, J. P. U., et al. 2023, The Messenger, 190, 38
- Kunder, A., Kordopatis, G., Steinmetz, M., et al. 2017, AJ, 153, 75
- Käppeler, F., Gallino, R., Bisterzo, S., & Aoki, W. 2011, Review of Modern Physics, 83, 157
- Laporte, C. F. P., Minchev, I., Johnston, K. V., & Gomez, F. A. 2019, MNRAS, 485, 3134
- Laurent, F., Kosmalski, J., Boudon, D., et al. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI, ed. C. J. Evans, L. Simard, & H. Takami, 99087V
- Lee, Y. S., Beers, T. C., An, D., et al. 2011, ApJ, 738, 187
- Lee, Y. S., Beers, T. C., Sivarani, T., et al. 2008, AJ, 136, 2022
- Lequeux, J., Peimbert, M., Rayo, J. F., Serrano, A., & Torres-Peimbert, S. 1979, A&A, 80, 155
- Li, H., Aoki, W., Matsuno, T., et al. 2022, ApJ, 931, 147
- Licquia, T. C. & Newman, J. A. 2015, ApJ, 806, 96
- Lindgren, L. 2010, in Relativity in Fundamental Astronomy: Dynamics, Reference Frames, and Data Analysis, ed. S. A. Klioner, P. K. Seidelmann, & M. H. Soffel, Vol. 261, 296
- Lindgren, L. & Bastian, U. 2010, in EAS Publications Series, Vol. 45, EAS Publications Series, 109
- Lindgren, L., Bastian, U., Biermann, M., et al. 2021a, A&A, 649, A4
- Lindgren, L., Klioner, S. A., Hernández, J., et al. 2021b, A&A, 649, A2
- Lodders, K., Palme, H., & Gail, H. 2009, Landolt Börnstein, 4B, 712
- Lucatello, S., Bragaglia, A., Vallenari, A., et al. 2023, The Messenger, 190, 13

- Lucey, M., Al Kharusi, N., Hawkins, K., et al. 2023, MNRAS, 523, 4049
- Luri, X., Palmer, M., Arenou, F., et al. 2014, A&A, 566, A119
- Lövdal, S. S., Ruiz-Lara, T., Koppelman, H. H., et al. 2022a, A&A, 665, A57
- Lövdal, S. S., Ruiz-Lara, T., Koppelman, H. H., et al. 2022b, VizieR Online Data Catalog, J/A+A/665/A57
- Magrin, D., Ragazzoni, R., Rauer, H., et al. 2018, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10698, Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave, ed. M. Lystrup, H. A. MacEwen, G. G. Fazio, N. Batalha, N. Siegler, & E. C. Tong, 106984X
- Maiolino, R., Nagao, T., Grazian, A., et al. 2008, A&A, 488, 463
- Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, AJ, 154, 94
- Malhan, K. & Ibata, R. A. 2018, MNRAS, 477, 4063
- Marigo, P., Girardi, L., Bressan, A., et al. 2017, ApJ, 835, 77
- Marocco, F., Eisenhardt, P. R. M., Fowler, J. W., et al. 2021, ApJS, 253, 8
- Martin, D., Perego, A., Arcones, A., et al. 2015, ApJ, 813, 2
- Martin, D., Perego, A., Kastaun, W., & Arcones, A. 2018, Classical Quantum Gravity, 35, 034001
- Martin, N. F., Starkenburg, E., Yuan, Z., et al. 2023, arXiv:2308.01344
- Martínez-Pinedo, G., Fischer, T., & Huther, L. 2014, J. Phys. G, 41, 044008
- Mascali, D., Santonocito, D., Amaducci, S., et al. 2022, Universe, 8, 80
- Mashonkina, L. & Gehren, T. 2001, A&A, 376, 232
- Matsuno, T., Hirai, Y., Tarumi, Y., et al. 2021, A&A, 650, A110
- McMillan, P. 2017, MNRAS, 465, 76
- McMillan, P. J. & Binney, J. J. 2008, MNRAS, 390, 429
- McWilliam, A. 1997, ARA&A, 35, 503
- Merloni, A., Alexander, D. A., Banerji, M., et al. 2019, The Messenger, 175, 42
- Merloni, A., Predehl, P., Becker, W., et al. 2012, arXiv:1209.3114
- Minchev, I., Famaey, B., Quillen, A. C., et al. 2012, A&A, 548, A127
- Miyamoto, M. & Nagai, R. 1975, PASJ, 27, 533
- Myeong, G. C., Evans, N. W., Belokurov, V., Amorisco, N. C., & Koposov, S. E. 2018a, MNRAS, 475, 1537
- Myeong, G. C., Evans, N. W., Belokurov, V., Sanders, J. L., & Koposov, S. E. 2018b, ApJL, 863, L28
- Myeong, G. C., Vasiliev, E., Iorio, G., Evans, N. W., & Belokurov, V. 2019, MNRAS, 488, 1235
- Naidu, R. P., Conroy, C., Bonaca, A., et al. 2020, ApJ, 901, 48
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563
- Nelson, D., Springel, V., Pillepich, A., et al. 2019, Comput. Astrophys. Cosmol., 6, 2
- Nissen, P. & Schuster, W. 2010, A&A, 511, L10
- n\_TOF Collaboration, Abbondanno, U., Aerts, G., et al. 2003, Document
- Papenfort, L. J., Gold, R., & Rezzolla, L. 2018, Phys. Rev. D, 98, 104028
- Pawlak, M., Mazeh, T., Faigler, S., et al. 2023, The Messenger, 190, 17
- Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825
- Peroux, C., Merloni, A., Liske, J., et al. 2023, The Messenger, 190, 42
- Perryman, M. & ESA. 1997, in ESA Special Publication, Vol. 1200, ESA Special Publication
- Perryman, M. A. C., de Boer, K. S., Gilmore, G., et al. 2001, A&A, 369, 339
- Pila-Díez, B., de Jong, J. T. A., Kuijken, K., van der Burg, R. F. J., & Hoekstra, H. 2015, A&A, 579, A38
- Piskunov, N. & Valenti, J. A. 2017, A&A, 597, A16
- Plez, B. 2012, Turbospectrum: Code for spectral synthesis, Astrophysics Source Code Library, record ascl:1205.004
- Prantzos, N., Abia, C., Cristallo, S., Limongi, M., & Chieffi, A. 2020, MNRAS, 491, 1832
- Price-Whelan, A., Sipőcz, B., Lenz, D., et al. 2020, adrn/gala: v1.3
- Price-Whelan, A. M. 2017, JOSS, 2, 388
- Quinn, P., Hernquist, L., & Fullagar, D. 1993, ApJ, 403, 74
- Recio-Blanco, A., de Laverny, P., Allende Prieto, C., et al. 2016, A&A, 585, A93
- Recio-Blanco, A., de Laverny, P., Palicio, P. A., et al. 2023, A&A, 674, A29
- Richard, J., Kneib, J. P., Blake, C., et al. 2019, The Messenger, 175, 50

- Riello, M., De Angeli, F., Evans, D. W., et al. 2021, *A&A*, 649, [A3](#)
- Rix, H.-W., Chandra, V., Andrae, R., et al. 2022, *ApJ*, 941, [45](#)
- Robin, A. C., Bienaymé, O., Fernandez-Trincado, J. G., & Reylé, C. 2017, *A&A*, 605, [A1](#)
- Robin, A. C., Luri, X., Reylé, C., et al. 2012a, *A&A*, 543, [A100](#)
- Robin, A. C., Marshall, D. J., Schultheis, M., & Reylé, C. 2012b, *A&A*, 538, [A106](#)
- Robin, A. C., Reylé, C., Derrière, S., & Picaud, S. 2003, *A&A*, 409, [523](#)
- Robin, A. C., Reylé, C., Fliri, J., et al. 2014, *A&A*, 569, [A13](#)
- Roederer, I. U., Hattori, K., & Valluri, M. 2018, *AJ*, 156, [179](#)
- Ruiz-Lara, T., Matsuno, T., Lövdal, S. S., et al. 2022, *A&A*, 665, [A58](#)
- Rybicki, J., Demleitner, M., Fouesneau, M., et al. 2018, *PASP*, 130, [074101](#)
- Sacco, G. G., Jeffries, R., Binks, A., et al. 2023, *The Messenger*, 190, [7](#)
- Sakari, C. M., Placco, V. M., Farrell, E. M., et al. 2018, *ApJ*, 868, [110](#)
- Sanders, J. L. & Das, P. 2018, *MNRAS*, 481, [4093](#)
- Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, *MNRAS*, 446, [521](#)
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *ApJ*, 500, [525](#)
- Schneider, P. 2006, *Extragalactic Astronomy and Cosmology*, 1st edn. (Berlin Heidelberg New York: Springer), p. 47ff.
- Schönrich, R., Binney, J., & Dehnen, W. 2010, *MNRAS*, 403, [1829](#)
- Seifert, W., Xu, W., Buschkamp, P., et al. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI, ed. C. J. Evans, L. Simard, & H. Takami, [990890](#)
- Sharpee, B., Zhang, Y., Williams, R., et al. 2007, *ApJ*, 659, [1265](#)
- Sheinis, A., Anguiano, B., Asplund, M., et al. 2015, *J. Astron. Telesc. Instrum. Syst.*, 1, [035002](#)
- Sheinis, A., Saunders, W., Gillingham, P., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9151, Advances in Optical and Mechanical Technologies for Telescopes and Instrumentation, ed. R. Navarro, C. R. Cunningham, & A. A. Barto, [91511X](#)
- Shen, J., Rich, R. M., Kormendy, J., et al. 2010, *ApJL*, 720, [L72](#)
- Shen, S., Cooke, R. J., Ramirez-Ruiz, E., et al. 2015, *ApJ*, 807, [115](#)
- Simmerer, J., Sneden, C., Cowan, J. J., et al. 2004, *ApJ*, 617, [1091](#)
- Skuladottir, A., Hansen, C. J., Salvadori, S., & Choplén, A. 2019, *A&A*, 631, [A171](#)
- Skuladottir, A., Puls, A. A., Amarsi, A. M., et al. 2023, *The Messenger*, 190, [19](#)
- Smee, S. A., Gunn, J. E., Uomoto, A., et al. 2013, *AJ*, 146, [32](#)
- Sneden, C., Cowan, J., & Gallino, R. 2008, *ARA&A*, 46, [241](#)
- Sneden, C., Cowan, J. J., Ivans, I. I., et al. 2000, *ApJ*, 533, [L139](#)
- Sneden, C., Cowan, J. J., Lawler, J. E., et al. 2003, *ApJ*, 591, [936](#)
- Soderblom, D. R. 2010, *ARA&A*, 48, [581](#)
- Sotillo-Ramos, D., Bergemann, M., Friske, J. K. S., & Pilipich, A. 2023, *MNRASL*, 525, [L105](#)
- Sotillo-Ramos, D., Pilipich, A., Donnari, M., et al. 2022, *MNRAS*, 516, [5404](#)
- Soubiran, C., Le Campion, J.-F., Brouillet, N., & Chemin, L. 2016, *A&A*, 591, [A118](#)
- Starkenburg, E., Martin, N., Youakim, K., et al. 2017, *MNRAS*, 471, [2587](#)
- Steinmetz, M., Guiglion, G., McMillan, P. J., et al. 2020a, *AJ*, 160, [83](#)
- Steinmetz, M., Matijević, G., Enke, H., et al. 2020b, *AJ*, 160, [82](#)
- Sterling, N. & Dinerstein, H. L. 2008, *ApJS*, 174, [158](#)
- Suda, T., Hidaka, J., Aoki, W., et al. 2017, *Publications of the Astronomical Society of Japan*, 69, [76](#)
- Suda, T., Katsuta, Y., Yamada, S., et al. 2008, *Publications of the Astronomical Society of Japan*, 60, [1159](#)
- Sutherland, W., Emerson, J., Dalton, G., et al. 2015, *A&A*, 575, [A25](#)
- Swann, E., Sullivan, M., Carrick, J., et al. 2019, *The Messenger*, 175, [58](#)
- Takahashi, K. & Yokoi, K. 1987, *Atomic Data and Nuclear Data Tables*, 36, [375](#)
- Taylor, E. N., Cluver, M., Bell, E., et al. 2023, *The Messenger*, 190, [46](#)
- Taylor, M. 2005, in *Astronomical Society of the Pacific Conference Series*, Vol. 347, *Astronom-*

- ical Data Analysis Software and Systems XIV, ed. P. Shopbell, M. Britton, & R. Ebert, [29](#)
- Thielemann, F.-K. 2023, Eur. Phys. J. A, [59](#), [12](#)
- Thielemann, F.-K., Eichler, M., Panov, I. V., & Wehmeyer, B. 2017, Annu. Rev. Nucl. Part. Sci., [67](#), [253](#)
- Tissera, P. B., Scannapieco, C., Beers, T. C., & Carollo, D. 2013, MNRAS, [432](#), [3391](#)
- Toloza, O., Rebassa-Mansergas, A., Raddi, R., et al. 2023, The Messenger, [190](#), [4](#)
- Tolstoy, E., Hill, V., & Tosi, M. 2009, ARA&A, [47](#), [371](#)
- Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, [613](#), [898](#)
- Truran, J. W., Cowan, J. J., Pilachowski, C. A., & Sneden, C. 2002, PASP, [114](#), [1293](#)
- van de Voort, F., Quataert, E., Hopkins, P. F., Kereš, D., & Faucher-Giguère, C.-A. 2015, MNRAS, [447](#), [140](#)
- Wanajo, S. & Ishimaru, Y. 2006, Nucl. Phys. A, [777](#), [676](#)
- Wanajo, S., Sekiguchi, Y., Nishimura, N., et al. 2014, ApJL, [789](#), [L39](#)
- Watson, D., Hansen, C. J., Selsing, J., et al. 2019, Nature, [574](#), [497](#)
- Wilson, J. C., Hearty, F. R., Skrutskie, M. F., et al. 2019, PASP, [131](#), [055001](#)
- Winkler, R., Barden, S. C., & Saviauk, A. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI, ed. C. J. Evans, L. Simard, & H. Takami, [990891](#)
- Winkler, R., Barden, S. C., Saviauk, A., et al. 2018, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10702, Ground-based and Airborne Instrumentation for Astronomy VII, ed. C. J. Evans, L. Simard, & H. Takami, [1070286](#)
- Wu, J., Nishimura, S., Lorusso, G., et al. 2017, Phys. Rev. Lett., [118](#), [072701](#)
- Xue, X.-X., Rix, H.-W., Ma, Z., et al. 2015, ApJ, [809](#), [144](#)
- Yanny, B., Rockosi, C., Newberg, H. J., et al. 2009, AJ, [137](#), [4377](#)
- York, D. G., Adelman, J., Anderson, John E., J., et al. 2000, AJ, [120](#), [1579](#)
- Youakim, K., Starkenburg, E., Martin, N. F., et al. 2020, MNRAS, [492](#), [4986](#)
- Yuan, Z., Chang, J., Beers, T. C., & Huang, Y. 2020, ApJL, [898](#), [L37](#)
- Zhao, G., Zhao, Y., Chu, Y., Jing, Y., & Deng, L. 2012, Res. Astron. Astrophys., [12](#), [723](#)
- Zhu, L., van de Ven, G., Leaman, R., et al. 2022, A&A, [664](#), [A115](#)
- Zolotov, A., Willman, B., Brooks, A. M., et al. 2009, ApJ, [702](#), [1058](#)
- Zucker, S. 2003, MNRAS, [342](#), [1291](#)

## A Appendix

### A.1 Reproduction of previous data

In order to check whether the methods used in this work to derive the kinematic properties are valid, the results of two works from the literature (Helmi et al. 2018; Lövdal et al. 2022a) are recreated using the approach as described in Section 5.2.

#### Helmi et al. (2018)

Helmi et al. (2018) use the Gaia DR2 6D dataset in order to showcase the presence of a slightly retrograde kinematic structure in the halo, identified as a remnant of the Galaxy's last major merger and named Gaia-Enceladus. Selection criteria for objects considered are that they have 6D information in Gaia DR2, a small relative parallax error  $\frac{\pi}{\sigma_\pi} > 5$ , and  $\pi > 0.2$  mas, while for some figures only even more closeby stars  $\pi > 0.4$  mas are included. Velocities that were derived from the 6D data are corrected for the Local Standard of Rest velocity  $v_{\text{LSR}} = 232$  km/s, as well as the peculiar motion of the Sun (see Schönrich et al. 2010). In order to discard as many disk stars as possible, a condition for halo stars is chosen as  $|v - v_{\text{LSR}}| > 210$  km/s. For nearby stars, the selection criteria for Gaia-Enceladus are  $-1500 < L_z < 150$  kpc·km/s and  $E_{\text{orb}} > -1.8 \cdot 10^5$  (km/s)<sup>2</sup>, but for stars farther away ( $0.2 \text{ mas} > \pi > 0.1 \text{ mas}$ ), the energy criterium is dropped as the Galaxy's gravitational potential is less well known there. Of course, a total separation of stars belonging to the substructure from other Milky Way stars is not possible from kinematics alone, in particular in the inner regions of the Galaxy where disk stars exhibit a higher velocity dispersion. For the Solar scaling used in some figures, the values are  $E_\odot = -1.63 \cdot 10^5$  (km/s)<sup>2</sup>,  $L_{z,\odot} = 1902.4$  kpc·km/s, and  $R_\odot = 8.2$  kpc.

Thus, in order to determine all parameters to recreate this data, the exact same approach as was described in Section 5.2 is applied to the Gaia DR2 6D subset with  $\frac{\pi}{\sigma_\pi} > 5$ , deriving velocities, distances and specific angular momenta. Only the choice of potential is different, as the authors use the potential described in Helmi et al. (2017), which includes a logarithmic halo, a Miyamoto-Nagai disk, and a Hernquist bulge,

$$\Phi_{\text{halo}} = v_{\text{halo}}^2 \ln(1 + R^2/d^2 + z^2/d^2), \quad (16)$$

$$\Phi_{\text{disk}} = -\frac{GM_{\text{disk}}}{\sqrt{R^2 + (a_d + \sqrt{z^2 + b_d^2})^2}}, \quad (17)$$

$$\Phi_{\text{bulge}} = -\frac{GM_{\text{bulge}}}{r + c_b}, \quad (18)$$

where  $v_{\text{halo}} = 173.2$  km/s and  $d = 12$  kpc for the halo,  $M_{\text{disk}} = 6.3 \cdot 10^{10} M_\odot$ ,  $a_d = 6.5$  kpc and  $b_d = 0.26$  kpc for the disk, and  $M_{\text{bulge}} = 2.1 \cdot 10^{10} M_\odot$  and  $c_b = 0.7$  kpc for the bulge. The orbital energy of the Sun in this potential is indeed found to be  $E_\odot = -1.63 \cdot 10^5$  (km/s)<sup>2</sup> after correcting for the zero-point offset, as mentioned in Helmi et al. (2018). Thus, some of the central figures showcasing Gaia-Enceladus in their work can be recreated. First,  $\sqrt{v_x^2 + v_z^2}$  vs.  $v_y$  is plotted for stars at  $\pi > 0.4$  mas and compared to their corresponding figure, as shown in Figure 62. The agreement between the two is excellent, with all important features being basically identical, such as the arc at  $v_y \sim -400$  km/s and the blob being clearly set apart from the thick disk near the selection cutoff. It should be noted that the range to select the stars colour-coded in blue that was given by Helmi et al. (2018) to be  $-1500 < L_z < 150$  kpc·km/s in the angular momentum would in fact preclude the objects at very high negative velocities, including the arc. Only by removing the lower boundary of that range could the figure be completely reproduced as shown in Figure 62.

In addition to this figure, the two left panels of their Extended Data Figure 1 is also reproduced, as depicted in Figure 63. This figure includes all stars up to a distance corresponding to  $\pi > 0.2$  mas, and

the additional constraint `phot_bp_rp_excess_factor < 1.27` is put on the data. Once again, the agreement is excellent, although the different colour coding alters the appearance of the left panel somewhat. Furthermore, the energy axis appears to be on a different scale after being transformed to Solar values, but this is just a matter of presentation. Figure 62 and Figure 63 suggest that the methods used to derive all relevant kinematic parameters in this work ( $v$ ,  $L_z$ ,  $E_{\text{orb}}$ ,  $R$ ) provide reliable data.

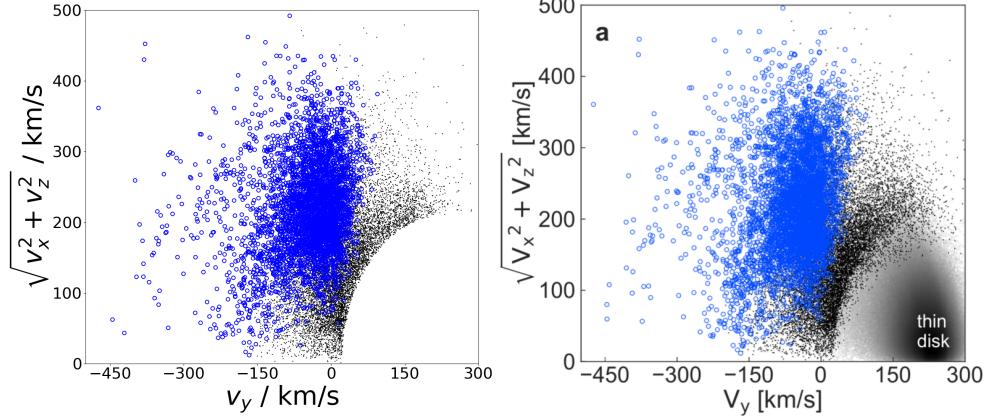


Figure 62: The recreated  $\sqrt{v_x^2 + v_z^2}$  vs.  $v_y$  plot (left panel) in comparison to the original Figure 1 (right panel) from Helmi et al. (2018). The blurred out thin disk stars were left out of the reproduced figure.

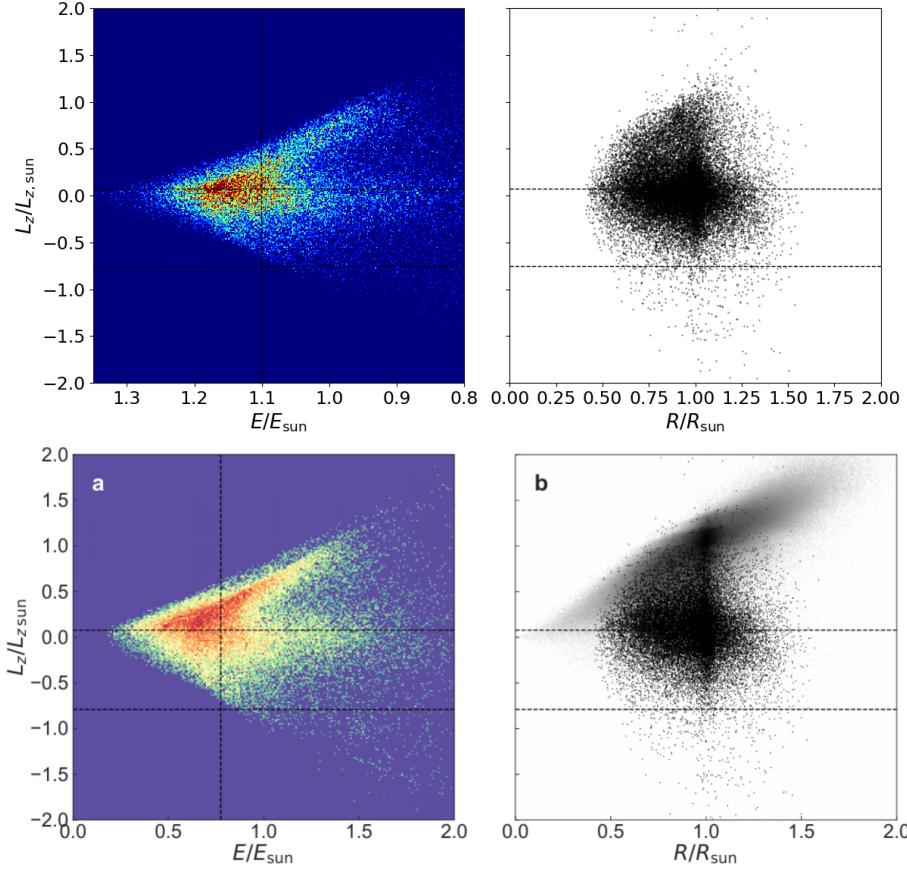


Figure 63: The recreated  $L_z$  vs.  $E_{\text{orb}}$  plot (upper panels) in comparison to the original Extended Data Figure 1 (lower panels) from Helmi et al. (2018) scaled by the Solar values. The blurred out thin disk stars in the bottom right panel were left out of the reproduced figure.

## Lövdal et al. (2022a) and Ruiz-Lara et al. (2022)

Lövdal et al. (2022a) apply a hierarchical clustering algorithm to local halo stars ( $d < 2.5$  kpc) in the 3D space defined by  $E_{\text{orb}}$ ,  $L_z$ , and  $L_{\perp}$  in order to obtain data-driven and statistics-based results on the substructure of the halo. They make use of data from Gaia EDR3 (meaning that their radial velocities are still from Gaia DR2) and complement this data with further radial velocities from different spectroscopic surveys. A detailed analysis and interpretation of the structures found is provided by Ruiz-Lara et al. (2022). The kinematic parameters as well as the cluster positions are provided by the authors (Lövdal et al. 2022b)<sup>20</sup>, facilitating a comparison of the parameters derived according to the methods in this work, and their parameters. For the set used by Lövdal et al. (2022a), data is retrieved from Gaia DR3 and the parameters are calculated as in Section 5.2, but using the alternative potential derived via GALA.

Figure 64 shows the comparison of the velocity distribution between this work and Lövdal et al. (2022a), and the comparison of the  $E_{\text{orb}}-L_z$ -distribution, whereas Figure 65 features a direct comparison of  $L_z$  and  $E_{\text{orb}}$ . Barring some outliers and a small accumulation of stars with disk-like orbits, the agreement is excellent. The arc at  $v_y \sim -400$  is clearly present in both figures in the upper panel of Figure 64, as is the Gaia-Enceladus blob in separation from the thick disk. In the lower panels of Figure 64, the slight overdensities marking Sequoia are visible, albeit less populated than in their data. Finally, the direct comparison of  $L_z$  and  $E_{\text{orb}}$  to the values of Lövdal et al. (2022a) in Figure 65 shows that the bulk of the stars is basically identical in  $L_z$  (indicated by the colour-coded high density along the identity line), and very similar in  $E_{\text{orb}}$ . Some stars deviate slightly, but only a very small number of stars are extreme outliers, likely those that were seen in the other two plots to be scattered across the whole space, but with a bias for disk-like kinematics, as the disk is the dominant in the Solar neighbourhood.

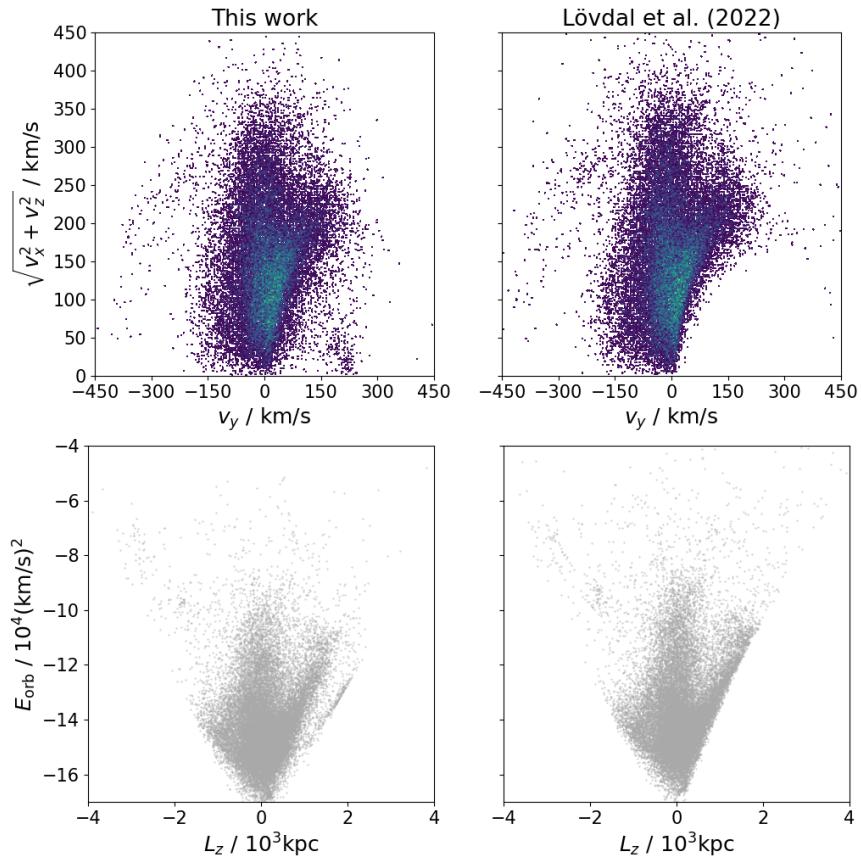


Figure 64:  $\sqrt{v_x^2 + v_z^2}$  vs.  $v_y$  (upper panels) and  $E_{\text{orb}}$  vs.  $L_z$  (lower panels) plots for parameters calculated in this work (left panels) and by Lövdal et al. (2022a) (right panels).

<sup>20</sup><https://vizier.cds.unistra.fr/viz-bin/VizieR?-source=J/A+A/665/A57>, retrieved 22.10.2023

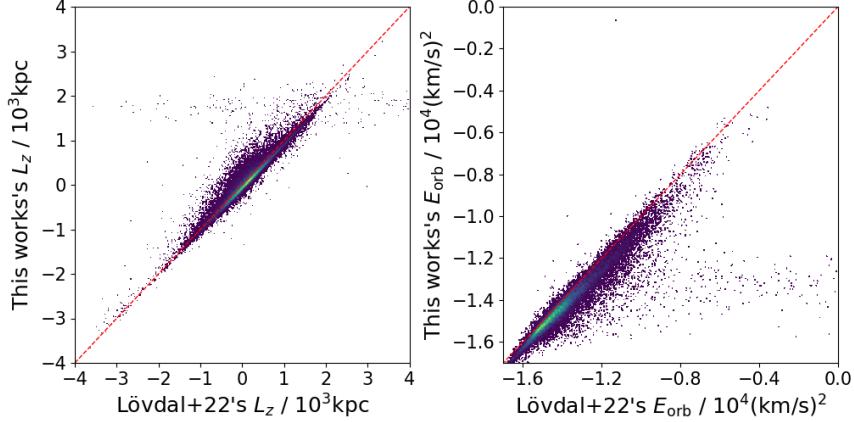


Figure 65: The direct comparison between the kinematic parameters calculated in this work on the  $y$ -axis and the parameters calculated by Lövdal et al. (2022a) on the  $x$ -axis for  $L_z$  (left) and  $E_{\text{orb}}$  (right).

It should be noted that in the reproduction of the data, Gaia DR3 radial velocities are used, whereas the original data used a combination of Gaia DR2 data and radial velocities from other sources. Thus, besides minor differences expected between the DR2 and DR3 radial velocities, it is expected that some of the stars have Gaia DR3 radial velocities in the left panel, but some other radial velocity in the right panel, which may in some cases deviate from each other significantly. This is most likely the explanation for the slightly increased scatter and the small number of outliers visible in the figures, which are also the reason why there are objects inside the velocity cut and at disk-like kinematics. However, the conclusion of this section is the same as for the Helmi et al. (2018) data and validates the derivations of kinematic parameters in Section 5.2.

## A.2 On the GSP-Spec abundances

The cross-identification of the 6D subset from Section 5 with Gaia DR3’s additional astrophysical parameters allows for the use of chemical abundance data derived from the RVS spectra via GSP-Spec (see Section 3.1). However, the number of available abundance measurements varies widely, and only the [M/H] and  $[\alpha/\text{Fe}]$  estimates are available for a significant number of stars, specifically just under 10%. Particularly the  $\alpha$ -abundances would be attractive, as mostly reliable [Fe/H]-estimates are already given for all stars by the Pristine-Gaia-Synthetic catalogue, whereas the  $\alpha$ -abundances contained in the cross-match with GALAH DR3 are less numerous. However, a simple verification of the [M/H] and  $[\alpha/\text{Fe}]$  abundances by using GALAH as a reference suggests that the  $\alpha$ -abundances provided by GSP-Spec in Gaia DR3 are at this stage not yet ready to be utilised on a large scale. For this, a crossmatch of the whole GALAH DR3 catalogue to Gaia DR3 is carried out, and basically all GALAH stars are identified in Gaia DR3, although not all have GSP-Spec abundance measurements.

The direct comparison between the abundances is shown in the top panels of Figure 66. As can be seen, [M/H] seems to agree with GALAH data for the most part, albeit with large scatter and many outliers, whereas the  $\alpha$ -abundances look almost completely random without any correlation to the GALAH  $\alpha$ -abundances. To improve the comparison, the GALAH flags are cleaned by asking `f1ag_sp = 0` and `f1ag_fe_h = 0`. For GSP-Spec, the flags provided are strings of length 41, and a completely clean sample is provided if the first 13 integers in the string are set to 0. However, with all 13 integers set to 0, the sample becomes very small, so instead the first 7 flags are cleaned. The resulting comparison is shown in the bottom two panels of Figure 66. The extreme outliers in the [M/H] comparison are now gone, and although the scatter remains rather high, the [M/H] estimates by GSP-Spec appear to provide reasonable metallicity data once the flags are cleaned. For  $[\alpha/\text{Fe}]$ , the comparison looks slightly improved, but is still without much correlation between the two estimates, and most stars are removed by the cleaning of the flags as well. Thus, only the GALAH DR3 elemental abundances are used in this work.

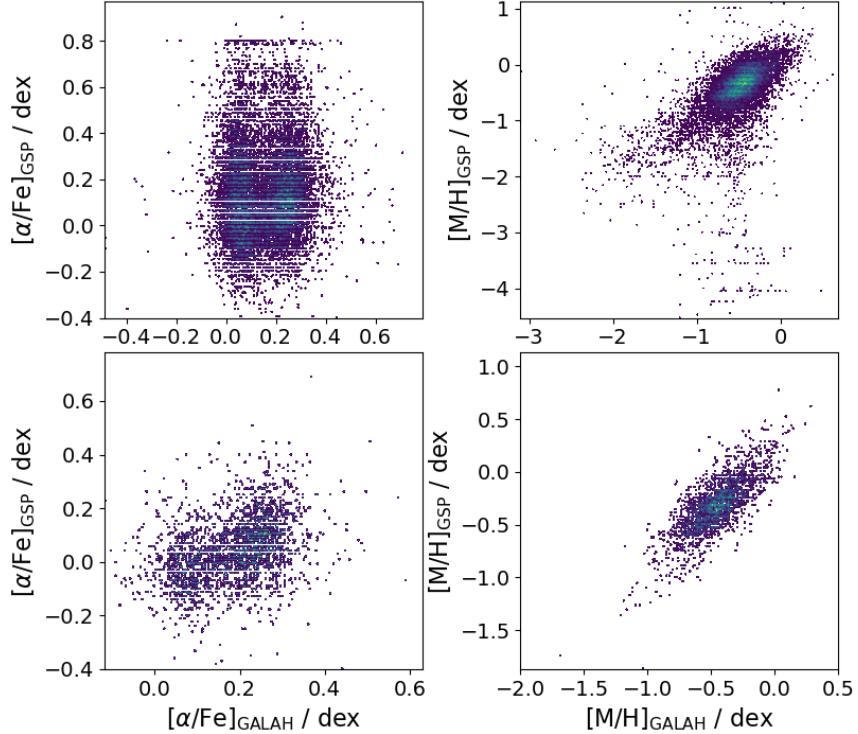


Figure 66: The comparison of  $[M/H]$  and  $[\alpha/Fe]$  between Gaia DR3’s GSP-Spec and GALAH DR3 for all available estimates (upper panels) and after cleaning the flags (lower panels).

### A.3 Extension of Section 5.5 to full GALAH DR3

The procedure described in Section 5.5 is applied to the full GALAH DR3 catalogue in the exact same way as for the crossmatch between GALAH and the 4MOST S2 target catalogue. In this case, there are about 92 300 stars left for the r-process comparison, 114 500 for the s-process comparison, and 98 600 for the weak r-process comparison after the flags are cleaned, the requirement of at least 6 available elements is ensured, and the references are scaled. The application of the selection criteria to these leave 1923 candidate stars for r-process-enrichment, 8757 candidate stars for s-process-enrichment, and 2023 candidate stars for weak-r-process-enrichment. The number for the s-process is remarkably high, and Figure 67 as well as Figure 68 show that the bulk of stars is indeed shifted to below  $\sigma < 2.5$  for the deviation from the s-process, whereas for the r- and weak-r process, this peak is at  $\sigma > 3$ . Clearly, there is some systematic effect that is not yet understood and needs to be investigated in the future. Finally, Figure 69 showing the statistical distribution of the sample in several parameters also indicates that almost all candidates are giants. However, while the s-process candidates contain almost no main sequence stars, the r-process and weak r-process do, and interestingly, the r-process candidates’ main sequence stars are located on the lower main sequence at cooler temperatures, whereas the weak r-process candidates’ main sequence stars are mainly turnoff stars. Also, the r-process candidates occupying the cool half of the giant branch and the weak r-process being more centrally set on the giant branch formed by the full sample is also visible here.

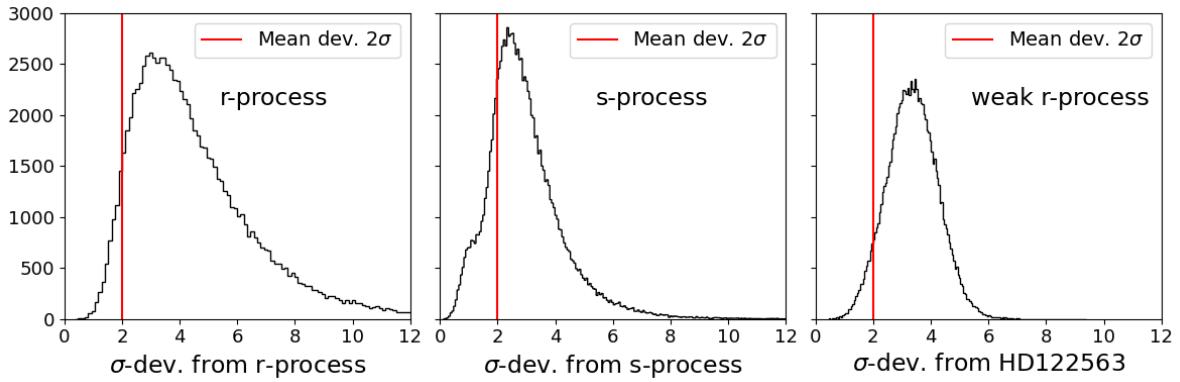


Figure 67: The distribution of the mean  $\sigma$ -deviations between the GALAH pattern and the reference pattern for the full GALAH DR3 catalogue. The red line indicates the  $2\sigma$ -range.

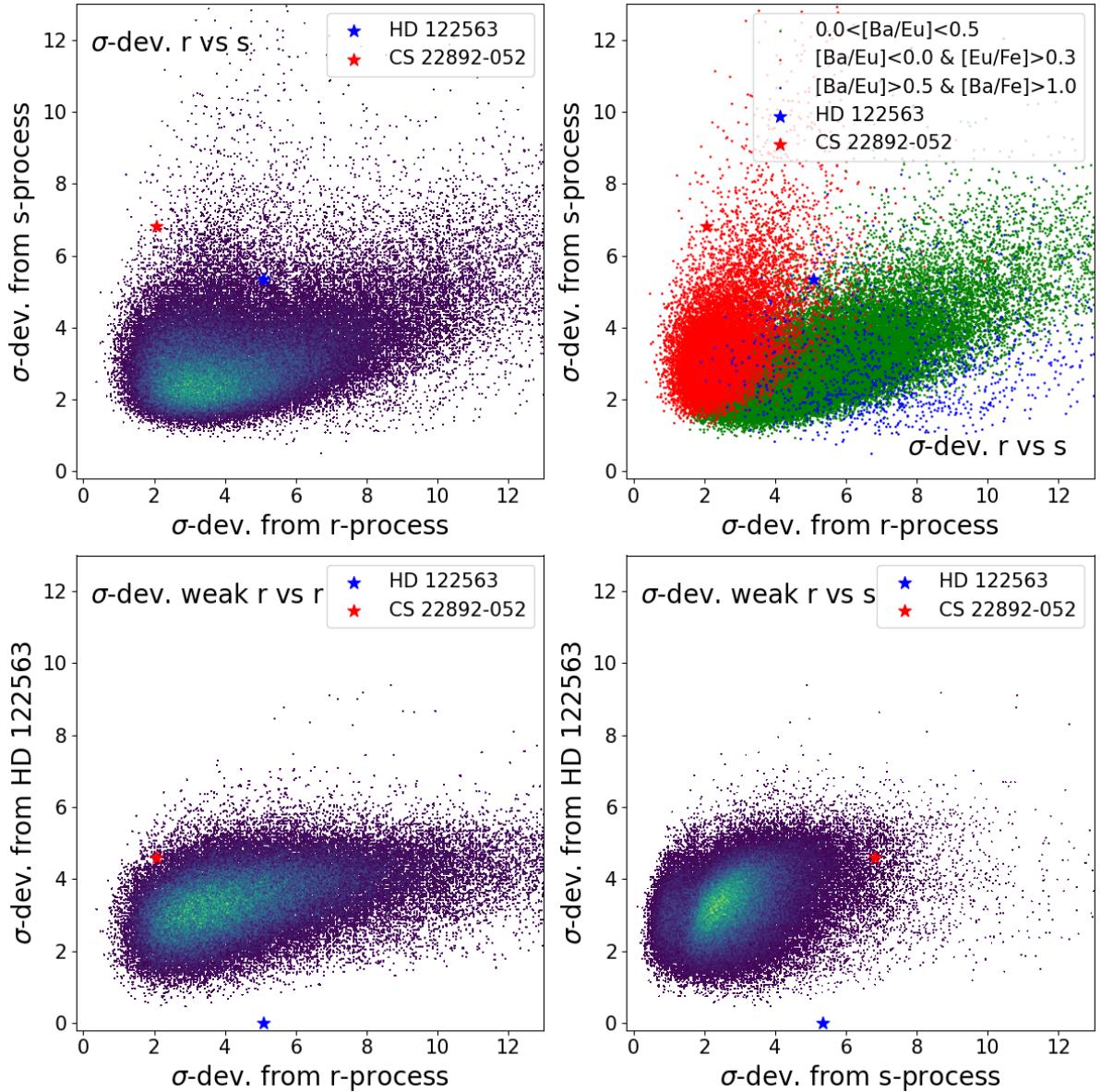


Figure 68: The deviations of the stars' elemental patterns from the three references for the full GALAH DR3 catalogue are shown in two dimensions.

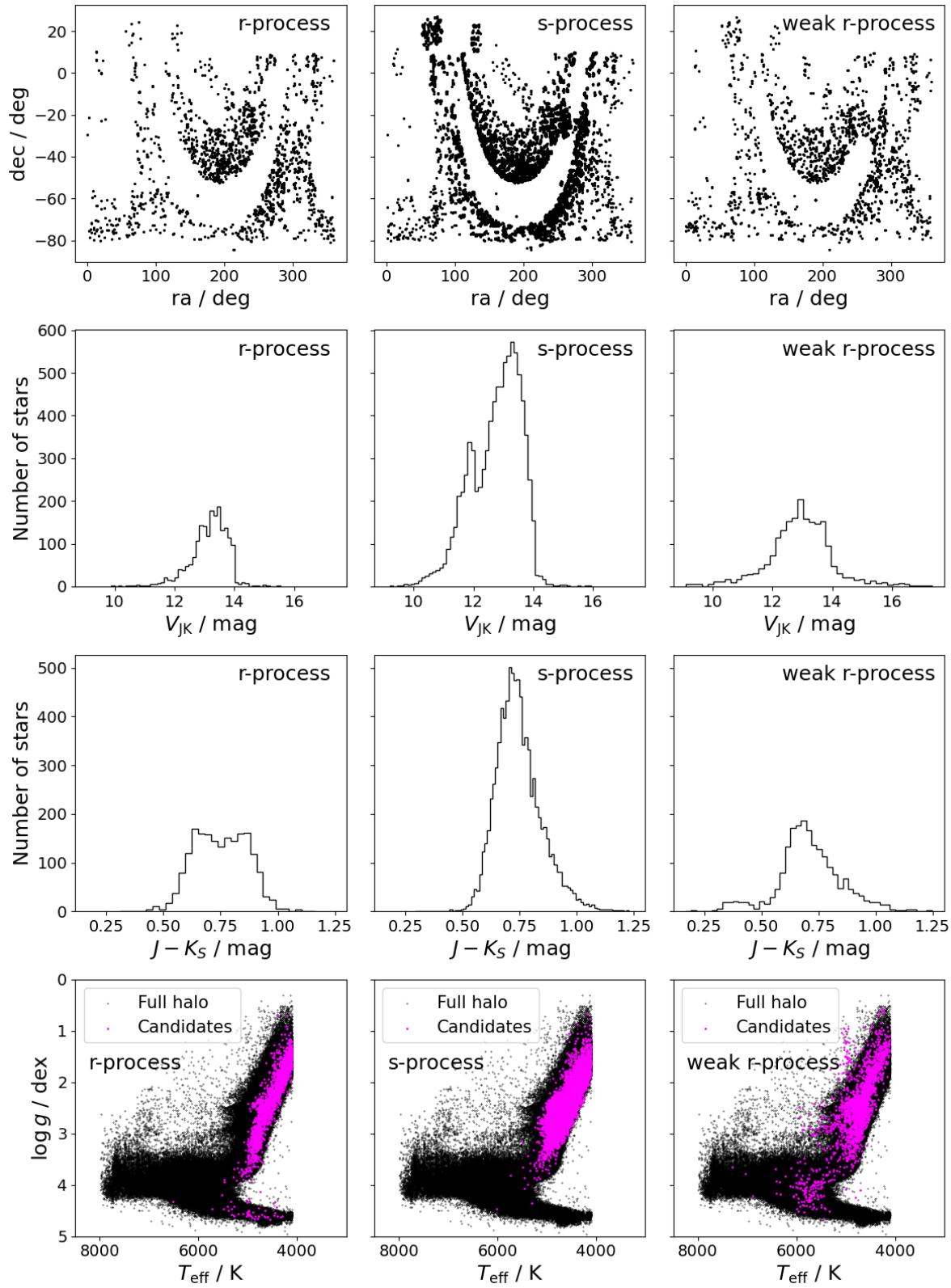


Figure 69: Some statistical properties of the selected candidates for the full GALAH DR3 catalogue. The stellar parameters in the plots in the lower panels were cleaned using `flag_sp = 0`.

## **Erklärung**

Ich versichere, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, den 07.11.2023

*Richter*