Galactic archaeology with metal-poor stars from the *Pristine* survey

Kris Youakim
Leibniz-Institut für Astrophysik Potsdam (AIP)

Dissertation
zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)
in der Wissenschaftsdisziplin Astrophysik

Eingereicht an der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Potsdam

Potsdam, den 11. August 2020

Cover: The Milky Way over Sandcut Beach, Vancouver Island, British Columbia, Canada.
Credit: Nic Annau. Pristine logo designed by: Anke Arentsen.

Betreuer: Dr. Else Starkenburg/Prof. Dr. Matthias Steinmetz

1. Gutachter: Dr. Else Starkenburg
Leibniz-Institut für Astrophysik Potsdam

2. Gutachter: Prof. Dr. Matthias Steinmetz
Leibniz-Institut für Astrophysik Potsdam / Universität Potsdam

3. Gutachter: Prof. Dr. Søren Larsen
Radboud University Nijmegen

Published online on the
Publication Server of the University of Potsdam:
https://doi.org/10.25932/publishup-47431
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-474314

Contents

1 Introduction **1**
- 1.1 Galactic archaeology . . . 2
  - 1.1.1 The metallicity content of stars . . . 3
  - 1.1.2 Stellar spectroscopy . . . 4
  - 1.1.3 The most metal-poor stars . . . 5
  - 1.1.4 Carbon content in metal-poor stars . . . 5
- 1.2 The *Pristine* Survey . . . 7
- 1.3 The Milky Way . . . 9
  - 1.3.1 The Galactic halo . . . 10
  - 1.3.2 Substructure in the halo . . . 11
- 1.4 Thesis summary . . . 13
  - 1.4.1 Spectroscopic confirmation of an efficient search for extremely metal-poor stars . . . 13
  - 1.4.2 The first three years of medium-resolution follow-up spectroscopy of *Pristine* EMP star candidates . . . 14
  - 1.4.3 The metallicity distribution function of the Milky Way halo down to the extremely metal-poor regime . . . 14
  - 1.4.4 A uniquely clean view of the Galactic outer halo using blue horizontal branch stars . . . 15
  - 1.4.5 Quantifying substructure in the Milky Way halo at different metallicities . 16

2 Efficiency of the *Pristine* survey **17**
- 2.1 Introduction . . . 19
- 2.2 The *Pristine* survey . . . 21
- 2.3 Spectroscopic follow-up . . . 23
  - 2.3.1 Data reduction and analysis . . . 23
  - 2.3.2 Spectral analysis . . . 23
- 2.4 Selection criteria . . . 24
  - 2.4.1 SDSS photometry . . . 25
  - 2.4.2 Infrared magnitudes from WISE and 2MASS . . . 27
  - 2.4.3 Applying a regularized regression technique . . . 28
- 2.5 Spectroscopic results . . . 29
- 2.6 Discussion . . . 32
  - 2.6.1 Purity and success rates of the selection . . . 32
  - 2.6.2 Comparison to other surveys . . . 34
  - 2.6.3 Future follow-up strategy . . . 34
- 2.7 Conclusions . . . 35

3 Medium resolution follow-up of *Pristine* candidates **37**
- 3.1 Introduction . . . 39
- 3.2 Data and Observations . . . 41
  - 3.2.1 Observations . . . 41
  - 3.2.2 Observational strategy . . . 43
  - 3.2.3 Data reduction . . . 43

3.3 Analysis with FERRE . . . 45
3.3.1 Stellar parameters . . . 46
3.3.2 Carbon abundance . . . 47
3.4 Results . . . 49
3.4.1 Comparison of the photometric metallicities and spectroscopic metallicities 49
3.4.2 Updated purity and success rates of the *Pristine* survey . . . 51
3.4.3 The carbon-enhancement present in the sample . . . 52
3.4.4 The full sample . . . 53
3.5 Future of the survey . . . 55
3.5.1 Pristine and Gaia . . . 56
3.5.2 Pristine and WEAVE . . . 56
3.6 Conclusions . . . 57

4 The metallicity distribution function of the Milky Way halo 59
4.1 Introduction . . . 61
4.2 Description of the data . . . 62
4.2.1 Photometric sample . . . 62
4.2.2 Spectroscopic sample . . . 63
4.3 Selection of the halo sample . . . 64
4.3.1 Distance slices into the Galaxy using MSTO stars . . . 65
4.3.2 The halo sample . . . 66
4.3.3 Distance ranges of the samples . . . 66
4.3.4 The MDF at different distances . . . 68
4.4 Correcting the MDF . . . 68
4.4.1 Corrections for the photometric metallicities . . . 69
4.4.2 Corrections for the halo sample selection . . . 70
4.5 Substructure in the halo MDF . . . 72
4.5.1 The MDF in different directions . . . 72
4.5.2 The Monoceros ring . . . 73
4.5.3 The contribution of the disc . . . 73
4.5.4 The Virgo overdensity . . . 74
4.6 Discussion . . . 75
4.6.1 Biases in the photometric metallicities . . . 76
4.6.2 Previous determinations of the halo MDF . . . 77
4.6.3 Comparison to the literature . . . 80
4.6.4 Comparison with Gaia-Enceladus . . . 82
4.6.5 The duality of the stellar halo . . . 82
4.6.6 Implications for the formation of globular clusters . . . 83
4.6.7 Future work . . . 83
4.7 Conclusions . . . 84
4.8 Appendix - Tables of the corrections and MDFs . . . 87

5 The Galactic outer halo with blue horizontal branch stars 91
5.1 Introduction . . . 93
5.2 Selection of BHB stars . . . 94
5.2.1 Towards a clean and pure BHB sample using SDSS . . . 95
5.2.2 Towards a clean and pure BHB sample using SDSS and *Pristine* . . . 98
5.2.3 Comparison between the two samples . . . 99
5.2.4 Completeness and purity as a function of magnitude . . . 100
5.2.5 Future outlook . . . 102

5.3 Results: Investigating the halo with BHB and BS stars . . . 102
5.3.1 The profile of the Galactic halo . . . 104
5.3.2 Clumpiness in the Galactic halo . . . 106
5.3.3 Tracing the Sagittarius stream further out . . . 109
5.4 Conclusions . . . 111

6 Quantifying substructure in the Milky Way halo 113
6.1 Introduction . . . 115
6.2 The data set . . . 117
6.2.1 Photometric metallicities from the *Pristine* survey . . . 117
6.2.2 The *Pristine* footprint . . . 118
6.2.3 Quality cuts on the data . . . 119
6.2.4 What Galactic environments are being probed? . . . 119
6.3 The two-point correlation function . . . 121
6.3.1 Tests of the two-point correlation function . . . 122
6.4 Creating a random sample . . . 124
6.4.1 A mock Milky Way . . . 124
6.4.2 Including holes and chip gaps . . . 125
6.4.3 Overlapping edges of the field . . . 126
6.5 Clustering results . . . 126
6.5.1 Regions 1 and 2 . . . 127
6.5.2 Regions 3, 4, and 6 . . . 128
6.5.3 Region 5 . . . 128
6.6 Discussion and conclusions . . . 129
6.6.1 Creation of the random sample . . . 130
6.6.2 From 2D to 3D . . . 130
6.6.3 Comparisons to simulations . . . 131
6.7 Appendix - Alternate random samples . . . 132

7 Conclusions & Future Outlook 137
7.1 Conclusions . . . 137
7.2 Preliminary impact of this thesis work . . . 137
7.3 Future Outlook . . . 138
7.3.1 The future of the *Pristine* survey: synergies with WEAVE and CFIS . . . 138
7.3.2 The Milky Way in a cosmological framework . . . 139
7.3.3 Gaia and spectroscopic surveys in the next decade . . . 140

Bibliography 141

Acknowledgements 149

Summary

The Milky Way is a spiral galaxy consisting of a disc of gas, dust and stars embedded in a halo of dark matter. Within this dark matter halo there is also a diffuse population of stars called the stellar halo, that has been accreting stars for billions of years from smaller galaxies that get pulled in and disrupted by the large gravitational potential of the Milky Way. As they are disrupted, these galaxies leave behind long streams of stars that can take billions of years to mix with the rest of the stars in the halo. Furthermore, the amount of heavy elements (metallicity) of the stars in these galaxies reflects the rate of chemical enrichment that occurred in them, since the Universe has been slowly enriched in heavy elements (e.g. iron) through successive generations of stars which produce them in their cores and supernovae explosions. Therefore, stars that contain small amounts of heavy elements (metal-poor stars) either formed at early times before the Universe was significantly enriched, or in isolated environments. The aim of this thesis is to develop a better understanding of the substructure content and chemistry of the Galactic stellar halo, in order to gain further insight into the formation and evolution of the Milky Way.

The *Pristine* survey uses a narrow-band filter which specifically targets the Ca II H & K spectral absorption lines to provide photometric metallicities for a large number of stars down to the extremely metal-poor (EMP) regime, making it a very powerful data set for Galactic archaeology studies. In Chapter 2, we quantify the efficiency of the survey using a preliminary spectroscopic follow-up sample of $\sim$ 200 stars. We also use this sample to establish a set of selection criteria to improve the success rate of selecting EMP candidates for follow-up spectroscopy. In Chapter 3, we extend this work and present the full catalogue of $\sim$ 1 000 stars from a three year long medium resolution spectroscopic follow-up effort conducted as part of the *Pristine* survey. From this sample, we compute success rates of 56% and 23% for recovering stars with [Fe/H] < -2.5 and [Fe/H] < -3.0, respectively. This demonstrates a high efficiency for finding EMP stars as compared to previous searches with success rates of 3–4 %.

In Chapter 4, we select a sample of $\sim$ 80 000 halo stars using colour and magnitude cuts to select a main sequence turnoff population in the distance range $6 < d_{\odot} < 20$ kpc. We then use the spectroscopic follow-up sample presented in Chapter 3 to statistically rescale the *Pristine* photometric metallicities of this sample, and present the resulting corrected metallicity distribution function (MDF) of the halo. The slope at the metal-poor end is significantly shallower than previous spectroscopic efforts have shown, suggesting that there may be more metal-poor stars with [Fe/H] < -2.5 in the halo than previously thought. This sample also shows evidence that the MDF of the halo may not be bimodal as was proposed by previous works, and that the lack of globular clusters in the Milky Way may be the result of a physical truncation of the MDF rather than just statistical under-sampling.

Chapter 5 showcases the unexpected capability of the *Pristine* filter for separating blue horizontal branch (BHB) stars from Blue Straggler (BS) stars. We demonstrate a purity of 93% and completeness of 91% for identifying BHB stars, a substantial improvement over previous works. We then use this highly pure and complete sample of BHB stars to trace the halo density profile out to $d > 100$ kpc, and the Sagittarius stream substructure out to $\sim$ 130 kpc.

In Chapter 6 we use the photometric metallicities from the *Pristine* survey to perform a clustering analysis of the halo as a function of metallicity. Separating the *Pristine* sample into four metallicity bins of [Fe/H] < -2, $-2 <$ [Fe/H] < -1.5, $-1.5 <$ [Fe/H] < -1 and $-0.9 <$ [Fe/H] < -0.8, we compute the two-point correlation function to measure the amount of

clustering on scales of < 5 deg. For a smooth comparison sample we make a mock *Pristine* data set generated using the Galaxia code based on the Besançon model of the Galaxy. We find enhanced clustering on small scales (< 0.5 deg) for some regions of the Galaxy for the most metal-poor bin ([Fe/H] < -2), while in others we see large scale signals that correspond to known substructures in those directions. This confirms that the substructure content of the halo is highly anisotropic and diverse in different Galactic environments. We discuss the difficulties of removing systematic clustering signals from the data and the limitations of disentangling weak clustering signals from real substructures and residual systematic structure in the data.

Taken together, the work presented in this thesis approaches the problem of better understanding the halo of our Galaxy from multiple angles. Firstly, presenting a sizeable sample of EMP stars and improving the selection efficiency of EMP stars for the *Pristine* survey, paving the way for the further discovery of metal-poor stars to be used as probes to early chemical evolution. Secondly, improving the selection of BHB distance tracers to map out the halo to large distances, and finally, using the large samples of metal-poor stars to derive the MDF of the inner halo and analyse the substructure content at different metallicities. The results of this thesis therefore expand our understanding of the physical and chemical properties of the Milky Way stellar halo, and provide insight into the processes involved in its formation and evolution.

Zusammenfassung

Die Milchstraße ist eine Spiralgalaxie, die aus einer Scheibe aus Gas, Staub und Sternen besteht und in einen Halo aus dunkler Materie eingebettet ist. Der Halo beherbergt auch eine diffuse Sternpopulation, den sogenanten stellaren Halo, welcher seit Milliarden von Jahren kleinere Galaxien verschlingt, die durch das große Gravitationspotenzial der Milchstraße angezogen werden. Wenn sie von der Gezeitenkraft der Milchstraße zerrissen werden, hinterlassen diese Galaxien lange Sternenströme, die ihrerseits Milliarden von Jahren überdauern können, bis sie sich komplett mit dem Rest der Halosterne vermischen. Die Menge an schweren chemischen Elementen (Metallizität) in den Sternen dieser Galaxien verrät uns ihre Sternentstehungsgeschichte: aufeinanderfolgende Generationen von Sternen und ihre Supernovae Explosionen reichern die Galaxie langsam mit schweren Elementen an. Sterne mit sehr wenig schweren Elementen (metallarme Sterne) bildeten sich daher entweder früh vor der signifikanten Anreicherung des Universums, oder in isolierten Umgebungen. Das Ziel dieser Arbeit ist eine genauere Charakterisierung der Struktur und Zusammensetzung des galaktischen stellaren Halos, um so die Entstehung und Entwicklung der Milchstraße besser zu verstehen.

Metallarme Sterne sind schwierig zu finden, da sie viel seltener sind als Sterne mit höherer Metallität, wie etwa unsere Sonne. Die *Pristine* Survey verwendet einen Schmalbandfilter, um metallarme Sterne photometrisch zu identifizieren, und kann daher Metallizitäten für eine große Anzahl von Sternen viel schneller bestimmen als spektroskopische Untersuchungen. Dies macht die *Pristine* Survey zu einem wertvollen Datensatz für Studien der Frühgeschichte der Milchstraße. In Kapitel 2 quantifizieren wir die Effizienz der Survey mit einer Stichprobe von ca. 200 Sternen, die spektroskopisch nachbeobachtet wurden. Wir verwenden diese Stichprobe auch, um eine Reihe von Auswahlkriterien festzulegen, um die Erfolgsrate bei der Auswahl metallarmer Kandidaten für die Folgespektroskopie zu verbessern. In Kapitel 3 erweitern wir diese Arbeit und präsentieren den vollständigen Katalog von ca. 1 000 Sternen aus einem dreijährigen spektroskopischen Follow-up-Projekt, das im Rahmen der *Pristine* Survey durchgeführt wurde. Diese Stichprobe zeigt, dass die *Pristine* Survey circa fünfmal effizienter im Finden extrem metallarmer Sterne ist als frühere Studien.

In Kapitel 4 wählen wir eine Stichprobe von ca. 80 000 Halosternen aus der *Pristine* Survey aus, um die Verteilung der Sterne an verschiedenen Metallizitäten im Halo zu analysieren. Wir verwenden die spektroskopische Stichprobe von Kapitel 3 für die Reskalierung der photometrischen Metallizitäten, um Verzerrungen in der Probe zu reduzieren. Die Steigung am metallarmen Ende dieser Verteilung ist deutlich flacher als in früheren spektroskopischen Studien gemessen. Das bedeutet, dass es viel mehr sehr metallarme Sterne im Halo geben könnte als bisher angenommen.

Blaue Horizontalaststerne (BHB-Sterne) sind sehr hell und ihre Entfernungen können sehr genau bestimmt werden. Deswegen sind sie ideale Standardkerzen für Studien von Sternpopulationen im galaktischen Halo. Kapitel 5 zeigt, wie die *Pristine* Survey dazu verwendet werden kann, BHB-Sterne besser zu identifizieren hinsichtlich ihrer Reinheit und Vollständigkeit als frühere Studien. Wir verwenden dann diese Stichprobe von BHB-Sternen, um das Dichteprofil des Halos sowie einige bekannte Halosubstrukturen über weite Entfernungen zu verfolgen.

In Kapitel 6 verwenden wir die photometrischen Metallizitäten aus der *Pristine* Survey, um eine Clustering-Analyse des Halos für verschiedene Metallizitäten durchzuführen. Wir finden stärkeres Clustering auf kleiner Skala (< 0,5 Grad) für einige Regionen der Galaxie

im metallärmsten Bereich, wohingegen wir in anderen Regionen große Signale sehen, die bereits bekannten Substrukturen in den jeweiligen Richtungen entsprechen. Dies zeigt die Fähigkeit dieser Methode, quantitative Vergleiche des Substrukturgehalts für verschiedene Proben und in verschiedenen galaktischen Umgebungen durchzuführen. Wir diskutieren auch die Schwierigkeiten bei der Entfernung systematischer Clustering-Signale aus den Daten und die Grenzen der Entflechtung schwacher Clustering-Signale von realen Substrukturen.

Zusammengefasst nähert sich diese Arbeit dem Problem eines besseren Verständnisses unseres Galaxienhalos aus mehreren Blickwinkeln. Als erstes stellt sie die *Pristine* Survey vor, die uns eine neue beträchtlichen Stichprobe von extrem metallarmen Sternen mit verbesserter Selektionseffizienz liefert. Dies ebnet den Weg für die Entdeckung weiterer metallarmer Sterne, die als Sonden für die frühe chemische Evolution verwendet werden können. Außerdem verbessert diese Arbeit das Auswahlverfahren für BHB-Sterne zur großräumigen Vermessung des galaktischen Halos, die Verwendung von großen Stichproben metallarmer Sterne zur Ableitung der Metallizitätsverteilung des inneren Halos, sowie die Analyse von Halosubstrukturen für verschiedene Metallizitäten. Die Ergebnisse dieser Arbeit erweitern somit unsere Kenntnis physikalischer und chemischer Eigenschaften des Milchstraßenhalos und vertiefen unser Verständnis von dessen Entstehung und Entwicklung.

1 Introduction

The Milky Way is just one Galaxy among several hundreds of billions in the observable Universe, but it is a particularly special Galaxy in that it is our home. In some ways it is the Galaxy that we can study the best, since nowhere else can we make such precise measurements of the chemical, kinematic, and physical properties of individual stars. However, from our perspective within the Galaxy, it can be difficult to reconstruct the true distributions of stars due to projection effects, and obfuscation by foreground stars, gas and dust. Combine this with the immense complexity of the many interacting and interdependent components, and it is no surprise that putting together a comprehensive and coherent view of our Galaxy has proven to be one of the greatest challenges of modern astronomy.

At the beginning of the 20th century, the idea that we live in a galaxy among billions of other galaxies was not yet widely accepted. In fact, in 1920, two very distinguished astronomers of the time, Harlow Shapley and Heber Curtis were advocating opposing views to explain our place in the Universe. Although Shapley correctly inferred that the Sun was not at the centre of the Galaxy due to the distribution of observed globular clusters (Shapley 1918), he incorrectly postulated that the observed spiral nebulae were merely clumps of stars that were still part of our own Galaxy, which constituted all of the stars in the Universe. Conversely, Curtis argued that these observed nebulae were actually each themselves external galaxies, much like our own Milky Way Galaxy, although he incorrectly calculated that the Sun was at the centre of the Milky Way. Just 100 years later, we have come a long way in understanding not only our position in our Galaxy, but also our place in the Universe. For instance, it is now known that the Sun resides 8.2 kiloparsecs (kpc) from the Galactic centre (see Abuter et al. 2019, for a recent, very precise measurement from the GRAVITY collaboration with an uncertainty of 0.3%) on the Orion spiral arm, and that our Galaxy is but one of many hundreds of billions of other Galaxies in the observable Universe. However, it should not be forgotten that these discoveries were far from trivial. Without advanced telescopes and instruments allowing for precise observations, as well as a solid foundation of sophisticated theories to interpret them, even knowing where we are in our own Galaxy is a challenging task. How then, can we expect to understand all of the complex processes through which our Galaxy formed and evolved? Although we now understand the complexity of our Galaxy and the immensity of this task better than ever before, reflecting on the past and realising how far we have come in just 100 years certainly gives us great hope for the future. On our collective quest towards understanding our Galaxy and its rich history, every novel insight, every new discovery, and every PhD thesis is a small step in the right direction.

In this introduction, I will present an overview of some of the relevant tools and methods used for studying our Galaxy, and the concept of using stars and their properties to learn about its formation and evolution. I will describe how we search for and use chemically primitive stars in our Galaxy to probe the earliest epochs of star formation, and how we can use the *Pristine* survey to efficiently find these metal-poor stars. I will discuss a general overview of the presently understood morphology and stellar content of our Milky Way Galaxy, with a particular emphasis on the Galactic halo, the Galactic environment that is the focus of this thesis. I will also review the ways in which the chemistry, kinematics and the substructure of halo stars can be used to reconstruct the assembly history of the Galactic stellar halo, and help to disentangle the complex build-up of the halo over time through the accretion of satellite galaxies. In addition, I will describe the special types of stars that we can use as distance tracers to map the morphology and substructures of the halo out to large distances, and finally conclude with a summary of the contents of each of the chapters in this thesis.

1.1 Galactic archaeology

Astronomers have the luxury of being able to look back in time. Due to the finite nature of the speed of light, the light emitted from distant galaxies can take millions, or billions of years to reach us. Therefore, we see these galaxies as they were in the past, and can directly study evolving galaxies in their nascent and developmental phases. Naturally, the limitation with this approach is that the further away and thus the fainter a galaxy is, the more difficult it is to resolve its features and to study it in detail. Not surprisingly, such observations of distant galaxies are limited in their resolution, and often are restricted to a single spectrum of the integrated light of an entire galaxy. Alternatively, we can also look for nearby, long-lived objects which have retained information from the past, and use them to infer what the conditions were like at the time of its formation. This field is known as near field cosmology (see Freeman & Bland-Hawthorn 2002 and Frebel & Norris 2015 for thorough reviews on this topic), in contrast to classical cosmology which observes distant galaxies. By combining present day observations of the chemistry, kinematics and spatial distributions of these stars, we can construct a picture of how our Galaxy evolved to the state that we observe it today. In this way, just as archaeologists use artefacts and relics from a distant time to learn about the diets, cultures, and lifestyles of ancient societies, so too can astronomers use old stars to learn about the Galaxy as it was at earlier times. This is the essence of Galactic archaeology.

After the Big Bang, only hydrogen, helium, and a trace amounts of lithium were produced before the Universe expanded to a large enough volume such that the temperature and density were too low for further nucleosynthesis. For the next few hundred million years the Universe was dark, until the first stars formed and emitted their light. It was finally again possible for atomic nuclei to fuse and for the production of heavier elements to occur in the central regions of these first stars. These stars then exploded as supernovae, ejecting these heavier elements into the surrounding interstellar medium. The next generation of stars were then formed from this material, slightly enriched in their heavy element content as compared to their predecessors. This process repeated itself until the present day, where most Galactic environments – and as a result the stars born within them – are relatively rich in heavy elements. Our Sun, for example, – born some 5 billion years ago – belongs to this relatively enriched group of stars. In contrast, stars that are very deficient in heavy elements, referred to as metal-poor stars, are also likely to be old and to have formed at the early stages of this enrichment process, or to have formed in isolated environments that avoided mixing with enriched regions. Furthermore, most stars retain the same elemental abundance signatures in their photospheres throughout their lifetimes, meaning that we can use the observed current

Table 1.1 — Nomenclature for metal-poor stars. This Table is reproduced from Table 1 in (Beers & Christlieb 2005)

[Fe/H]	Term	Acronym
>+0.5	Super metal-rich	SMR
~0.0	Solar	-
<-1.0	Metal-poor	MP
<-2.0	Very metal-poor	VMP
<-3.0	Extremely metal-poor	EMP
<-4.0	Ultra metal-poor	UMP
<-5.0	Hyper metal-poor	HMP
<-6.0	Mega metal-poor	MMP

chemical abundance trends, kinematics, and spatial distributions of stellar populations in the Galaxy to discern how it evolved, especially when interpreted in conjunction with galaxy simulations. Galactic archaeology is therefore a powerful tool with which we can learn about the past of our Galaxy from how we observe it today.

1.1.1 The metallicity content of stars

In contrast to its everyday meaning, or how it is defined in a chemistry textbook, the word "metal" in the context of astronomy refers to any element heavier than hydrogen or helium. By definition, $X + Y + Z = 1.0$, where $X = \frac{m_H}{M}$, $Y = \frac{m_{He}}{M}$ and $Z = \sum \frac{m_i}{M}$, where m_i is the mass of element i, and M is the total mass of the star, such that X and Y are the mass fractions of hydrogen and helium contained in the star, respectively. Z is known as the metallicity, and is defined as the mass fraction of the star containing all elements heavier than hydrogen and helium. It is necessary to define a reference value for metallicity, so we use the Sun, since it is the star for which the most precise abundance determinations are possible. However, there is not a consensus on the specific value of the solar metallicity since it still relies upon hydrodynamical modelling of the solar atmosphere, but two commonly used values are Z = 0.0134 (Asplund et al. 2009) and Z = 0.0153 (Caffau et al. 2011b). The relative metallicity of a star is defined as $[M/H] = \log\left(\frac{Z}{Z_\odot}\right)$, such that a value of [M/H] = 0 corresponds to a metallicity equal to that of the Sun. However, the total metallicity of a star can be a difficult quantity to measure, since practically it is unfeasible to measure all elemental abundances from a spectrum. One of the easiest heavy elements to measure is iron (Fe), since there are a large number of iron absorption lines at a wide range of wavelengths. Therefore, the relative metal abundance of a star is often estimated using the Fe metallicity, $[Fe/H] = \log_{10}\left(\frac{N_{Fe}}{N_H}\right) - \log_{10}\left(\frac{N_{Fe}}{N_H}\right)_\odot$. This logarithmic scale means that a star with [Fe/H] = -1, has 1/10th the amount of iron as the Sun. The term metallicity can also be used generally to refer to [Fe/H], as is commonly done in the literature and also in this thesis, as long as this is explicitly stated to avoid confusion. A common classification system to describe stars of different metallicity was defined in Beers & Christlieb (2005), and is summarized in Table 1.1. This nomenclature is broadly used in the literature when discussing the iron content of stars, particularly at low metallicities, and is also used throughout this thesis, with the pertinent acronyms being redefined at the beginning of each chapter for the convenience of the reader.

The degree to which iron can be used as a proxy for the total metallicity of a star depends on how similar the relative abundance pattern is to that of the Sun. To first order, [M/H] ~ [Fe/H] is a reasonable approximation, however, this does not hold for all stellar populations. The iron

content differs from the total metal content for stars with enhanced α-element[1] abundances relative to the Sun, according to the approximation $[M/H] \sim [\mathrm{Fe/H}] + \log(0.694 \times 10^{[\alpha/Fe]} + 0.306)$ (Salaris & Cassisi 2005). For example, some stellar populations such as halo stars exhibit a typical alpha-enhancement of $[\alpha/\mathrm{Fe}] \sim 0.4$, which would equate to a correction of $[\mathrm{M/H}] \sim [\mathrm{Fe/H}] + 0.3$.

These deviations in the relative abundances of certain elements with respect to the solar abundance can be attributed to specific physical processes and thus can be used to trace the chemical evolution of the Galaxy. For example, type II (core collapse) supernovae eject a relatively large amount of α elements, and a relatively small amount of iron peak elements. These types of supernova explosions occur from massive stars, which are shorter lived and therefore this enrichment channel dominates on shorter timescales. In contrast, supernova type Ia eject a much higher relative amount of iron peak elements, and are typically the result of merging binary white dwarf stars, which evolve on longer timescales. Therefore, the relative ratio $[\alpha/\mathrm{Fe}]$ of a given stellar population contains information about the relative contribution of these two processes, and the "knee" (the location at which the $[\alpha/\mathrm{Fe}]$ decreases with respect to [Fe/H]) encodes the star formation history and thereby the rate of chemical evolution (e.g. McWilliam 1997; Matteucci & Recchi 2001). More specifically, a population in which chemical enrichment proceeds slowly will have a knee at a lower [Fe/H] values than a population in which chemical enrichment proceeds quickly (for a nice comparison of the knee of the Milky Way and several local group dwarf Galaxies see Tolstoy et al. 2009). For this reason, Galactic components with different timescales of formation such as the thin disc, thick disc and halo of the Milky Way can be distinguished chemically by their relative $[\alpha/\mathrm{Fe}]$ abundances for a given [Fe/H].

1.1.2 Stellar spectroscopy

The most commonly used method currently employed to measure the elemental abundances of stars is stellar spectroscopy. In simple terms, this is based on the principle that different atomic nuclei, each with specific energy levels of electron orbitals, absorb the light from the star as it is emitted. When observing the light emitted form a star, we see the light as it is emitted from the outer most layer, the photosphere. Therefore, from the light observed by a detector on a telescope, it is possible to measure the absorption lines in the spectrum, whose specific wavelengths and strengths encode the elemental composition and abundances present in the photosphere of the star.

Commonly used methods for spectral analysis include measurements of equivalent widths, and full spectral fitting of synthetic spectra (for example the FERRE code, used in Chapters 2 and 3) to determine the sizes and shapes of the atomic and molecular absorption lines, as well as the continuum flux level of the stellar spectrum. Both of these methods rely on physical models of convective and radiative physics in stars to produce synthetic spectra with absorption lines corresponding to a given abundance pattern for a star with a given set of stellar parameters. Typically, this is done under the assumption of 1-Dimensional modelling with Local Thermodynamic Equilibrium (1D LTE), since this provides a moderate approximation of the relevant stellar physics for abundance analyses, and is relatively computationally untaxing. However, with improved computing power in recent years, more complex modelling has been possible, namely in 3-D and in the regime of Non-Thermodynamic Equilibrium (NLTE; e.g. Auer & Mihalas 1972; Asplund 2005; Mashonkina et al. 2007, 2011; Bessell et al. 2015; Nordlander et al. 2017). Taking these effects into account can be quite important, especially for stars of very low metallicity where small changes in the line profiles can result in large

[1] Elements produced through the alpha process, i.e. O, Mg, Si, S, Ca and Ti

changes in the derived abundances. For example, it has been shown that for the most metal-poor star, SMSS0313-6708, that the [Fe/H] abundance can differ by 0.8 dex from 1D LTE analysis when full NLTE and 3D effects are considered (Nordlander et al. 2017). Full 3D NLTE analyses are still very computationally expensive, making them currently unfeasible for large scale spectroscopic analyses. However, as computing capabilities continue to improve, these methods are likely to become standard and ubiquitous treatments for detailed stellar abundance analyses in the future.

1.1.3 The most metal-poor stars

Some of the most metal-poor stars in the Galaxy can be traced back to belonging to the very earliest generations of stars, with their chemical abundance patterns being consistent with the yields of just a single metal-free supernovae explosion (e.g. Keller et al. 2014; Takahashi et al. 2014; Ishigaki et al. 2014; Iwamoto et al. 2005). The most metal-poor stars in the Galaxy are therefore powerful probes to look back at the conditions present at early times in the Galaxy, and particularly to reveal information about the very first stars formed in the Universe. These metal-poor stars –and in particular metal-poor stars that are also old– have been shown in simulations to preferentially reside in the outskirts of the Galaxy, in low mass dwarf galaxy systems, and in the Galactic centre (e.g. Starkenburg et al. 2017b; El-Badry et al. 2018). Therefore, the stellar halo of the Milky Way is a prime environment to search for these stars, especially since observing the halo largely avoids the abundant obscuring gas and dust found on lines of sight towards the Galactic center. Nevertheless, these stars are rare amongst the much more common metal-rich stars which dominate most regions of the Galaxy. In Chapter 2, we make rough estimates for the frequency of EMP stars using spectroscopic follow-up data from the *Pristine* survey, and calculate that ~ 1 in 800 stars in the Galactic halo in the magnitude range $14 < V < 18$ are EMP stars. Extrapolating this to the UMP regime using the rough estimates from Schörck et al. 2009 and Allende Prieto et al. 2014, namely that there are $\sim 1\%$ as many UMP stars as EMP stars (this estimate is further corroborated by our work in Chapter 3), then only ~ 1 in 80 000 stars will have [Fe/H] < -4. It is these ultra, hyper, and mega metal-poor stars which are the most interesting in terms of revealing information about the epoch of reionization and the first generations of stars in the Universe. At the time of writing of this thesis, less than 50 stars have been discovered with [Fe/H] < -4 (e.g. Bonifacio et al. 2012; Hansen et al. 2014; Placco et al. 2015; Allende Prieto et al. 2015a; Aguado et al. 2017a), and only a handful of stars with [Fe/H] < -5 (Christlieb et al. 2004; Frebel et al. 2005; Keller et al. 2014; Bonifacio et al. 2015; Aguado et al. 2018b,a; Nordlander et al. 2019). These stars are shown in Figure 1.1 as black points, with their corresponding carbon abundances as the y-axis. All stars with [Fe/H] < -4.5 are labelled with the citation of their original discovery. Stars which have updated abundances with 3D and/or NLTE analyses are shown with blue points and connected with a dashed line to the original 1D LTE abundance values.

1.1.4 Carbon content in metal-poor stars

Another commonly used element for Galactic archaeology studies is carbon. Carbon is relatively easy to measure in a stellar spectrum since it can be measured in several wavelength regions either with atomic lines or large molecular band features (e.g Asplund et al. 2009). It is formed predominantly in the cores of evolved stars through the triple-alpha process (the fusion of three helium-4 nuclei), and its prevalence makes it an important tracer of chemical evolution in the Galaxy. Figure 1.1 shows the measured [C/Fe] and [Fe/H] values for many of the current known most metal-poor stars. At these low metallicities, an interesting trend appears, with the number of stars enhanced in carbon increasing with decreasing metallicity. Stars are defined to

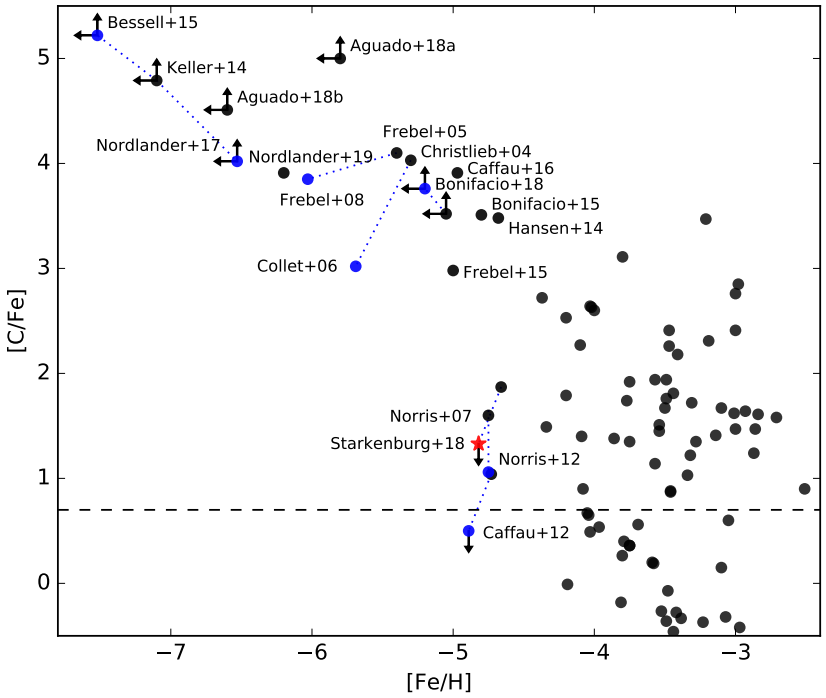

Figure 1.1 — Figure showing the ratios of iron-to-hydrogen and carbon-to-iron abundances for the most metal-poor stars. All stars with [Fe/H] < −4.5 are labelled with the reference corresponding to their original discovery. Black points are the 1D LTE abundances and blue points are the 3D or NLTE measurements (for the stars for which they are available), connected for a given star with a blue dashed line. The arrows show upper or lower limits on the abundance measurements, and the dashed line shows the definition for carbon enhancement from Aoki et al. (2007) of [C/Fe] = +0.7.

be enhanced in carbon if the ratio of [C/Fe] > +0.7 (Aoki et al. 2007), although other works have used different definitions (e.g. [C/Fe] > +1.0; Beers & Christlieb 2005). Stars of low metallicity that are enhanced in carbon are appropriately referred to as carbon enhanced metal-poor (CEMP) stars. About 15-20% of stars with [Fe/H] < −2 and ∼ 40% of stars with [Fe/H] < −3 are CEMP stars (e.g. Yong et al. 2013b; Placco et al. 2014, see also Chapter 3). Interestingly, despite their extremely low iron content, nearly all of the stars with [Fe/H] < −4.5 are significantly enhanced in carbon. At least 3 stars show significantly lower [C/Fe] than other stars of similar [Fe/H], namely HE0557-4840 (Norris et al. 2007, 2012), J1029+1729 (Caffau et al. 2011a) and Pristine_221.8781+9.7844 (Starkenburg et al. 2018), the latter of which was discovered as part of the *Pristine* survey, and is labelled as the red star[2] in Figure 1.1. This suggests that there are differential chemical enrichment channels responsible for the formation of these carbon enriched and non-carbon enriched subgroups of UMP stars.

Additional elemental abundances, such as [Ba/Fe] and [Eu/Fe] can help to differentiate the enrichment processes responsible for this observed carbon enhancement. The rapid neutron-capture process (r-process) and slow neutron-capture process (s-process) are the two main

[2]The NLTE corrections for Pristine_221.8781+9.7844 shown in the Figure are based on the corrections computed in Caffau et al. (2012) for the very similar star J1029+1729

mechanisms through which atomic nuclei heavier than iron form. The r-process proceeds in environments of high free neutron flux, where nuclei are bombarded with neutrons such that they accrete them faster than they can decay, resulting in very fast build up of large, exotic atomic isotopes. These then decay down to the more stable forms of various heavy atoms, such as Europium (Eu), a commonly used tracer for this process. The s-process occurs when neutrons are added slowly, such that only stable isotopes will continue to build while unstable ones will decay. This process is common in asymptotic giant branch (AGB) stars, a later evolutionary phase of low to intermediate mass stars in the mass range of $\sim 1-4M_{\odot}$. A common tracer for this process is Barium (Ba), as it is relatively easy to measure. Stars enriched in Ba are therefore referred to as CEMP-s stars, stars enriched in Eu are CEMP-r stars, stars enhanced with both are CEMP-rs stars, and stars not enhanced in s- or r-process elements are CEMP-no stars. It has also been shown that CEMP-s and CEMP-rs stars typically have higher absolute carbon (A(C)) abundances than CEMP-no stars, such that these two populations can be separated based on their locations in an A(C) vs [Fe/H] diagram (Spite et al. 2013; Yoon et al. 2016). Most of the stars that have CEMP-s abundance patterns are in binary systems, compared to CEMP-no stars which seem to have a binary fraction that is consistent with that of carbon normal giant stars (e.g. Starkenburg et al. 2014; Hansen et al. 2016a,b; Arentsen et al. 2019). This favours the scenario that CEMP-s stars receive their carbon (along with s-process elements) from the stellar winds of a nearby AGB companion (e.g. Abate et al. 2015), whereas CEMP-no stars are likely to have obtained their carbon directly from their initial gas cloud at the time of formation. It has been suggested that spin stars which, due to their high rotation, eject carbon into the surrounding interstellar medium (ISM) could be one mechanism to explain how carbon can be enriched in a primordial gas cloud while iron is not (e.g. Meynet et al. 2006, 2010; Chiappini 2013). Additionaly, faint supernovae have also been suggested as another mechanism that can contribute to this observed abundance ratio. This theory refers to supernovae that only partially explode, and due to or mixing-and-fallback mechanisms eject only the light elements in the outer layers into the ISM while retaining heavier elements –like iron– in their cores (Umeda & Nomoto 2003, 2005; Nomoto et al. 2013; Tominaga et al. 2014).

Clearly, the detailed chemical fingerprints of the most metal-poor stars contain a wealth of information about the processes involved in their formation and the interactions with their environment that have led to their present day abundance patterns. However, due to the low number of known stars at these lowest metallicities, the CEMP ratios and other abundance characteristics are still not statistically well constrained. It is therefore highly desirable to find more of these stars, and to obtain high-resolution follow-up with detailed measurements of as many elements as possible. This, in combination with the kinematic signatures and spatial distributions of these stars are pivotal contributions to the advancement of our understanding of our Galaxy and its formation.

1.2 The *Pristine* Survey

Given the rarity of extremely metal-poor stars in the Milky Way, searching for them randomly with single slit spectrographs is an ill-advised usage of telescope time. However, many metal-poor star searches have successfully mined general, large scale spectroscopic surveys such as SDSS/SEGUE/BOSS and LAMOST. Stars that look to be metal-poor in the low- or medium-resolution spectra of these surveys can then be followed up at high-resolution to obtain precise elemental abundances (e.g. Caffau et al. 2011b, 2013a; Aoki et al. 2013; Allende Prieto et al. 2015a; Bonifacio et al. 2015; Aguado et al. 2016, 2017a,b, 2018a,b; Bonifacio et al. 2018b; Li et al. 2018; Mardini et al. 2019a,b). In contrast to this approach, other studies have gone for more targeted searches and increased the efficiency by specifically targeting metal-poor stars based

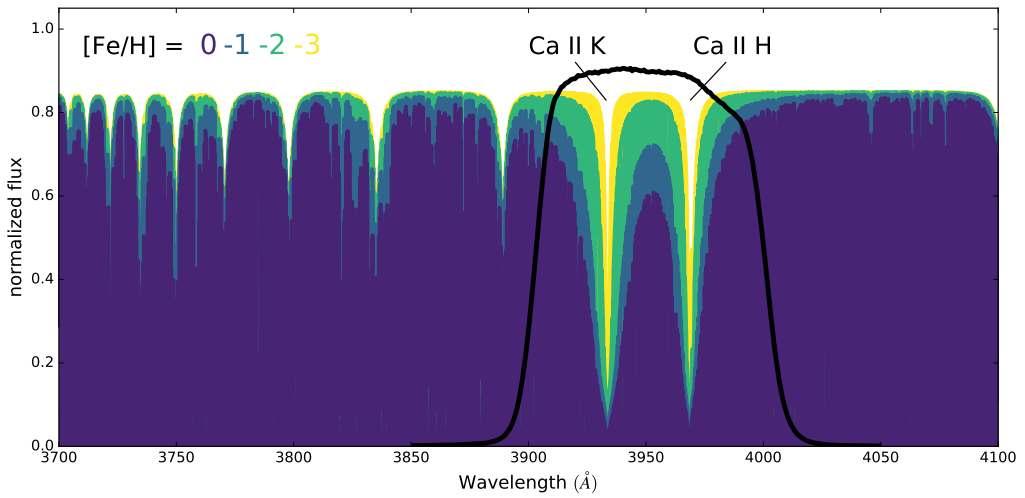

Figure 1.2 — The narrow-band *Pristine* filter is shown as the black dotted line. Also shown are synthetic spectra with effective temperatures of 4 500 K, surface gravities of 1.5, and varying metallicities as indicated by the legend.

on distinguishing characteristics, namely their small metallic absorption lines. One spectral region of particular interest for this approach is at blue optical wavelengths and contains the Ca II H & K absorption lines at 3969 and 3933 Å, respectively. These singly ionized absorption lines of Calcium are the strongest metallic absorption lines in the optical wavelengths and are generally a good proxy for the iron content of a star. In metal-poor stars there are fewer Ca atoms in a star's photosphere to absorb outgoing light from the star, resulting in smaller absorption lines. Furthermore, in hot stars the Ca atoms in the photosphere are more likely to be doubly ionized, and this increased excitation of the Ca II atoms results in a further weakening of their absorption lines. Therefore, there is a degeneracy between metallicity and temperature which must be broken with independent temperature measurements. The *Pristine* survey addresses this by combining the narrow-band observations with broad-band photometry from large scale surveys such as SDSS or Gaia. The photometric colours from these broad-band surveys can then be used to estimate the continuum shape of the spectrum and approximate the temperature of a given star. Finally, it is worth noting that most of the metallicity information is contained in the Ca II K line, as the Ca II H line is blended at low resolutions with an adjacent the H_ϵ line, which can remain strong in hotter stars.

Early works including the HK survey (Beers et al. 1985, 1992) and the Hamburg ESO survey (HES Christlieb et al. 2002) used Ca II H & K objective-prism surveys to identify stars with weak Ca II K lines for spectroscopic follow-up. Other studies have used broad-band photometry, or specifically designed narrow-band filters to target this spectral region and derive metallicities photometrically (e.g., Ivezić et al. 2008; Anthony-Twarog et al. 1991, 2000; Keller et al. 2007; Gu et al. 2015; Ibata et al. 2017a,b; Casagrande et al. 2019; Da Costa et al. 2019).

The *Pristine* survey expands on these works, by utilising a specially designed narrow-band filter which specifically targets the Ca II H & K absorption lines. Figure 1.2 shows the wavelength coverage and throughput of the *Pristine* filter, overlaid on synthetic spectra with fixed temperatures and gravities but varying [Fe/H]. This clearly shows the difference in flux for stars of high and low metallicity. The great advantage of the *Pristine* filter is that it is mounted on the MEGACAM instrument at the Canada France Hawaii Telescope, with a large 1 x 1 deg

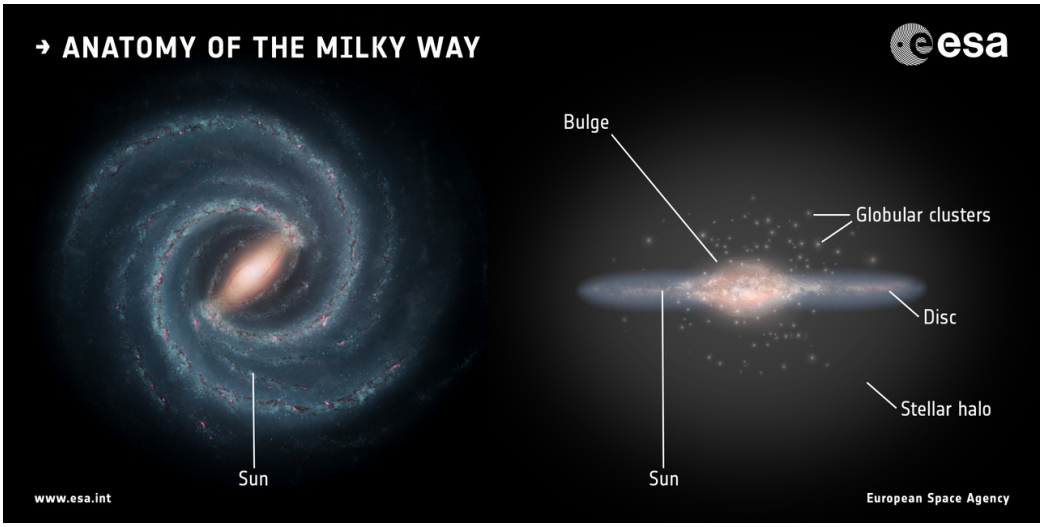

Figure 1.3 — An artist's impression of the anatomy of the Milky Way. The left image shows the spiral arms which make up the disc with the pronounced bar and bulge at the centre, and the Sun located about 8.2 kpc from the Galactic centre. The right image shows the canonical components of the Milky Way described in the text, including the bulge, disc, and stellar halo, with globular clusters dispersed throughout. Image credit: NASA/JPL-Caltech (left), ESA (right), ESA/ATG medialab (layout).

field of view and a 4.3m diameter primary collecting mirror. In combination with the filter being very narrow and nearly top hat shape, this makes for a very efficient tool for searching for metal-poor stars. In just a 100s exposure, the survey can go to a depth of $V = 21$ with a S/N ~ 10 (for further technical details of the survey see Starkenburg et al. 2017). The survey currently covers a region of ~ 5 000 deg^2, predominantly in the northern Galactic halo, and includes auxiliary programs targeting the Galactic bulge region (Arentsen et al. 2020) and nearby dwarf Galaxies (Longeard et al. 2018, 2019).

So far, the *Pristine* survey has been shown to be very efficient at finding EMP stars (see Chapters 2 and 3), and has already found one of the most metal-poor stars discovered to date, Pristine_221.8781+9.7844 (Starkenburg et al. 2018), in the course of its spectroscopic follow-up campaign. The data products of the survey are used throughout this thesis, as well as in many other recent and ongoing works (Starkenburg et al. 2017; Caffau et al. 2017; Starkenburg et al. 2018; Longeard et al. 2018; Sestito et al. 2019a; Longeard et al. 2019; Bonifacio et al. 2019; Venn et al. 2019; Sestito et al. 2019b; Arentsen et al. 2020).

1.3 The Milky Way

The Milky Way is a barred spiral Galaxy with a central box/peanut bulge, a dominant disc, and a diffuse stellar halo (for a recent review of the structural, kinematic and integrated properties of the Milky Way see Bland-Hawthorn & Gerhard 2016). Figure 1.3 shows a schematic view of the Milky Way, with its spiral arms and various components. The disc contains most of the baryonic mass of the Galaxy (i.e. stars, gas, and dust) while the halo contributes most of the total mass of the Galaxy but in the form of dark matter, with only a small amount of mass in the form of stars. The disc is commonly further separated into a thin and thick component, and the halo is sometimes separated into an inner and outer component, although these separations can vary depending on whether they are done morphologically, chemically or kinematically. Although

it is sometimes convenient to separate the Galaxy into these discrete components to quantify their structural parameters so that they can be compared to the equivalent observed structures in other galaxies, in reality there is large overlap between these components at any given position in the Galaxy. For example, the local neighbourhood around the Sun is comprised of a complex superposition of stars from the thin disc, thick disc, and halo. This speaks to the complex, interconnected way in which the Galaxy has formed and evolved over time.

The total mass of the Milky Way contained within the virial radius is $M_{vir} =\sim 0.5-2.5\times 10^{12}M_{\odot}$, and varies substantially depending on the methods used and assumptions made to measure it (for a summary of the different methods and mass estimates see Section 6.3 of Bland-Hawthorn & Gerhard 2016). One recent estimate favours an intermediate value for the mass of the Galaxy of $M_{vir} =\sim 1.54^{+0.75}_{-0.44}\times 10^{12}M_{\odot}$ (Watkins et al. 2019). A large part of this mass is in the form of dark matter, with the stars totalling a mass of $\sim 5\times 10^{10}M_{\odot}$, of which the halo contributes $\sim 1\%$ of this, and the rest coming from the disc and bulge (Bland-Hawthorn & Gerhard 2016, and references therein). The mass of the stellar halo is difficult to estimate since a lot of it is contained in substructures, but summing up the largest of these and adding in the measurable smooth component gives a total stellar mass of $\sim 4-7\times 10^{8}M_{\odot}$ (Bland-Hawthorn & Gerhard 2016). Despite the Galactic halo only contributing a small fraction of the total stellar mass of the Galaxy, it is still an important environment in which to study its evolutionary processes and global properties of the Galaxy, such as the accretion of satellite galaxies and star clusters and the gravitational potential of the Galaxy.

1.3.1 The Galactic halo

The Galactic halo is distinct from the disc and bulge of our Galaxy in its chemistry, kinematics and stellar density. In velocity space, disc stars typically have higher tangential velocities than halo stars do and are confined to the plane of the disc. Halo stars typically have higher vertical and radial velocities since their orbits take them high out of the plane of the disc. These components can also be separated by their chemistry, as disc stars have higher [Fe/H] and lower [α/Fe] on average than halo stars do. This can be traced back to differing formation channels, with halo stars typically having formed early and coming, at least partially, from smaller dwarf galaxy systems in which star formation was short lived, whereas star formation in the disc continued on longer timescales such that type Ia supernovae could contribute more significantly to enrich the environment in iron-peak elements (see for example McWilliam 1997; Helmi 2008; Bland-Hawthorn & Gerhard 2016, and references therein).

It has also been suggested that the halo of the Galaxy is made up of two distinct components, which can also be distinguished by their kinematics and chemistry (Carollo et al. 2007, 2010; Beers et al. 2012). In this paradigm, the inner halo dominates at heliocentric distances of $10-15$ kpc, and the outer halo dominates beyond $15-20$ kpc. These two populations are characterized by different metallicity distribution functions (MDFs), with the inner halo having a peak [Fe/H] ~ -1.5, and the outer halo being more metal-poor with a peak of [Fe/H] ~ -2.0 (Carollo et al. 2010; An et al. 2013, 2015; Zuo et al. 2017). However, an independent analysis of the spectroscopic data set has suggested that the two distinct halo components in velocity space seen in Carollo et al. (2007, 2010) and Beers et al. (2012) can be attributed to selection biases and large uncertainties on the distance determinations (Schönrich et al. 2011, 2014). Furthermore, other studies investigating the MDF of the of halo samples do not show strong evidence for two distinct components, but rather just one with a peak at [Fe/H] ~ -1.6 and a tail extending to very low metallicity (Ryan & Norris 1991; Schörck et al. 2009; Allende Prieto et al. 2014, see also Chapter 4).

Many previous works have used distance tracers to investigate the density profile of the Galacitic halo out to large radii, including RR Lyrae (Watkins et al. 2009; Hernitschek et al. 2018),

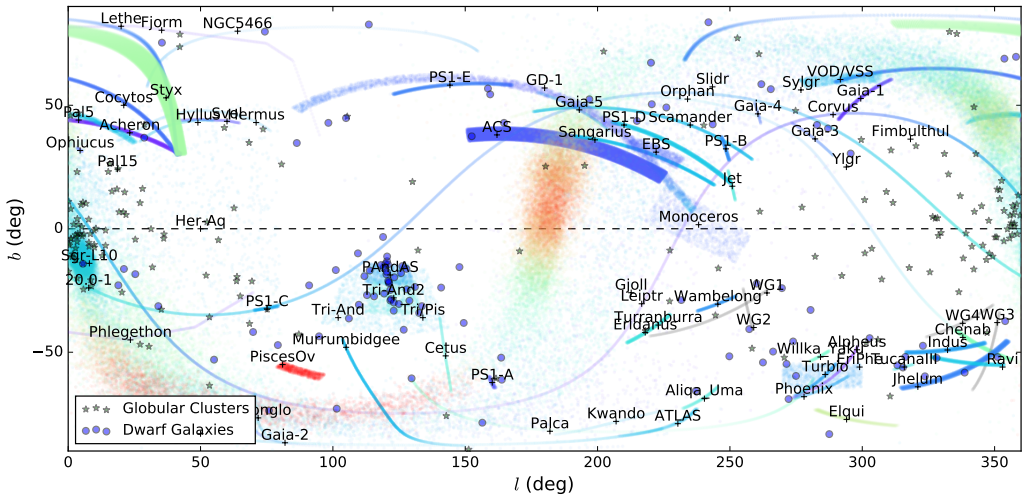

Figure 1.4 — The rich substructure content of the Milky Way. The known local dwarf galaxies are shown as blue dots and the known Milky Way globular clusters as green triangles. Stellar streams and diffuse structures are plotted to show their approximate observed distributions on the sky. This figure was produced using the python package galstreams (Mateu et al. 2018).

BHB (Deason et al. 2014; Fukushima et al. 2018; Thomas et al. 2018b), and giant stars (Xue et al. 2015; Slater et al. 2016). Some of these studies have suggested that there is a break in the stellar density profile between 20-40 kpc (Thomas et al. 2018b; Deason et al. 2014; Watkins et al. 2009). However, the existence of this break is not ubiquitously found, as some other authors present profiles that are consistent with a single component (De Propris et al. 2010; Slater et al. 2016; Hernitschek et al. 2018, see also Chapter 5). Importantly, these fits to the smooth density profiles depend on the assumed flattening of the halo at different radii, and on the large amount of substructure present in the halo. Therefore, all of these studies are limited by their ability to remove known substructures, and the extent of the coverage of the data they use, which is highly variable among the studies and the tracers used. Furthermore, it does not necessarily follow that a halo that is dual in its kinematics and chemistry should necessarily also show a break in stellar density, given that the mixing processes and resulting distributions of stars in the Galaxy are very complex, and that the halo components have had a long time to phase mix. It is also unclear to what extent trends observed using distance tracers such as blue horizontal branch stars, RR Lyrae, and selected samples of giants can be generalized to the entire stellar population.

1.3.2 Substructure in the halo

Although the stellar halo may appear smooth at first glance, a closer look reveals that it is in fact full of structure. These substructures are the remnants of accreted satellites whose jettisoned stellar material have built up the stellar halo over time. These include stars that are already highly phase mixed (making up the field star population), stars that are in the process of mixing (in the form of stellar streams and shells), and stars that are still bound in their original dwarf galaxy systems but are currently falling into the Milky Way potential. As a result, we can currently observe all stages of this accretion process, and learn about the formation processes that led to the halo that we see today. The prevalence of coherent structure in the halo has

become increasingly apparent in recent years with high quality, multi-band photometry over large portions of the sky made available by large scale surveys such as SDSS (Sloan Digital Sky Survey, York et al. 2000), PanSTARRS (Panoramic Survey Telescope and Rapid Response System Chambers et al. 2016), DES (Dark Energy Survey Abbott et al. 2018) and the Gaia mission (Gaia Collaboration et al. 2016, 2018). These fantastic data sets have allowed for the discovery of some 50 dwarf galaxies in the neighbourhood of the Milky Way, and many stellar streams. McConnachie (2012) provides a fairly recent compilation of the known dwarf galaxies in and around the Local Group, although there have been many new discoveries added in the past few years (e.g. Torrealba et al. 2019). Grillmair & Carlin (2016) provide a compilation of known streams and their properties, although there have also been many new recent discoveries (e.g. Koppelman et al. 2018; Shipp et al. 2018; Malhan et al. 2018; Ibata et al. 2019), particularly due to the exquisite astrometric and kinematic information provided by the Gaia DR2. Figure 1.4 shows an up to date map of the known substructures in the Milky Way. From this view, it is quite apparent that the halo contains a substantial amount of structure in the form of dwarf galaxies, globular clusters and stellar streams, that are direct evidence of its active accretion history. In addition, several studies have also demonstrated that a large fraction of the stars in the halo contribute to substructure, and are not smoothly distributed even if they can not be resolved into coherent clusters or streams (Bell et al. 2008; Starkenburg et al. 2009; Lancaster et al. 2019a, and Chapter 6). Clearly, there is a rich and active accretion history in our Galaxy, and there is much to be learned about the formation of the halo from the remains of these accreted objects.

Reconstructing the assembly history of the Galaxy is highly non-trivial, although with very good quality data it is possible to identify individual past accretion events by finding groups of stars that are still coherent today in their kinematic, chemical, and spatial distributions. Recently, the broad coverage and excellent quality of the second data release of the Gaia mission(Gaia DR2, Gaia Collaboration et al. 2018) has allowed for the identification of large accreted structures which seem to be dominating the inner region of the Galactic halo. Belokurov et al. (2018) identify a structure by selecting stars on highly eccentric orbits, and name this structure the Gaia Sausage, due to the elongated distribution of stars in the $v_\theta - v_r$ plane. This accretion event seems to be consistent with another large structure which has been named the Sequoia, and is characterized by a large number of stars on retrograde orbits (Myeong et al. 2019). In a different study, Helmi et al. (2018) identify a potentially related structure consisting of stars with high energy orbits, which they name Gaia Enceladus. This structure consists of both high eccentricity and retrograde stars, and it is currently unclear to what degree this structure is related to the Sausage and Sequoia.

In both of these scenarios, a large merger event with a single –or possibly multiple– progenitor(s) with a mass of $\sim 10^9 M_\odot$ (slightly more massive than the Small Magellanic Cloud) likely occured early on in the formation of the Galaxy, $\sim 8-11$ Gyr ago. It has also been suggested that these large merger events contributed most of the stellar content that presently makes up the inner halo, within a galactocentric radius of $\sim$ 30 kpc, and also heated the Galactic disc to produce what we see today as the thick disc. This inner halo component of stars moving on radial orbits has also been suggested to correspond to the density break separating the inner and outer halo, due to a pile-up of stars at the apocenter of their orbits from this early accretion event (Deason et al. 2018c). Furthermore, the estimated mass of these massive merging galaxies is consistent with recent mass estimates of the stellar mass of the halo within 100 kpc, of $1.5 \pm 0.4 \times 10^9 M_\odot$, and the peak of the metallicity distribution function (MDF) at [Fe/H] = -1.6 (Helmi et al. 2018) is qualitatively in preliminary agreement with what is found for the MDFs of several different halo samples (Ryan & Norris 1991; An et al. 2013, 2015; Allende Prieto et al. 2014, and Chapter 4). Although these recent findings have had a profound impact on our understanding of the accretion history of the halo and have provided insight into the potentially largest and most influential accretion events, the specific details of this are

still unknown, and more work is needed before we have a fully self consistent picture of the formation of the halo.

In addition to the Galactic halo component made up of accreted material from in-falling external satellites, there may also be a non-insignificant component of halo stars formed in-situ and kicked up from the disc. Several observations have been made of some strongly visible over-densities at low Galactic latitudes which are thought to be the result of kicked up disc material, ejected from a past accretion events of large satellites that passed through and perturbed the disc. In particular toward the Galactic anticenter, where there is a large amount of substructure at low galactic latitudes, such as the large Monoceros Ring structure, A13, TriAnd, the Anti-Center Stream (ACS) and the Easter Banded Structure (EBS). Although these structures were originally thought to be disrupted debris from externally accreted satellites, several recent studies have shown that the chemical (Bergemann et al. 2018), kinematic (Carollo et al. 2010; Li et al. 2017; de Boer et al. 2018; Deason et al. 2018a) and stellar population (Price-Whelan et al. 2015; Sheffield et al. 2018) properties suggest that some or all of these structures may not only be associated with each other but also are consistent with a disc origin. This has been corroborated with dynamical modelling showing that such substructures ejected from the disc can been attributed to the interaction of the pericentric passage of Sagittarius (Laporte et al. 2019, 2018). Furthermore, galaxy simulations have shown that for galaxies similar in size to the Milky Way it is not uncommon for a significant fraction of stars within R < 20 kpc to formed in the initial parent galaxy rather than from accreted material (Abadi et al. 2006). However, it is still not entirely clear if the Milky Way has a real population of halo stars that formed as an "in situ" component, or if this can be entirely attributed to kicked up disc stars, as is suggested in Di Matteo et al. (2019).

Taken together, these results show that despite significant recent progress in understanding the processes which contributed to the assembly of the Galactic halo, a clear picture of the full accretion history of our Galaxy is still lacking. Furthermore, the precise contribution of in situ and accreted material to the total stellar content of the halo still remains an open question, and the relative frequency of metal-poor stars in the halo (i.e. the metal-poor tail of the MDF) remains somewhat unconstrained. This thesis aims to contribute to the characterization of the chemical, morphological, and kinematic properties of the stellar halo population, in order to further constrain the possible formation and evolution scenarios, with the hopes that it will be possible to one day converge on a comprehensive depiction of the Galactic halo.

1.4 Thesis summary

1.4.1 Spectroscopic confirmation of an efficient search for extremely metal-poor stars

In Chapter 2, we use a sample of 205 spectroscopic follow-up stars from *Pristine* metal-poor star candidates to assess and improve the criteria used for the selection of candidates. We investigate several different parameter spaces, including multiple colour combinations and infrared selection criteria using WISE magnitudes (Schlaufman & Casey 2014), and summarize a list of selection criteria to use in subsequent follow-up efforts. These criteria predominantly remove brighter stars that are saturated in the SDSS g, r and i bands, or objects which do not have a reliable, single peaked point-spread function (PSF) determination. In addition, we restrict the selection to stars with a high probability of being metal-poor, based on a probability that is computed during the photometric metallicity determination.

We apply these criteria to the sample, which reduce it down to 149 stars, and compute success rates for the *Pristine* survey of 70% and 22% for recovering stars with [Fe/H] < -2.5

and [Fe/H] < −3.0, respectively. Our success rate of 22% to recover EMP stars represents a significant improvement with respect to previous searches which find success rates of ∼ 3 − 4% (Schlaufman & Casey 2014; Schörck et al. 2009). We also use the spectroscopic sample and rough estimates of the frequency of UMP stars in the halo from the literature (Schörck et al. 2009; Allende Prieto et al. 2014) to extrapolate preliminary estimates of the number of UMP stars that we expect to find in the future with the *Pristine* survey. This depends crucially on the soon to be operational large scale, multi-object spectroscopic (MOS) surveys such as WEAVE and 4MOST, which will make it possible to follow-up a significant portion of the total candidates. We estimate that we will find ∼ 1 000-1 200 EMP stars and ∼ 10-12 UMP stars per 1 000 deg^2 of sky covering in the magnitude range $V < 18$. If this prediction comes to fruition, this would mark a significant contribution to the current known sample of metal-poor stars and the field of Galactic archaeology.

1.4.2 The first three years of medium-resolution follow-up spectroscopy of *Pristine* EMP star candidates

In Chapter 3, we present the results of the first 3 years of spectroscopic follow-up of *Pristine* metal-poor star candidates. This chapter also focuses in more detail on the spectroscopic analysis of the sample using the FERRE code, which implements a χ^2 minimization algorithm to fit the observed spectrum with a grid of synthetic model spectra. Compared to the analysis in Chapter 2, we expand the grid of synthetic spectra to include higher metallicities ([Fe/H < -2), and to obtain carbon abundances for a subset of the stars. We perform an in-depth analysis of the parameter space for which reliable carbon abundances can be obtained, which depends on the effective temperature, [Fe/H], and S/N of the stars. In total, we present [Fe/H] abundances for 1 007 stars, [C/Fe] abundances for 169 of these, and report the discovery of > 700 new VMP stars with [Fe/H] < −2.0 and just under 100 new EMP stars with [Fe/H] < −3.0.

We use this significantly larger sample to expand upon the work of Chapter 2, and recompute the success rates of finding metal-poor stars. We update the success rates to 56% and 23% for recovering stars with [Fe/H] < −2.5 and [Fe/H] < −3.0, respectively. Furthermore, we find an average carbon enhancement fraction ([C/Fe] >= 0.7) of 41 ± 4% for stars with −3[Fe/H] < −2 and 58 ± 14% for stars with [Fe/H] < −3 which is slightly higher, but still in rough agreement with what has been reported earlier in the literature (Yong et al. 2013b; Placco et al. 2014). We then discuss in more detail the plans to follow-up *Pristine* candidates with WEAVE (with a memorandum of understanding now in place between the two surveys), which will dedicate up to twenty fibers per 3.14 deg^2 field of view to Pristine-selected EMP candidates in the magnitude range $15 < G < 19$, in the low-resolution Galactic archaeology survey of high Galactic latitudes. We project that over the 5 000 deg^2 of overlap between *Pristine* and WEAVE, this will yield ∼ 5 000-7 000 EMP star discoveries, for which ∼ 3 000 will have multiple abundance measurements, such as C, Na, Mg, Al, Si, Ca, Ti, Fe. Furthermore, we predict the discovery of ∼ 150 − 200 stars with [Fe/H] < −4.0, including ∼ 5 − 10 hyper metal-poor stars with [Fe/H] < −5.0, effectively doubling the samples currently available.

1.4.3 The metallicity distribution function of the Milky Way halo down to the extremely metal-poor regime

In Chapter 4, we use the large sky coverage and large number of stars for which we have *Pristine* photometric metallicities to produce a metallicity distribution function (MDF) of the Galactic halo. As described in Chapters 2 and 3, stars that are predicted to be EMP by *Pristine* are confirmed to be EMP by spectroscopy ∼ 23 % of the time. Therefore, if we want a realistic representation of the MDF, we need to rescale the photometric distribution

accordingly. Using the spectroscopic sample described in Chapter 3, along with a sample of overlapping SDSS/SEGUE stars to supplement the higher metallicity end at [Fe/H] > -2.5, we create a statistical relation between the photometric and spectroscopic metallicities and use this to rescale the *Pristine* photometric metallicities.

We then separate the total sample into heliocentric distance bins by selecting main sequence turn-off stars (MSTO) using a colour cut of $0.15 < g_0 - i_0 < 0.6$ and binning the sample by magnitude. The brightest of these bins, $15 < g < 16$, spans a distance range of $\sim 0.8 - 3.5$ kpc, and the faintest bin, $19 < g < 20$ spans a distance range of $\sim 4 - 21$ kpc. The MDFs of these bins show a clear decrease in peak metallicity with increasing distance, demonstrating a metallicity gradient going out into the halo. We further select a halo sample consising of $\sim$ 80 000 stars, by making additional cuts to the faintest MSTO sample, namely a narrower colour cut of $0.15 < (g_0 - i_0) < 0.4$. After investigation of other potential biases introduced in the MDF from the rescaling procedure and halo sample selection, and applying corrections where needed, we derive a halo MDF that is largely unbiased at metallicities of [Fe/H] < -1.5.

We compare the resulting halo MDF to those derived in previous works (Ryan & Norris 1991; Schörck et al. 2009; An et al. 2013; Allende Prieto et al. 2014; An et al. 2015) and to a simple, closed-box chemical evolution model (Hartwick 1976) and find that the *Pristine* halo sample MDF peaks at [Fe/H] $= -1.6$, which is in good agreement with the peak value of previous works. The metal-poor slope of the MDF substantially shallower than the slope presented with previous spectroscopic samples (Schörck et al. 2009; Allende Prieto et al. 2014) suggesting that there may be a larger fraction of stars at [Fe/H] < -2 in the halo than previously thought. We can not confirm a bimodal or multi-component MDF which would suggest a halo that is dual in nature, as was shown in previous works (Carollo et al. 2007; An et al. 2015; Zuo et al. 2017). Finally, based on a simplistic comparison, the *Pristine* halo sample suggests that the lack of globular clusters with [Fe/H] < -2.5 in the Milky Way is due to a physical truncation of the globular cluster MDF, rather than simply due to statistical undersampling.

1.4.4 A uniquely clean view of the Galactic outer halo using blue horizontal branch stars

In Chapter 5, we use the *Pristine* filter to differentiate between blue horizontal branch (BHB) stars and blue straggler (BS) stars. Such hot, A-type stars can be fairly easily selected based on their temperatures in broad-band colours, however, in this colour range there is a mixture of BHB and BS stars, which have characteristically different brightnesses. Therefore, in order to get a clean sample of standard candle BHB stars, it is necessary to have a means to separate these two populations. This can be done spectroscopically (e.g. Xue et al. 2008) by measuring the broadening of the balmer lines, since BS stars are more compact and have larger surface gravities than BHB stars. This difference in gravity is also very pronounced at blue wavelengths as a difference in the continuum near the Balmer jump of the spectrum. The *Pristine* filter is conveniently located right where the difference in the spectra between the two populations is largest, such that combining the $CaHK$ flux information with that of the SDSS u-band gives fairly strong discrimination power to differentiate between BHB and BS stars.

This method is similar to that implemented in Deason et al. (2011), but we show that the $u - CaHK$ colour combination is better able to separate BHB and BS stars than the $u - g$ combination used in that work. Using a sample of spectroscopically identified BHB stars (Xue et al. 2008), we compute the completeness of both the SDSS only sample from Deason et al. (2011) and the SDSS + *Pristine* sample derived in this work. We find that in the magnitude range $14.0 < g < 19.2$, the SDSS only sample has a purity of 93% and a completeness of 46%. With the added *Pristine* information the purity stays very similar, but the completeness rises substantially to 91%.

Using this highly pure and more complete sample, we trace the trailing and leading arms of the Sagittarius stream out to distances of > 100 kpc. Compared to previous work using RR Lyrae variable stars (Hernitschek et al. 2017), we find a similar morphology but but note that the BHB sample traces a smoother halo than the RR Lyrae stars do, and therefore we see the stream as a less pronounced overdensity. We tentatively confirm the possible existence of a spur feature that is part of the Sagittarius stream at a distance of ~ 130 kpc, and probe it slightly further out than previously possible with RR Lyrae stars. Further exploration of this spur feature is left to future work either with spectroscopic follow-up of distance BHB candidates, or with deeper u-band photometry from the Canada-France Imaging Survey (CFIS, Ibata et al. 2017a,b), expanding on the work of Thomas et al. (2018b). In addition, we find that the distance profile of the halo follows a power-law with an almost constant slope of 3.5–4.0, depending on if the Sagittarius stream is removed or not.

1.4.5 Quantifying substructure in the Milky Way halo at different metallicities

In Chapter 6, we seek to statistically quantify the amount of substructure in the Galactic halo as a function of metallicity. We use the photometric metallicities form the *Pristine* survey to separate the sample into metallicity bins of 0.5 dex, and also separate the ~ 2 500 deg^2 *Pristine* footprint into 6 different regions, corresponding to the 6 separate contiguous regions that make up the footprint. These regions cover a range of Galactic latitudes ($0 < l < 360$) and longitudes ($20 < b < 80$) in the northern Galactic halo, with the exception of region 7 which is in the south ($110 < l < 130$ and $-40 < b < -30$), allowing for the comparison between a variety of different halo environments.

To quantify the strength of clustering at different metallicities in each region we use the angular two-point correlation function (2-PCF), implementing the discrete estimator described in Landy & Szalay (1993). The resulting quantity, $\hat{\omega}(\theta)$, represents the excess probability of two stars having an angular separation of θ, with respect to a randomly distributed sample. It is therefore necessary to generate a random sample which appropriately models the smooth, large scale density variations introduced from the underlying Galactic components such as the disc, bulge and halo. For this purpose, we use the Galaxia code (Sharma et al. 2011) to generate a synthetic mock catalogue of the *Pristine* sample based on the Besançon model of the Galaxy (Robin et al. 2003). In order to further limit clustering signatures introduced from systematics such as the depth of the survey or the scanning law, we limit the depth of the survey to $16 < V < 20$ and mask the random sample to have the same holes and the regular chip gap patterns that result from the observations and photometric reductions.

We find that for regions 1 and 2, there is an enhanced clustering signal at small scales (< 0.5 deg) for the lowest metallicity bin ([Fe/H] < −2). We discuss the known substructures such as compact dwarf galaxies and globular clusters that are found in these regions, and conclude that these can not be the main cause of this signal. For region 1 we cautiously suggest that this signal may correspond to a conglomeration of diffuse material detected in this region of sky by the STREAMFINDER algorithm in Ibata et al. (2019), which despite being detected with high significance cannot be resolved into coherent stellar streams. We also note, however, that this signal may be the result of residual systematics, and suggest future work to confirm whether this observed signal is real or not. In region 5, in a low latitude anticenter direction, we find a large scale clustering signal at intermediate metallicities which we associate to the Monoceros/A13/ACS/EBS overdensity. The rest of the regions show no distinguishable difference in the clustering signals between the metallicity samples. We conclude the chapter by discussing the current limitations of this work in progress and the future directions in which it can be taken.

2 Spectroscopic confirmation of an efficient search for extremely metal-poor stars

K. Youakim, E. Starkenburg, D. S. Aguado, N. F. Martin, M. Fouesneau, J. I . González Hernández, C. Allende Prieto, P. Bonifacio, M. Gentile, C. Kielty, P. Côté, P. Jablonka, A. McConnachie, R. Sánchez Janssen, E. Tolstoy, and K. Venn

— Published as K. Youakim et al., 2017, MNRAS 472, 2963 —

ABSTRACT

The *Pristine* survey is a narrow-band, photometric survey focused around the wavelength region of the Ca II H & K absorption lines, designed to efficiently search for extremely metal-poor stars. In this work, we use the first results of a medium-resolution spectroscopic follow-up to refine the selection criteria for finding extremely metal-poor stars ([Fe/H] ≤ -3.0) in the *Pristine* survey. We consider methods by which stars can be selected from available broad-band and infrared photometry plus the additional *Pristine* narrow-band photometry. The spectroscopic sample presented in this chapter consists of 205 stars in the magnitude range $14 < V < 18$. Applying the photometric selection criteria cuts the sample down to 149 stars, and from these we report a success rate of 70% for finding stars with [Fe/H] ≤ -2.5 and 22% for finding stars with [Fe/H] ≤ -3.0. These statistics compare favourably with other surveys that search for extremely metal-poor stars, namely an improvement by a factor of $\sim 4-5$ for recovering stars with [Fe/H] ≤ -3.0. In addition, *Pristine* covers a fainter magnitude range than its predecessors, and can thus probe deeper into the Galactic halo.

2.1 Introduction

During Big Bang nucleosynthesis, only the lightest elements were produced in any significant quantity: mainly hydrogen, helium, and trace amounts of lithium. Nearly all heavier elements were formed later in the interiors of stars and released in their supernovae explosions, thereby enriching the metal content of the Universe over time (Alpher et al. 1948; Burbidge et al. 1957). The oldest stars that formed in the early Universe from this pristine gas should therefore be largely free from heavier elements due to their early time of formation. Analysis of these metal-poor stars, their stellar parameters, chemical abundances, dynamics, and spatial distributions in the Galaxy can offer insight into the local environments in which they formed, and thus help to invoke constraints on our understanding of the first generation stars that came before them and the Galaxy at early times (e.g., Freeman & Bland-Hawthorn 2002; Beers & Christlieb 2005; Frebel & Norris 2015).

Such stars are rare among the overwhelming numbers of more metal-rich populations in the Galaxy. From the Besançon model of the Galaxy (Robin et al. 2003), the expectation is that in a high Galactic latitude field towards the anti-centre direction ([l,b] = [0°, 60°]) and in the magnitude range $14 < V < 18$, only ~ 1/2 000 stars will be extremely metal-poor (EMP), with a metal content less than 1/1 000 of the Sun ([Fe/H] ≤ -3). This ratio increases to ~ 1/500 for $18 < V < 20$, as this fainter magnitude range probes more of the metal-poor halo rather than the metal-rich Galactic disc. Although these are only projections, since this model relies on assumptions about the EMP tail of the metallicity distribution function, they emphasize that an efficient pre-selection method is needed to find and study these very rare stars. This is one of the principal goals of the *Pristine* survey.

For stars close to the main sequence turnoff, broad-band optical colours hold some metallicity information due to line blanketing at blue wavelengths (Schwarzschild et al. 1955; Sandage & Eggen 1959; Wallerstein 1962; Ivezić et al. 2008). However, this relation typically breaks down at metallicities just below [Fe/H] = −2 (Starkenburg et al. 2017), which is the metallicity range of greatest interest to study the oldest and most pristine star formation environments. More recent work has shown that with good u-band data, the SDSS photometric metallicities can be extended into the [Fe/H] ~ −2.5 regime when a technique of multiple fitting to calibrated isochrones is used (An et al. 2013, 2015). The Canada-France Imaging Survey (CFIS) has also shown increased success in photometric metallicity determination with their high quality u-band observations (Ibata et al. 2017a,b). Although these recent advances have improved the capabilities of photometric metallicity calibrations, they still do not provide information for the EMP regime at [Fe/H] ≤ -3.

A recent study by Schlaufman & Casey (2014) implemented a combination of optical and infrared broad-band filters from the Wide-Field Infrared Survey Explorer (WISE, Wright et al. 2010), the Two Micron All-Sky Survey (2MASS, Skrutskie et al. 2006), and the AAVSO Photometric All-Sky Survey (APASS, Henden et al. 2009; Henden & Munari 2014; Henden et al. 2015) to photometrically identify EMP stars. Most of the selection power of this method is based on the strong molecular absorption down to [Fe/H] = −2 in the wavelength region covered by the WISE W2 filter (4.6 μ), such that colour combinations with the WISE W1 (3.4 μ) and the 2MASS J (1.2 μ) filters can effectively select metal-poor stars from photometry alone (see Figure 1 from Schlaufman & Casey (2014)). Nevertheless, due to the limiting magnitudes of the existing infrared photometry from WISE, and the quality cuts required for this method to work, this technique is mostly suited for very bright targets, and thus is mainly sensitive to local halo stars (both Schlaufman & Casey 2014 and Casey & Schlaufman 2015 adopt a faint threshold of $V = 14$ for their sample).

Whenever large samples of stars are targeted at a certain phase in their stellar evolution, their relative brightness will trace different distances and hence various environments inside

our Galaxy. Many of the most metal-poor stars known have magnitudes brighter than $V = 16$, due to the techniques with which they were discovered. Up until the last decade, the main sources for extremely metal-poor (EMP; [Fe/H] ≤ -3) stars, as well as a few ultra metal-poor (UMP; [Fe/H] ≤ -4) stars, were through Ca H & K objective-prism surveys such as the HK survey (with a magnitude limit of $B \sim 15.5$, Beers et al. 1985, 1992) and the Hamburg ESO survey (HES, with a magnitude limit of $B \sim 17 - 17.5$, Christlieb et al. 2002). In line with expectations, the fainter HES was more successful in finding EMP and UMP stars because it reached deeper into the metal-poor outer halo.

Although magnitude ranges often limit the distance range probed, there are still significant differences in the chemical properties of EMP stars and their present-day location and kinematics in the Galaxy (Cayrel et al. 2004; Frebel et al. 2006; Bonifacio et al. 2009; Carollo et al. 2010, 2012; An et al. 2013; Starkenburg et al. 2013; Skúladóttir et al. 2015). This dependence of chemical composition on Galactic environment has been further emphasized by recent studies in the bulge of the Galaxy, namely that EMP stars in the halo are often enhanced in carbon, whereas EMP stars in the bulge rarely exhibit carbon enhancement (Howes et al. 2015, 2016; Koch et al. 2016; Lamb et al. 2017).

The need for large samples of EMP stars across various environments and magnitude ranges has been somewhat mitigated in recent years by large scale, blind spectroscopic surveys. Some examples of these include the Sloan Digital Sky Survey (SDSS, York et al. 2000), as well as its dedicated constituent spectroscopic campaigns, the Sloan Extension for Galactic Understanding and Exploration (SEGUE, Yanny et al. 2009; Fernández-Alvar et al. 2015, 2016), and the Baryonic Oscillations Spectroscopic Survey (BOSS, Eisenstein et al. 2011; Dawson et al. 2013) (for higher-resolution follow-up of metal-poor stars based on these samples, see Caffau et al. 2013a; Aoki et al. 2013; Allende Prieto et al. 2015a; Aguado et al. 2016). These large surveys have the advantage that they probe deeper than the objective-prism surveys, and can obtain large numbers of good quality spectra. Nevertheless, the success rates for discovering metal-poor stars in these surveys are naturally low because they do not specifically target these stars, and EMP stars compose a very small fraction of the total stellar content of the Galaxy.

Several previous studies have described the approach of targeted narrow-band photometry on the Ca II H & K wavelength region as a means of providing metallicity information (e.g., Anthony-Twarog et al. 1991, 2000). One recent example which has been particularly successful is the *Skymapper* survey in the Southern Hemisphere (e.g., Keller et al. 2007), which uses a v filter (wavelength coverage $\sim 3\,650 - 4\,000$ Å, Bessell et al. 2011) to photometrically pre-select metal-poor star candidates for spectroscopic follow-up. Operating on a similar concept, the *Pristine* survey uses a specially designed filter (wavelength coverage $\sim 3\,900 - 4\,000$ Å) that is even narrower and more targeted on the Ca II H & K absorption lines. Although it covers less sky area compared to *Skymapper*, the *Pristine* survey is better suited to efficiently study fainter targets because it utilizes the 4m-class Canada-France-Hawaii Telescope (CFHT), which provides a large aperture and excellent image quality, and is located in the Northern Hemisphere where SDSS broad-band photometry is readily available.

In this work we use the first results of medium-resolution follow-up spectroscopy of 205 stars within the *Pristine* survey to assess the performance of the survey's photometric pre-selection. This sample consists of targets in the magnitude range $14 < V < 18$, and therefore represents only the brighter end of the full *Pristine* target sample. We use this sample to assess and improve the criteria used for selecting follow-up candidates for spectroscopy. By doing so, we pave the way to the successful follow-up of even fainter targets, opening up the possibility of the *Pristine* dataset to be used to efficiently select targets for large multiplexing spectroscopic surveys in the near future, such as the William Herschel Telescope Enhanced Area Velocity Explorer (WEAVE, Dalton et al. 2018), the 4-metre Multi-Object Spectroscopic Telescope (4MOST, de Jong et al. 2019), the Subaru Prime Focus Spectrograph (PFS, Takada

et al. 2014), or the Maunakea Spectroscopic Explorer (MSE, McConnachie et al. 2016b). These survey efforts are expected to probe more of the pristine environments in the outskirts of the Galactic halo; one of the regions in the Galaxy expected to harbour possible first star environments, as highlighted in recent analyzes of cosmological simulations (e.g., Starkenburg et al. 2017b).

The chapter is organized as follows: In Sections 2.2 and 2.3 we introduce the *Pristine* survey and its spectroscopic follow-up programme. In Section 2.4, we summarize and discuss improvements to our candidate selection criteria, as well as investigate whether infrared photometry and regularized regression techniques can be used to further improve our results. In Section 2.5, we present the results of the medium-resolution spectroscopic follow-up, in particular comparing the predicted photometric metallicities from *Pristine* to spectroscopically determined metallicities. Finally, in Section 2.6 we discuss the current purity and success rates of our target selection, compare them to expectations and other works, and discuss projections and strategies for the continuation of the *Pristine* survey. In this chapter, we demonstrate that the *Pristine* survey shows unparalleled efficiency for finding the most metal-poor stellar populations of the Galaxy. This is key for the eventual completion of two of the survey's main objectives, which include finding large numbers of EMP stars to contribute to the characterization of the extremely metal-poor tail of the metallicity distribution function, as well as uncovering the exceedingly rare UMP stars.

2.2 The *Pristine* survey

For a full and detailed description of the *Pristine* survey we refer the reader to Starkenburg et al. (2017, the first *Pristine* survey paper, hereafter referred to as Paper I). We recapitulate the essential elements of the survey here.

The *Pristine* survey uses a narrow-band Ca II H & K filter (hereafter referred to as the $CaHK$ filter) mounted on MegaPrime/MegaCam at the Canada France Hawaii Telescope (CFHT) on Mauna Kea in Hawaii. The filter was specifically designed by members of the *Pristine* team to cover the wavelength region of the singly ionized Ca II H & K lines, located at 3968.5 and 3933.7 Å, respectively. The narrow width of the filter reduces the influence of other spectral features, such as the nearby CN molecular absorption bands at 3839 and 4142 Å. MegaPrime/MegaCam fields are ~ 1 deg^2, and with integrations of 100 seconds, a signal to noise (S/N) of 10 at a depth of $g_0 \sim 21.0$ can be achieved (Paper I). As of September 2016 the sky coverage was $\sim 1\,000$ deg^2, and data collection is ongoing with the aim to cover at least $\sim 3\,000$ deg^2. The footprint of the survey targets the Galactic halo and intentionally spans a range in Galactic latitude ($30° < b < 78°$) to sample a diverse range of halo environments. Observations are made in the Northern Hemisphere, and overlap by design with regions of sky previously observed photometrically by SDSS. *Pristine* can therefore cross-match its targets with SDSS to obtain $ugriz$ broad-band photometry, which is useful for temperature determination and point source identification, allowing for the elimination of most objects that are not stars. Another important advantage to the overlap with SDSS is that there is a sample of several thousand stars distributed over the *Pristine* footprint for which moderate-resolution ($R \approx 1\,800$) spectra are already available from the SDSS and SEGUE surveys. Thus, *Pristine* has a large sample of spectroscopic metallicities which can be used to calibrate the assignment of photometric metallicities, thereby greatly reducing the amount of overhead and telescope time required to make the survey operational.

Figure 2.1 depicts the parameter space used for assigning photometric metallicities. The y-axis shows the colour of the SDSS g-band minus the $CaHK$ magnitude obtained from the *Pristine* narrow-band filter. An extra combination with the SDSS g and i or r magnitudes is added to stretch the plot vertically and make it easier to see the metallicity gradient. The x-axis

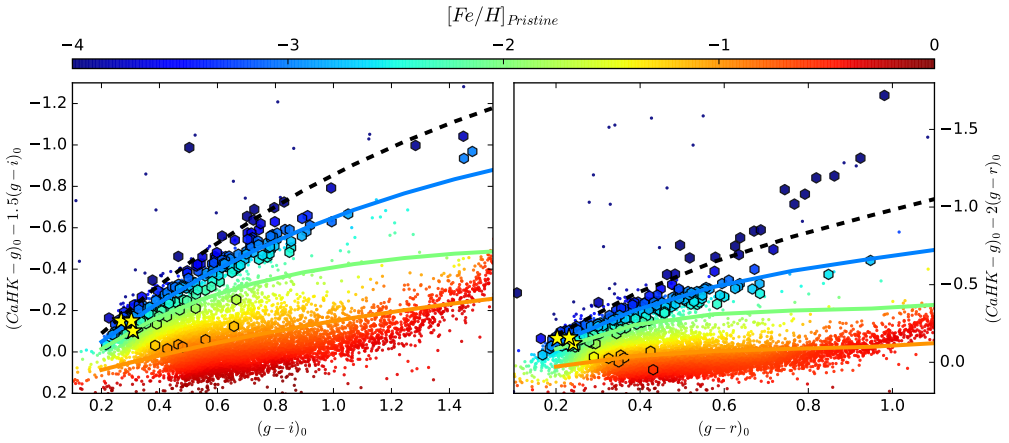

Figure 2.1 — The colour-colour space used to assign photometric metallicities for the *Pristine* sample. The left and right panels show the calibrations for determining photometric metallicities using SDSS $g-i$ and $g-r$ colours, respectively. Stars that are included in the spectroscopic follow-up sample presented in this chapter are shown with large symbols. The coloured lines trace constant metallicities of [Fe/H] = −1, −2, and −3, and the black dashed line represents the expected limit for stars that have no metal absorption lines in their spectra. The small points are 10 000 randomly selected *Pristine* stars, to show the parameter space covered by the survey, and the yellow stars are the 3 stars for which example spectra are shown in Figure 2.2. All symbols are coloured according to their derived photometric metallicities (see text for details).

displays the SDSS $g-i$ or $g-r$ colours, which are proxies for stellar effective temperature. Unless specified otherwise, all magnitudes from SDSS and *Pristine* discussed in the text of the rest of this chapter refer to the deredened magnitudes (see Paper I for details on the deredening procedure). Lines of constant metallicity are also plotted, with orange, green, and blue representing [Fe/H] of −1, −2, and −3, respectively. These lines were produced using synthetic spectra models, generated with Model Atmospheres in Radiative and Convective Scheme (MARCS, Gustafsson et al. 2008) stellar atmospheres and the Turbospectrum code (Alvarez & Plez 1998; Plez 2008). The black dashed line shows the theoretical limit for stars with no metal absorption lines present in their spectra (see Paper I for details). Using information from both these synthetic models and all stars with overlapping *Pristine* photometry and SDSS/SEGUE spectra, this colour-colour space is divided into different photometric metallicity bins. *Pristine* stars are then assigned a metallicity depending on the bin in which they fall, corresponding to their position in this plot. This procedure is followed for both $g-i$ and $g-r$ colours. The minimum metallicity that can be assigned is −4.0, and any object that falls outside of the calibrated regions (approximately the areas shown in Figure 2.1, up to 0.2 dex above the black dashed lines) receives a metallicity of -99. In cases where reliable metallicities are derived for both $g-i$ and $g-r$, the $g-i$ metallicity is preferentially used, since the sample space spans a larger colour range than in $g-r$, and therefore separates the sample more effectively by metallicity over the same range in temperature. We find that the photometric metallicity calibration has a standard deviation of 0.2 dex when compared with the spectroscopic metallicities from SDSS/SEGUE for the metallicity range from [Fe/H] = −0.5 down to [Fe/H] = −3.0 (Paper I).

The small coloured points shown in Figure 2.1 are a random selection of 10 000 *Pristine* stars coloured according to their photometrically derived [Fe/H] values. Large hexagons are the 205

stars from the medium-resolution follow-up sample used in this chapter (see Section 2.3), also coloured by their corresponding *Pristine* photometric metallicities. These stars are almost all selected from the upper regions of the plot between the [Fe/H] = −2 line, the [Fe/H] = −3 line, and the black dashed (no-metals) line, which are the regions expected to contain the most promising EMP star candidates. The stars that lie significantly above the no-metals lines – particularly in the $g-r$ panel – are either stars that were chosen before the full selection criteria described in this chapter were implemented, or are stars that have moved in the plot as a result of improvements to the photometric reduction pipeline and calibration.

2.3 Spectroscopic follow-up

In conjunction with the photometric component of *Pristine* , a spectroscopic follow-up programme has been observing the most promising, bright ($V < 18$) metal-poor candidates on 2 – 4m class telescopes with medium- and high-resolution spectrographs. In this chapter, we focus on the homogeneous follow-up sample of 205 candidate stars observed with the Intermediate Dispersion Spectrograph (IDS) on the 2.5m Isaac Newton Telescope (INT) over the period of March 18 - 27, May 15 - 23, July 20 - 24, and September 2 - 6, 2016, and with the Intermediate dispersion Spectrograph and Imaging System (ISIS) on the 4.2m William Herschel Telescope (WHT) over the period of May 1-2, and July 29-31, 2016 (Programs C71 and N5). Both telescopes are located at the Roque de Los Muchachos Observatory in La Palma, Canary Islands. For the INT, the EEV10 CCD and the R900V grating with a 1.0" slit width were used, resulting in a resolution of 3333 at 4500 Å over 2 pixels at the detector. For the WHT, the R600b and R600R gratings were used, along with the GG495 filter in the red arm. In conjunction with the default dichroic (5 300) and a 1.0" slit, the set-up provided a mean resolution power of 2 400 and 5 200 in the blue and red arms, respectively.

2.3.1 Data reduction and analysis

Spectra were reduced using the Image Reduction and Analysis Facility (IRAF, Tody 1986) software package. All basic reduction steps were implemented, including: image preprocessing (bias subtraction, flat fielding), spectrum extraction, sky subtraction, wavelength calibration, and heliocentric radial velocity correction. Although fringing has been shown to sometimes be a problem for the EEV10 CCD on the INT, the amplitude is lower than 5 % when $\lambda < 6500$ Å[1]. The spectral range used for stars in our sample cover a wavelength range of $\sim 3750-5210$ Å, therefore, it was not necessary to apply a fringing correction.

2.3.2 Spectral analysis

The spectra were analyzed using FERRE[2] (Allende Prieto et al. 2006). We provide here some basic information about the analysis process, but for a detailed account we refer the reader to (Aguado et al. 2017a,b). Both the observed and synthetic spectra were normalized using a running-mean filter 30 pixels wide. The FERRE code fits the entire available spectral region and searches for the atmospheric parameters that best match the observed spectrum by interpolating within the grid. The grid of synthetic spectra used is similar to the one described by Allende Prieto et al. (2014), but with [C/Fe] as a free parameter (Allende Prieto et al. 2015b; Aguado et al. 2016, 2017a,b). This grid has four dimensions and limits of $-6 <$ [Fe/H] < -2, $1 < \log g < 5$, $-1 <$ [C/Fe] < 5, and 4 750K $< T_{\rm eff} <$ 7 000K.

[1] http://www.ing.iac.es/astronomy/instruments/ids/ids_eev10.html
[2] FERRE is available from http://github.com/callendeprieto/ferre

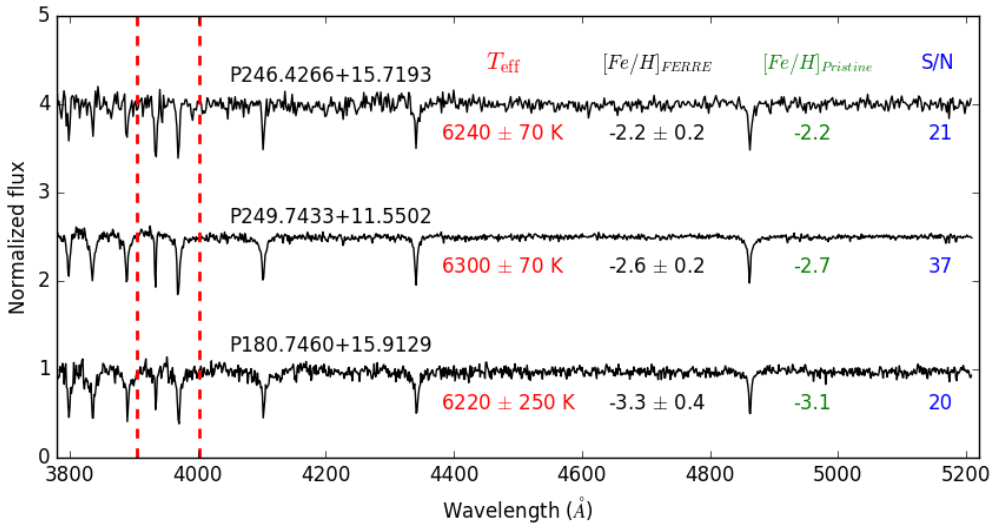

Figure 2.2 — Sample spectra of three *Pristine* target stars with different metallicities but similar temperatures, as determined spectroscopically by FERRE. These stars are marked in the *Pristine* colour-colour space as yellow stars in Figure 2.1. The dotted red lines show the wavelength region (limits at which the transmission falls below 50% of the maximum) for the $CaHK$ filter, and demonstrate the sensitivity of the filter to detecting changes in the strength of the Ca II H & K lines.

To determine the uncertainties, the metallicities are re-derived 50 times after injecting random noise, according to the noise model provided by the data reduction pipeline, and a normal distribution for each instance. The standard deviation of the resulting metallicity distribution is taken as the metallicity uncertainty. Following Aguado et al. (2017b) we add an additional 0.1 dex to the uncertainties to account for other systematic effects.

Figure 2.2 illustrates some typical spectra obtained from the INT, and the relevant wavelength region used for the analysis, $\sim 3750-5210$ Å. The three sample spectra shown were specifically chosen to have similar temperatures, such that the line absorption in the wavelength region targeted by the narrow-band filter can easily be compared. Both the Ca II H (3968.5 Å) and the Ca II K (3933.7 Å) lines are weaker in more metal-poor stars of similar stellar parameters. In relatively warm stars, such as those shown here, the Ca II H line remains somewhat stronger, since it is blended with the H_ϵ line (3970 Å). Therefore, it is particularly the Ca II K line that is a good indicator of whether a star is deficient in all metals, including calcium (e.g., Beers et al. 1999). We note that at this resolution, we cannot typically resolve the interstellar calcium lines form the Ca II H & K lines. However, since any additional blended features only increase the strength of the lines, this will result in stars appearing more metal rich than they actually are, but not more metal-poor.

2.4 Selection criteria

One of the main goals of this chapter is to assess and improve the selection of spectroscopic follow-up stars based on the photometric parameters. Throughout the spectroscopic follow-up, we have developed a specific set of criteria to remove the most contaminants while keeping the completeness as high as possible. These criteria are described in this section.

2.4.1 SDSS photometry

SDSS was chosen as the principal survey to combine to *Pristine* because of its large footprint in the Northern Hemisphere and excellent quality of well-calibrated, deep broad-band photometry. We evaluate the photometric information in several of the SDSS broad-band filters combined with the *Pristine* narrow-band information. The selection criteria that we refined with the spectroscopic sample are described below:

- *Non-star contamination*: Objects that are not stars may exhibit strange spectral signatures that could make them appear to be metal-poor stars from our photometric selection. We therefore identify and remove as many of these sources as possible during the photometric reduction to minimize this source of contamination. The photometry was reduced using the Cambridge Astronomical Survey Unit pipeline (CASU, Irwin & Lewis 2001), and modified to work specifically for CFHT/MegaCam data (Ibata et al. 2014). Objects identified as being stars are flagged as such, and we impose that requirement for objects to be considered for further follow-up. In addition, when matching *Pristine* to SDSS, we only consider sources labelled as stars, thereby providing another means to remove non-point source objects.

- *White dwarf contamination*: Most white dwarfs have very weak CaHK absorption features and therefore could be mistaken for metal-poor stars with the *Pristine* narrow-band filter. Stars at $u-g$ magnitude < 0.6 are likely to be white dwarfs, and are as such easily separated from most main-sequence and giant stars (Ibata et al. 2017a; Lokhorst et al. 2016). We use this colour cut to remove white dwarfs from the sample.

- *Variability*: Since the SDSS $ugriz$ broad-band observations and the *Pristine* narrow-band $CaHK$ observations were taken several years apart, any variable objects could show large variations in brightness between the two data acquisitions, and therefore move significantly in the vertical direction on the colour-colour plot shown in Figure 2.1. This would result in the scattering of non-metal-poor stars into the metal-poor regime and contaminate the sample of stars selected for follow-up. In order to remove these variable objects, the Chi-square variability parameter measured from Pan-STARRS1 photometry was used (Hernitschek et al. 2016), namely that the Pan-STARRS1 variability flag < 0.5. It should be noted that this variability index is only sensitive to brightness variations over the period of the Pan-STARRS1 survey. Since this timescale is shorter than the difference in time between the SDSS and *Pristine* observations, this flag will fail to remove variable objects with periods longer than the Pan-STARRS1 survey. Thus, these objects remain as a source of contamination, although the total number of these in our sample is expected to be quite small.

- *Quality of SDSS i-, r-, and g-band photometry*: We consider SDSS photometric quality flags for saturation, blending, interpolation problems, objects too close to the edge of the frame, or suspicious detections, in each of the g-, r- and i-bands. Since our sample is crossmatched with SDSS and metallicity determinations depend upon the $g-i$ and $g-r$ colours, stars flagged in SDSS as having bad photometry that affects both of these colours must be removed from the sample. We therefore immediately remove all stars which are flagged for bad photometry in the g-band. For the r- and i-bands, we only remove a star if it is flagged as having bad photometry in both bands, because if only one of them is flagged, we may still be able to obtain a reliable photometric metallicity from the other.

- *Probability of a star to have* [Fe/H] ≤ -2.5 *in both $g-r$ and $g-i$*: We choose specifically not to use computed uncertainties in the metallicities because a Monte Carlo estimated

uncertainty probability distribution in metallicity space is distinctly non-Gaussian in shape. Instead, we compute probabilities to reflect the likelihood that a given star has an $[\mathrm{Fe/H}]_{Pristine} \leq -2.5$. To compute these, we take the uncertainties in $CaHK$, g, and i or r photometry, and re-draw these magnitudes in a Monte Carlo fashion for 10^4 instances. For each re-draw, the *Pristine* photometric metallicity is calculated from the fiducial $CaHK$, g, and i or r magnitudes. The probability of a star to have $[\mathrm{Fe/H}]_{Pristine} \leq -2.5$ is subsequently determined by the fraction of the draws for which it gets assigned a photometric metallicity below $[\mathrm{Fe/H}]_{Pristine} = -2.5$. This procedure is done for both the $g-i$ and $g-r$ photometric metallicities. We discard any star for which this probability is less than 0.25 for both $g-i$ and $g-r$.

- *Photometric metallicity grid*: If a star falls outside of the parameter space for which the assignment of photometric metallicities has a valid calibration, it is assigned a metallicity of −99. This is approximately the region contained within Figure 2.1, up to 0.2 dex above the black dashed no-metals line. If a star has a metallicity of −99 for both $g-r$ and $g-i$, it is not considered for follow-up. Similarly, if a star is assigned a metallicity of −99 for only one of $g-r$ or $g-i$, it is removed from the sample if it also has a probability $(\mathrm{[Fe/H]} \leq -2.5) < 0.25$, or bad photometry in the other band.

- *Colour range*: The colour ranges over which *Pristine* most successfully separates stars of different metallicities are $0.25 < g-i < 1.5$ and $0.15 < g-r < 1.2$. These colour ranges correspond roughly to temperatures of 4 200K $< T_{\mathrm{eff}} <$ 6 500K, covering the tip of the red giant branch and the cooler main sequence, all the way to the main sequence turn-off. For hotter stars, the different [Fe/H] populations exhibit more overlap and thus assignment of a metallicity in this regime suffers from larger uncertainty and is more susceptible to contamination by more metal-rich stars. For cooler stars, the main-sequence population at the [Fe/H] = −1 line begins to turn upward and contaminate the more metal-poor red giant regimes. Some stars that fall outside of these colour ranges may still be assigned valid photometric metallicities and may still be interesting targets, but these are followed-up at a lower priority because these regions have a higher contamination rate.

To summarize the selection criteria, a star is removed from the sample if any of the following are true:

(1) $\mathrm{P}([\mathrm{Fe/H}]_{g-r} \leq -2.5) < 0.25$ and $\mathrm{P}([\mathrm{Fe/H}]_{g-i} \leq -2.5) < 0.25$

(2) g-band phot flag

(3) point source flag (CASU flag) $\neq -1$

(4) r-band phot flag and $\mathrm{P}([\mathrm{Fe/H}]_{\mathrm{g-i}} \leq -2.5) < 0.25$

(5) i-band phot flag and $\mathrm{P}([\mathrm{Fe/H}]_{\mathrm{g-r}} \leq -2.5) < 0.25$

(6) $[\mathrm{Fe/H}]_{g-i} = -99$ and $\mathrm{P}([\mathrm{Fe/H}]_{g-r} \leq -2.5) < 0.25$

(7) $[\mathrm{Fe/H}]_{g-r} = -99$ and $\mathrm{P}([\mathrm{Fe/H}]_{g-i} \leq -2.5) < 0.25$

(8) $[\mathrm{Fe/H}]_{g-r} = -99$ and $[\mathrm{Fe/H}]_{g-i} = -99$

(9) i-band phot flag and r-band phot flag

(10) $u-g$ mag < 0.6

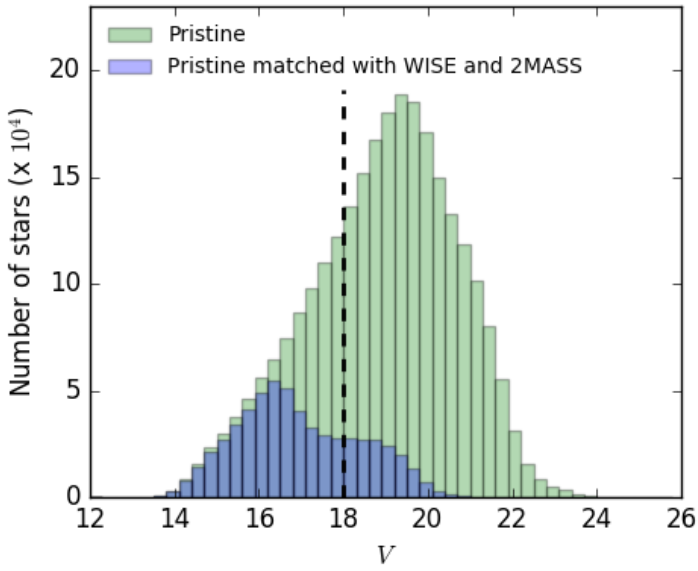

Figure 2.3 — Distribution of V magnitudes[3] for *Pristine* stars (green), and stars that have available WISE and 2MASS magnitudes and satisfy the quality cuts used in SC14 (blue). The black dashed line shows the magnitude limit for the spectroscopic sample in this chapter.

(11) variability > 0.5

2.4.2 Infrared magnitudes from WISE and 2MASS

A study conducted by Schlaufman & Casey (2014, hereafter referred to as SC14) has shown that there is metallicity information contained in the infrared wavelength regions. In their paper, they devise a set of novel selection criteria using the infrared broad-band filters of WISE and 2MASS to select for metal-poor stars. The main selection power of this method comes from the WISE W2 band, centred at 4.6 μ, which contains molecular bands that are strongly metallicity dependent. In this section, we investigate whether the addition of this infrared magnitude information could increase the selection efficiency of *Pristine* .

We applied the selection criteria from SC14 to the *Pristine* photometric sample with $V < 18$, and the analysis revealed two significant limitations on its ability to improve the *Pristine* selection. The first of these is that the WISE and 2MASS magnitudes are only available for a subset of the brightest stars in the *Pristine* sample. Figure 2.3 illustrates the overlap between the total *Pristine* SDSS-matched sample and the stars for which WISE and 2MASS broad-band information is available. To this sample, we have also applied the quality cuts defined in SC14 (only the flag criteria, but not the colour criteria; see their appendix), such that it is a true representation of the subsample of *Pristine* for which this analysis could be performed. When applied to the brightest subset of this sample ($V < 15$), the WISE and 2MASS selection criteria increased the relative number of stars with [Fe/H] < -2.5 from 0.7% to 3.9%, where the metallicities are those derived photometrically from *Pristine* . In the $15 < V < 16$ magnitude bin, the improvement was less pronounced (from 0.7% to 1.7%), and for the fainter magnitude

[3] V magnitudes are calculated using SDSS g and r magnitudes according to the relation described in `https://www.sdss3.org/dr8/algorithms/sdssUBVRITransform.php` (Lupton 2005)

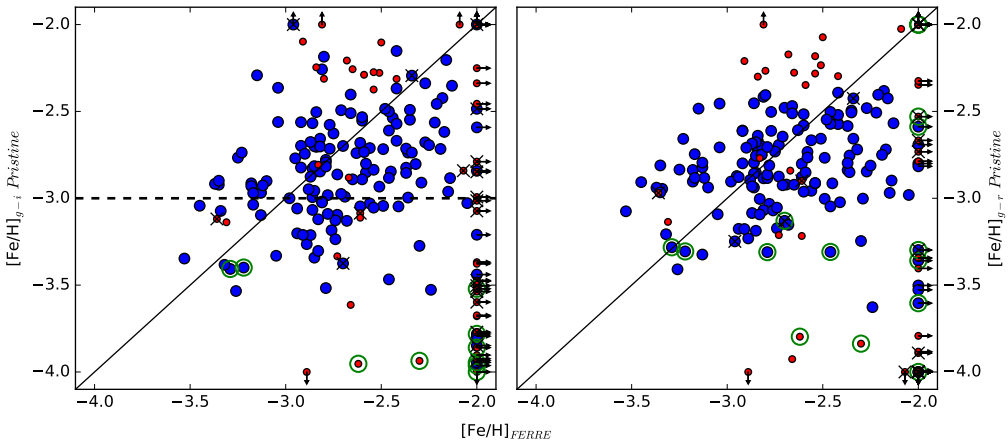

Figure 2.4 — Photometric metallicities for both $g-i$ and $g-r$ plotted against spectroscopic metallicity. The data points are labelled as follows: stars that pass the selection criteria (blue large circles), stars that do not pass the selection (red smaller circles), stars with bad photometry in i for the left panel and r for the right panel (marked with an x), and stars that are above the theoretical no-metals line in Figure 2.1 (circled in green). Data points that fall outside of the plotted region are forced to the border of the plot and marked with arrows showing their true positions.

samples ($V > 16$) this selection power was completely lost. This is to be expected, as the uncertainties in the WISE catalogue for these fainter magnitudes quickly become larger than the range allowed by the selection criteria (i.e. $-0.04 \leq W1 - W2 \leq 0.04$). To account for this, SC14 limit their sample to bright stars with $V < 14$. SDSS photometry is limited to $V \gtrsim 14$ (for a typical star this corresponds roughly to a $CaHK \sim 15$) due to saturation, and because *Pristine* is matched with SDSS, it also inherits this limit. In principle, *Pristine* can observe brighter stars, down to a magnitude of $CaHK \sim 12$, but for these stars the narrow-band information must be used in conjunction with broad-band photometry with a brighter saturation limit than SDSS. This has been successfully demonstrated using APASS, in a recent paper by the *Pristine* collaboration (Caffau et al. 2017). However, even for the bright samples where SC14 provides selection power, it removes a large number of the stars of primary interest with [Fe/H] ≤ -2.5. Therein lies the second major limitation of the SC14 selection criteria and its application to the *Pristine* sample, its low completeness in the metal-poor regime.

For these reasons, we conclude that the WISE and 2MASS selection criteria as implemented by SC14 are quite limited in their application to the *Pristine* sample, and we therefore do not include them in our selection criteria. However, if a large sample of bright, moderately metal-poor candidates are indifferentiable by the *Pristine* selection criteria and follow-up telescope time is limited such that they can not all be observed, then the WISE and 2MASS selection criteria may be useful as a final means to prioritize the sample.

2.4.3 Applying a regularized regression technique

We applied a regularized regression technique, namely Lasso LARS (Tibshirani 1996; Efron et al. 2004), to further assess the need to add other photometric data to predict [Fe/H] metallicities. Such a technique can tell us the leverage of any photometric colour or flux, thereby giving us an independent – and unbiased – view on the most valuable photometric information. It defines a model from a polynomial combination of all the photometric inputs, which also includes

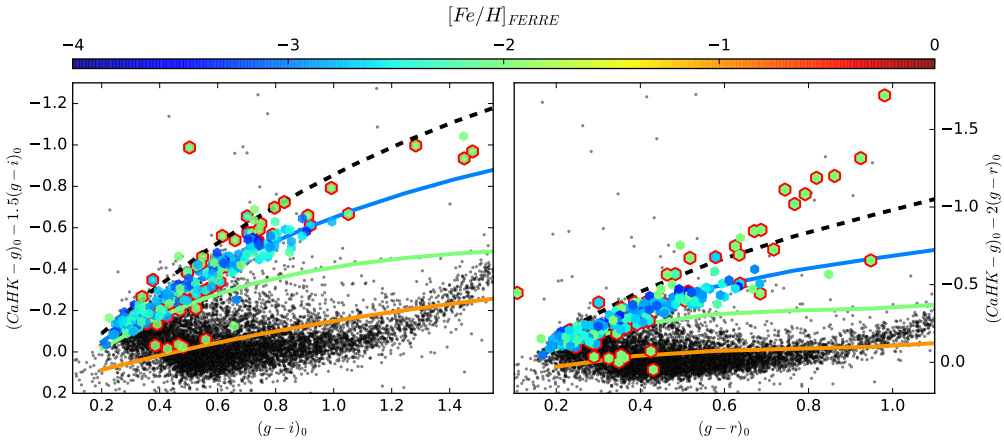

Figure 2.5 — The same colour-colour space as in Figure 2.1, but with stars that have been selected for spectroscopic follow-up coloured according to their spectroscopic $[\mathrm{Fe/H}]_{\mathrm{FERRE}}$. The coloured lines trace along constant metallicities and the black dashed line is the expected limit of stars that have no metal absorption lines in their spectra. The grey points are 10 000 randomly selected *Pristine* stars, to show the parameter space covered by the survey. Data points circled in red represent the stars that are removed from the sample by the selection criteria, many of which are contaminants with $[\mathrm{Fe/H}]_{\mathrm{FERRE}} \geq -2$. These stars circled in red are listed at the bottom of Table 2.1.

colours and cross-terms between the different bands. While doing so, the regularization in this method additionally acts to prefer solutions with fewer parameter values. Such a complete model allows us to effectively explore the importance of various datasets, such as WISE, 2MASS, Spitzer, and SDSS.

During this procedure, we found the $u-g$ colour to be efficient in flagging peculiar objects that could contaminate our sample, which corresponds to our usage of this colour to select out white dwarf contaminants. We also confirmed that adding infrared data such as Spitzer or WISE photometry does not contain significantly independent information from our initial SDSS $ugriz + CaHK$ dataset. Indeed, the metallicity information seems to be mostly contained in the $CaHK-g$, and $g-i$ colour combinations, with extra information in the SDSS u- and r-bands.

At this stage of the Pristine survey, our training set (i.e., the cleaned photometric sample and the SDSS/SEGUE metallicities) does not contain many stars with $[\mathrm{Fe/H}] \leq -2.5$. Therefore, it may be fruitful to repeat this analysis once a larger training set becomes available after additional follow-up spectroscopy.

Taken together, the analysis of the infrared information contained in WISE and 2MASS and an analysis using regularized regression did not result in any changes to the selection criteria described at the end of Section 2.4.1. We therefore proceed with this list of selection criteria to choose stars for future follow-up spectroscopy.

2.5 Spectroscopic results

Figure 2.4 shows the photometrically predicted [Fe/H] with *Pristine* for both $g-i$ and $g-r$ and the spectroscopically determined [Fe/H] with FERRE for all of the stars followed up at the INT and WHT. In the following discussion of the results, we use the terms *Pristine* metallicities

Table 2.1 — Metallicities of Pristine stars from photometry and spectroscopy. Column $CaHK$ is the magnitude obtained from the Pristine narrow-band filter, columns [Fe/H]$_{g-i}$ and [Fe/H]$_{g-r}$ are the photometric metallicities determined using the $g-i$ and $g-r$ colours, respectively. Each is followed by the corresponding derived probability that this metallicity is ≤ -2.5. The next two columns are the spectroscopic metallicities derived form FERRE and their associated uncertainties. Column S/N is the signal to noise ratio of the analyzed spectrum, and column Inst. indicates the instrument used for the observations: either INT/IDS or WHT/ISIS. The last column shows which selection criteria a given star did not pass, and the flags are encoded according to the numbers assigned in the summary list of selection criteria in Section 2.4. We include here only 10 of the 205 stars observed, to show the form of the table. The full table is available online, along with a supplementary table which gives the SDSS coordinates and magnitudes for each of the stars in the sample.

Name	$CaHK$	(±)	[Fe/H] g-i	prob [Fe/H] g-i < -2.5	[Fe/H] g-r	prob [Fe/H] g-r < -2.5	[Fe/H] FERRE	(±)	S/N	Inst.	Flags
Pristine_183.5424+13.6790	15.42	0.02	-3.0	0.98	-3.0	0.99	≥ -2.0	0.2	24	IDS	-
Pristine_184.7471+10.6008	15.97	0.02	-3.4	1.00	-3.5	1.00	≥ -2.0	-	27	IDS	-
Pristine_185.0736+15.1006	15.84	0.02	-2.8	0.97	-2.8	0.97	-2.4	0.2	27	IDS	-
Pristine_185.6263+06.1900	15.46	0.02	-3.3	1.00	-3.2	1.00	-2.9	0.2	31	IDS	-
Pristine_186.5993+15.0468	15.49	0.02	-2.5	0.57	-2.6	0.63	-2.4	0.3	27	IDS	-
Pristine_245.1095+08.8947	14.94	0.02	-2.8	0.96	-99	-0.01	-2.1	0.2	15	IDS	2,5
Pristine_237.5278+12.2989	16.18	0.02	-3.0	1.00	-0.0	0.00	≥ -2.0	-	10	IDS	3,5
Pristine_240.8957+08.4476	16.93	0.02	-3.5	1.00	-99	-0.01	≥ -2.0	-	32	ISIS	3,5
Pristine_182.5908+06.1748	17.28	0.02	-99	-0.01	-99	-0.01	-2.9	0.3	16	IDS	1,3,6,7,8
Pristine_230.9962+07.4789	15.66	0.02	-99	-0.01	-99	-0.01	≥ -2.0	-	43	IDS	1,3,6,7,8

and photometric metallicities synonymously to refer to the metallicity values derived from the narrow-band photometric *Pristine* + SDSS $ugriz$ data, and the terms FERRE metallicities and spectroscopic metallicities to refer to the metallicities derived from analysis of the spectra with FERRE. Only spectra of sufficient quality to be reliably analyzed with FERRE (this was decided visually by the authors, but approximately follows a cut of $S/N = 10$) are included in the sample, which totals 205 stars. The blue large circles represent the 149 stars which pass all of the selection criteria summarized in the list at the end of Section 2.4.1, and the red smaller circles are removed on the basis of at least one of the selection criteria. In both panels, the *Pristine* metallicities are skewed toward the metal-poor end when compared to FERRE. This reflects two characteristics of the sample: 1) stars predicted by photometry to be more metal-poor were preferentially selected for spectroscopic follow-up and 2) because of the shape of the metallicity distribution function, there will be more stars at higher metallicities that will scatter into our photometrically selected sample than the other way around. As a check, we looked to see if there was a correlation between $[\mathrm{Fe/H}]_{Pristine}$ and the computed probabilities of having $[\mathrm{Fe/H}]_{Pristine} \leq -2.5$. Indeed, these showed a tight anti-correlation, which was the expected behaviour, given that the current sample is relatively bright and has small photometric uncertainties.

All objects in Figure 2.4 at $[\mathrm{Fe/H}]_{\mathrm{FERRE}} = -2.0$ should be interpreted as having $[\mathrm{Fe/H}] \geq -2.0$, since the spectral grid used for this analysis was specifically optimized for metal-poor stars and only assigned metallicities in the range $-6 \leq [\mathrm{Fe/H}] \leq -2$. Future work will extend the grid of synthetic spectra to higher metallicities to determine metallicity values for these more metal-rich stars, but for the purposes of this work it is sufficient just to classify them as contaminants. Points marked with an X are flagged with bad photometry in either the i- or r-bands, and points circled in green fall above the no-metals line in Figure 2.1. In both the $g-i$ and $g-r$ panels, most of the stars that fall above the no-metals line are contaminants located at $[\mathrm{Fe/H}]_{\mathrm{FERRE}} \geq -2$. However, a few of the stars that do remain are some of the most metal-poor in the sample, and therefore removing stars based solely on this criteria may be detrimental as it could potentially remove the very rare UMP stars that we are searching for. Dealing with the stars above this line is therefore a matter of completeness versus purity, and given that finding a large number of EMP stars and finding the extremely rare UMP stars are both major objectives of this survey, a choice needs to be made. With the current, small sample of stars that fall in this regime, it is difficult to make a quantitatively driven decision about this matter. Fortunately, many of the stars that fall above the no-metals line are already removed by other selection criteria. We therefore decide not to eliminate the stars that fall above the no-metals line from the sample, in order to mitigate the risk of missing potential UMP stars, but at the cost of a slightly increased contamination rate.

Figure 2.5 shows the same colour-colour space as Figure 2.1, but with the spectroscopic sample coloured by their FERRE metallicities, with stars that do not meet the selection criteria highlighted in red. Again, stars with a light green colour corresponding to $[\mathrm{Fe/H}]_{\mathrm{FERRE}} = -2.0$ actually have $[\mathrm{Fe/H}] \geq -2.0$ and are contaminants. Many of these are successfully removed with the implementation of the selection criteria. Finally, Table 2.1 tabulates the photometric and spectroscopic metallicities for all of the stars in the sample. The stars that do not pass all the selection criteria are listed last, with the rightmost column showing exactly which selection criteria they failed to meet. This table also provides the uncertainties for the spectroscopic metallicities. In this chapter, we report and use only the [Fe/H]values of this sample in order to assess the follow-up success of Pristine. The full sample, as well as determinations of stellar parameters and other abundances are presented in Chapter 3.

It is clear from Table 2.1, Figure 2.4, and Figure 2.5 that there is still some useful information in the *Pristine* photometry, even at these low metallicities of $[\mathrm{Fe/H}] \leq -2.5$. In Figure 2.4,

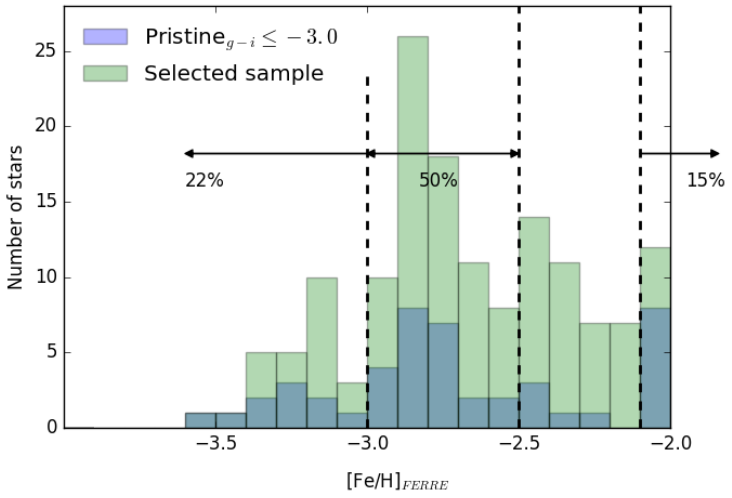

Figure 2.6 — The FERRE metallicity distribution for the selected sample of 149 stars that pass the selection criteria (green), and the metallicity distribution of the 46 stars that have a [Fe/H]$_{Pristine} \leq -3.0$ (blue). The percentages show the fraction of stars from the [Fe/H]$_{Pristine} \leq -3.0$ sample that are contained in the given metallicity ranges.

although there is a scatter around the one-to-one line, the lowest metallicity stars from their spectroscopic metallicities also typically have a lower photometric metallicity determination.

2.6 Discussion

2.6.1 Purity and success rates of the selection

The current spectroscopic sample can be divided into three groups: a total sample of all 205 stars that were observed, a subsample of the 149 stars that pass all of the selection criteria, and a subsample of the 46 stars with [Fe/H]$_{Pristine} \leq -3.0$, which represents the best candidates (all stars below the black-dotted line in the left panel of Figure 2.4). Table 2.2 presents the numbers of stars in various photometric and spectroscopic metallicity bins for these three samples. Firstly, it shows the number of stars in each sample that were predicted by *Pristine* photometric metallicities to have [Fe/H] ≤ -2.5 or ≤ -3.0, respectively. In addition, it provides the same numbers according to the spectroscopic [Fe/H]$_{\rm FERRE}$, and the success rates, which we define as the fraction of stars predicted to be below a certain [Fe/H]$_{Pristine}$ that were actually found to have [Fe/H]$_{\rm FERRE}$ below that value. The selection criteria (as defined in Section 2.4) increase the relative fraction of metal-poor stars in all cases, and eliminate a large number of contaminants with [Fe/H]$_{\rm FERRE} \geq -2$, as compared to the total sample. The [Fe/H]$_{Pristine} \leq -3.0$ sample increases the the relative fractions even more, but concedes a higher contamination rate than the sample which passes the selection criteria.

Figure 2.6 shows the FERRE metallicity distribution for the sample that satisfies the selection criteria. To visualize the success of the selection based on photometry, we plot in blue the metallicity distribution of the [Fe/H]$_{Pristine} \leq -3.0$ sample. The percentage of stars that fall in each region are shown, namely that 22% of these stars still end up below [Fe/H] ≤ -3.0, 50% fall between $-3.0 <$ [Fe/H] ≤ -2.5, and 15% are contaminants with [Fe/H] ≥ -2.0.

Table 2.2 — Numbers of stars with photometric predictions $[Fe/H]_{Pristine}$ below −2.5 and −3.0, the numbers of stars that are spectroscopically confirmed below those metallicities, and the success rates, given for the full spectroscopic sample, the sample after application of the selection criteria (described in Section 2.4), and the sample of stars with $[Fe/H]_{Pristine} \leq -3.0$.

	Total observed	Selection criteria	$[Fe/H]_{Pristine} \leq -3.0$
Total number	205	149	46
$[Fe/H]_{Pristine} \leq -2.5$	163/205 (80%)	130/149 (87%)	46/46 (100%)
$[Fe/H]_{Pristine} \leq -3.0$	73/205 (36%)	46/149 (31%)	46/46 (100%)
$[Fe/H]_{FERRE \leq -2.5}$	119/205 (58%)	98/149 (66%)	33/46 (72%)
$[Fe/H]_{FERRE \leq -3.0}$	27/205 (13%)	25/149 (17%)	10/46 (22%)
$[Fe/H]_{FERRE \geq -2.0}$	42/205 (20%)	11/149 (7%)	7/46 (15%)
success $[Fe/H] \leq -2.5$	101/163 (62%)	91/130 (70%)	-
success $[Fe/H] \leq -3.0$	12/73 (16%)	10/46 (22%)	10/46 (22%)

Over the *Pristine* footprint, covering ~ 1 000 deg^2 as of September 2016, we have photometrically identified 10 243 metal-poor star candidates with $V < 18$ that pass all of the selection criteria laid out in Section 2.4. The *Pristine* survey does go deeper than this ($V \sim 20$), but this is the magnitude range accessible with 2 − 4m class telescopes. The selected sample constitutes 1.3% of all of the stars in the survey present in this magnitude range, and more than half of these have a predicted photometric metallicity $[Fe/H]_{Pristine} \leq -2.5$. Table 2.3 summarizes the number of candidate stars split into magnitude ranges, where the first number given for each entry is the number of stars followed up and the second is the number of candidates in the full sample.

Although the success rates reported in this chapter are based on a small sample of stars, we can still estimate the number of stars with $[Fe/H] \leq -3.0$ that we would expect to find in the entire $V < 18$ sample. However, since we have selected the best candidates available first, (i.e., we have observed a higher fraction of stars with $[Fe/H]_{Pristine} \leq -3.0$), we cannot directly scale the number of EMP stars found in our 205 star sub-sample to the number expected for the full sample. We therefore separate the sample into ranges of photometric metallicities, compute the relative fraction of EMP stars recovered in each metallicity range, and then scale these numbers to the total candidate sample. This calculation yields an expected number of ~ 1 000 − 1 200 EMP stars over the ~ 1 000 deg^2 *Pristine* footprint (the final number is somewhat dependant on the bin size chosen for the metallicity ranges). Considering all observed stars, we therefore estimate a frequency of ~ 1/800 (1.25%) for stars to have a metallicity of $[Fe/H] \leq -3.0$ for $14 < V < 18$ in the Galactic halo.

Based on the Besançon Model of stellar population synthesis of the Galaxy (Robin et al. 2003) – for a similar sky region to the *Pristine* footprint and a magnitude range of $14 < V < 18$ – we expect a frequency of 1/2 000 (0.05%) for randomly selected halo stars to have $[Fe/H] \leq -3.0$. It should be noted that this is only a first order approximation, as the model relies on several assumptions about the metal-poor tail of the halo metallically distribution function. It should also be noted that our projections have been made based on a small sample of stars that preferentially occupy the brighter part of this magnitude range. However, as a coarse comparison, frequencies of expected EMP stars from simulated galaxy model predictions are in reasonable agreement with our observations.

Table 2.4 summarizes the comparisons of the relative return for HES, SC14 and *Pristine* . Other efforts have yielded similar or lower return rates as HES and SC14 (e.g., Allende Prieto et al. 2000).

*These percentages are computed from the scaled sample presented in Table 3 of Schörck et al. (2009).

Table 2.3 — Number of candidate stars in different magnitude bins and metallicity ranges. The first number in each cell is the number of stars followed up with spectroscopy from the sample in this chapter, and the second is the total number of candidates as of September 2016 over the ∼ 1 000 deg^2 *Pristine* survey footprint. [Fe/H] values shown are photometric *Pristine* $g-i$ metallicities.

	# Candidates	[Fe/H]≤-2.5	[Fe/H]≤-3.0
$V < 15$	47/213	30/166	13/48
$15 < V < 16$	114/797	91/554	33/92
$16 < V < 17$	29/2 388	28/1 549	17/242
$17 < V < 18$	15/6 845	14/4 354	10/674
Total	205/10 243	163/6 623	73/1 056

Table 2.4 — The relative fractions of metal-poor stars in *Pristine* compared to other surveys.

Survey	[Fe/H] <-3	[Fe/H] <-2.5	-3 <[Fe/H] <-2
Pristine	17%	66%	76%
HES	4%	22%*	40%*
SC14	3.8%	-	32%

2.6.2 Comparison to other surveys

In order to compare these results to other surveys, we use the relative fractions of metal-poor stars from the selected sample of 149 stars. Although the success rates are a more telling quantification of the capabilities of *Pristine* for finding EMP stars, it is more appropriate to use the relative fractions for a quantitative comparison to other works. This is because the *Pristine* survey has the advantage over other metal-poor star searches that it can quantify the metallicity of its candidates, and select for example candidates with $[\mathrm{Fe/H}]_{Pristine} \leq -3.0$, instead of labelling objects in a binary fashion as EMP candidates or not.

SC14 report that $3.8^{+1.3}_{-1.1}$% of their candidate stars have an [Fe/H] $\lesssim -3.0$, and $32^{+3.0}_{-2.9}$% have $-3.0 \lesssim$ [Fe/H] $\lesssim -2.0$, from high-resolution follow-up of their selection with WISE and 2MASS magnitudes. Although we report significantly higher rates of 17% and 76%, respectively, it should be taken into consideration that they are using publicly available survey data and are specifically targeting bright stars, so they enjoy the advantage of large sky coverage and ease of spectroscopic follow-up. They also use near infrared magnitudes, which offers the advantage of being able to probe the crowded regions of the disc in the direction of the bulge (Casey & Schlaufman 2015), although they are limited in the distance they can reach due to the bright nature of their sample.

The stellar content and metallicity distributions of HES are presented in Schörck et al. (2009). In that paper, they report a fraction of stars with [Fe/H] ≤ -3.0 of 7% for their best-selected sample, and $3-4$% for the other samples. Their best-selected sample totals 105 out of 1 638 stars, and constitutes only 6.4% of their total accepted follow-up sample. This sample can be compared to the 22% success rate of the best *Pristine* sample, stars with an assigned photometric metallicity of [Fe/H] ≤ -3.0. Taking the entire HES sample as a whole then yields 65 out of 1 638 stars with [Fe/H] ≤ -3.0, a fraction of 4%, and this can be compared to the relative fraction from the whole *Pristine* sample of 17%.

2.6.3 Future follow-up strategy

Schörck et al. (2009) report that for the bias-corrected HES metallicity distribution function, around 1-3% of all [Fe/H] ≤ -3.0 stars had a metallicity [Fe/H] ≤ -4.0. Allende Prieto et al.

(2014) report similar numbers for SDSS/BOSS, with 1 star at [Fe/H] ≤ -4.0 out of 118 at [Fe/H] ≤ -3.0 (see their Table 2).

We can use these statistics to make projections of how many UMP stars we expect to find. Taking a conservative estimate, we expect one star with [Fe/H] ≤ -4.0 for every $\sim$ 100 stars with [Fe/H] ≤ -3.0. We therefore predict that we will find $\sim 10-12$ UMP stars over the $\sim$ 1 000 deg^2 footprint in the magnitude range $V < 18$, given that we will uncover a projected $\sim$ 1 000 - 1 200 stars with [Fe/H] ≤ -3.0. Furthermore, it is not surprising that we have not yet found any UMP stars in our current sample of 205 stars (27 with [Fe/H] ≤ -3.0). This sample is still quite small when compared to other surveys that have successfully found UMP stars, such as SDSS (with SEGUE and BOSS), and HES, which have both followed up many thousands of stars with low-resolution spectroscopy.

Projecting forward into the future, with a larger footprint of $\sim$ 3 000 deg^2, we expect to find a statistical sample of several tens of these stars. Furthermore, if we can follow-up the fainter magnitude range of *Pristine* ($18 < V < 20$), this not only would provide many more candidates, but also probe deeper into the halo and potentially result in the discovery of many more UMP stars.

Given the availability of time on $2-4$m class telescopes, it may be possible for our team to obtain a complete follow-up sample for the brighter magnitude ranges of our candidate sample, up to $V < 16$. For the magnitude ranges fainter than this, there are too many candidates to feasibly follow up with single slit spectrographs. However, this task would be well-suited to the upcoming new generations of multi-object spectrographs, such as WEAVE, 4MOST, PFS, and MSE.

2.7 Conclusions

Through an analysis of the first medium-resolution spectroscopic sample from the follow-up programme of *Pristine* , we have demonstrated that the narrow-band survey is very efficient at uncovering EMP stars in the Galactic halo. We used this sample to assess and refine the selection criteria for selecting photometric candidates for spectroscopic follow-up. This included investigating whether infrared magnitudes from WISE and 2MASS could improve the selection efficiency, as was done by Schlaufman & Casey (2014), but this added information was only useful for the brightest Pristine stars ($V < 15$) and even then resulted in low completeness in the metal-poor regime of [Fe/H] ≤ -2.5. Analyzing the selection criteria with a regularized regression technique, we confirmed that the u, g, r, i, and $CaHK$ magnitudes contain the most useful information for separating the sample by metallicity.

The total spectroscopic sample consisted of 205 stars, of which 27 were found to have [Fe/H] ≤ -3.0 and 119 were found to have [Fe/H] ≤ -2.5. This sample was reduced to 149 stars by the refined photometric selection criteria, of which 25 had [Fe/H] ≤ -3.0 (17%) and 98 were found to have [Fe/H] ≤ -2.5 (66%). This return rate for finding EMP stars is unprecedented, with other surveys typically reporting values of 3-4%. For stars predicted by *Pristine* to be EMP with [Fe/H] ≤ -3.0, we report a success rate of 22% for confirming them as EMP, and for stars predicted to have a metallicity of [Fe/H] ≤ -2.5 we report a success rate of 70%.

The *Pristine* survey is ongoing, both with increasing sky coverage of the photometric footprint with CFHT/Megacam and with its spectroscopic follow-up campaign. Based on our statistics, we expect to uncover $\sim$ 1 000 $-$ 1 200 stars with [Fe/H] ≤ -3.0 and $\sim 10-12$ stars with [Fe/H] ≤ -4.0 per $\sim$ 1 000 deg^2 of survey area. In the future, we hope to expand our spectroscopic follow-up towards fainter magnitudes with the next generation of multi-object spectrographs.

Acknowledgements

We would like to thank the referee for their useful comments which helped to improve the chapter. KY would also like to thank Gal Matijevič and Jenn Wojno for their comments and insightful discussions. We gratefully thank the CFHT staff for performing the observations in queue mode, for their reactivity in adapting the schedule, and for answering our questions during the data-reduction process. We also thank the support astronomers and staff at the INT/WHT for their expertise and help with observations. We thank Nina Hernitschek for granting us access to the catalogue of PanSTARRS1 variability catalogue. ES and KY gratefully acknowledge funding by the Emmy Noether program from the Deutsche Forschungsgemeinschaft (DFG). NFM gratefully acknowledges funding from CNRS/INSU through the Programme National Galaxies et Cosmologie, and through the CNRS grant PICS07708. ES and KY benefited from the International Space Science Institute (ISSI) in Bern, CH, thanks to the funding of the Team "The Formation and Evolution of the Galactic Halo". DA acknowledges the Spanish Ministry of Economy and Competitiveness (MINECO) for the financial support received in the form of a Severo-Ochoa PhD fellowship, within the Severo-Ochoa International PhD Program. DA, CAP, and JIGH also acknowledge the Spanish ministry project MINECO AYA2014-56359-P. JIGH acknowledges financial support from the Spanish Ministry of Economy and Competitiveness (MINECO) under the 2013 Ramón y Cajal program MINECO RYC-2013-14875.

This chapter is based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Science de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii.

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

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

3 The first three years of medium resolution follow-up spectroscopy of *Pristine* metal-poor star candidates

D. S. Aguado, K. Youakim, J. I. González Hernández, C. Allende Prieto, E. Starkenburg, N. Martin, P. Bonifacio, A. Arentsen, E. Caffau, L. Peralta de Arriba, F. Sestito, R. Garcia-Diaz, N. Fantin, V. Hill, P. Jablonca, F. Jahandar, C. Kielty, N. Longeard, R. Lucchesi, R. Sánchez-Janssen, Y. Osorio, P. A. Palicio, E. Tolstoy, T. G. Wilson, P. Côté, G. Kordopatis, C. Lardo, J. F. Navarro, G. F. Thomas, and K. Venn

— Published as D.S. Aguado, K. Youakim et al., 2019, MNRAS 490, 2241 —

ABSTRACT

We present the results of a three year long, medium-resolution spectroscopic campaign aimed at identifying very metal-poor stars from candidates selected with the $CaHK$, metallicity-sensitive Pristine survey. The catalogue consists of a total of 1 007 stars, and includes 146 rediscoveries of metal-poor stars already presented in previous surveys, 707 new very metal-poor stars with [Fe/H] < -2.0, and 95 new extremely metal-poor stars with [Fe/H] < -3.0. We provide a spectroscopic [Fe/H] for every star in the catalogue, and [C/Fe] measurements for a subset of the stars (10% with [Fe/H] < -3 and 24% with $-3 <$ [Fe/H] < -2) for which a carbon determination is possible, contingent mainly on the carbon abundance, effective temperature and S/N of the stellar spectra. We find an average carbon enhancement fraction ([C/Fe] $\geq +0.7$) of 41 $\pm$ 4% for stars with $-3 <$ [Fe/H] < -2 and 58 $\pm$ 14% for stars with [Fe/H] < -3, and report updated success rates for the *Pristine* survey of 56% and 23% to recover stars with [Fe/H] < -2.5 and [Fe/H] < -3, respectively. Finally, we discuss the current status of the survey and its preparation for providing targets to upcoming multi-object spectroscopic surveys such as WEAVE.

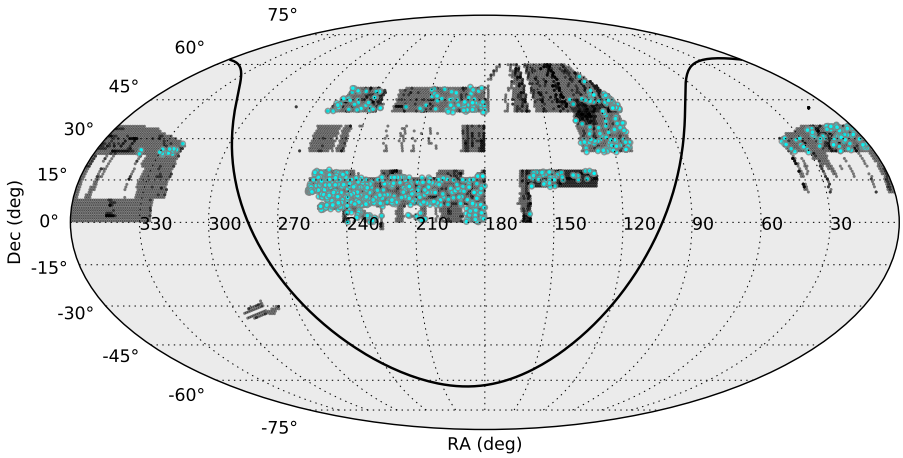

Figure 3.1 — The current footprint of the *Pristine* survey (black) covering ~ 5 000 deg^2. The stars making up the spectroscopic follow-up sample are plotted as cyan points and were selected from ~ 2 500 deg^2 of the total region. The Galactic plane is shown as the black line.

3.1 Introduction

The current picture of Galactic chemical enrichment is based on the production of elements heavier than He in the interiors of stars, their subsequent release into the interstellar medium through supernova explosions, and their eventual reintegration into ensuing stellar generations. Apart from a few exceptions, such as mass transfer binaries, the current elemental compositions of stars are expected to maintain the chemical imprint of their birth environments, which in turn reflect this enrichment process. Based on this principle, it is possible to use stars with primitive elemental abundance patterns, also known as very metal-poor (VMP: [Fe/H] < -2), to study the early Universe.

One issue that hampers our ability to study the detailed abundance trends of metal-poor stars, is their scarcity in our local environment with respect to the younger, more metal-rich populations. However, metal-poor stars are more abundant in certain Galactic environments, making them promising searching grounds. Cosmological simulations demonstrate that the outer regions of the Galaxy are the most dominated by old and/or metal-poor stars (for recent studies using hydrodynamical simulations see Starkenburg et al. 2017b and El-Badry et al. 2018). If one has a good method to efficiently distinguish metal-poor from more metal-rich populations, and is interested in stars that are both metal-poor and old, then the Galaxy's inner regions and some of its satellites are also promising hunting grounds (e.g., White & Springel 2000; Tumlinson 2010; Starkenburg et al. 2017b).

Naturally, a substantial amount of effort has gone into finding and studying these rare stars, and they remain a strong focus of current and future surveys dedicated to Galactic archaeology. Given that they are so rare among the far more numerous foreground populations, there are two options when searching for metal-poor stars: (i)observing a large sample of stars from general science purpose surveys to find the few metal-poor stars among them, or (ii)targeted searches which aim for these stars specifically. The former approach has been quite successful and has contributed significantly to the current sample of the most metal-poor stars (e.g. Caffau et al. 2013a; Aoki et al. 2013; Allende Prieto et al. 2015a; Aguado et al. 2016; Li et al. 2015; Aguado

Table 3.1 — Technical Information for Facilities Used in this Analysis

Instr.	Telescope	Detector	Grating	Dispersion	Range	$\lambda/\Delta\lambda$ @ 4500Å	Slit
IDS	2.5-m INT	EEV10	R900V	0.69 Å px^{-1}	3600-5200 Å	3300	1.″0
ISIS	4.2-m WHT	EEV12	R600B	0.45 Å px^{-1}	3600-5100 Å	2400	1.″0
EFOSC2	3.6-m NTT	CCD40	600	0.95 Å px^{-1}	3600-5200 Å	930	1.″0

et al. 2017a,b, 2018a,b), mostly with the help of large spectroscopic surveys such as the Sloan Digital Sky Survey (SDSS, York et al. 2000), the Sloan Extension for Galactic Understanding and Exploration (SEGUE, Yanny et al. 2009), the Baryonic Oscillations Spectroscopic Survey (BOSS, Eisenstein et al. 2011; Dawson et al. 2013), and more recently the large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST, Deng et al. 2012). More targeted searches have also been in use for many years, from early efforts using a Ca H & K objective-prism technique, such as the HK survey (Beers et al. 1985, 1992) and the Hamburg ESO survey (Christlieb et al. 2002), to more recent efforts using targeted narow/medium-band photometry at blue wavelengths, like the *SkyMapper* survey (Keller et al. 2007; Wolf et al. 2018; Casagrande et al. 2019; Huang et al. 2019), and the *Pristine* survey (Starkenburg et al. 2017).

Future metal-poor star searches will be even more effective by combining both of these strategies. The upcoming generation of multi-object spectroscopic (MOS) surveys such as the William Herschel Telescope Enhanced Area Velocity Explorer (Dalton et al. 2018, WEAVE), the Dark Energy Spectroscopic Instrument (Levi et al. 2013, DESI), the 4-metre Multi-Object Spectroscopic Telescope (de Jong et al. 2019, 4MOST), the Galactic archaeology with HERMES (Buder et al. 2018, GALAH), the Sloan Sky Digital Survey-V (Kollmeier et al. 2017, SDSS-V), and the Maunakea Spectroscopic Explorer (McConnachie et al. 2016a; McConnachie 2019, MSE) will provide, together with Gaia (Gaia Collaboration et al. 2018), an unprecedented number of spectra over the whole sky. Although these surveys will have the capability to observe tens of millions of stars, it will still be necessary to target metal-poor stars specifically in order to maximize the output for Galactic archaeology studies. When used in tandem with pre-selection surveys such as *SkyMapper*, and *Pristine*, it will be possible to obtain high quality observations of metal-poor stars across an unprecedented range of magnitudes, wavelengths, and Galactic environments. In order for the target pre-selection from such surveys to be maximally effective, they must be validated beforehand by dedicated spectroscopic follow-up programs.

In this chapter, we present the results of the first three years of spectroscopic follow-up for the *Pristine* survey, using low- and medium-resolution spectroscopic facilities. This not only provides a detailed understanding of the selection of candidates to target with future MOS surveys, but also has the added value of providing the Galactic archaeology community with a sizeable catalogue of new, metal-poor stars, a subset of which also have measurements of carbon abundances.

Carbon is a well studied element in metal-poor stars, and has important implications for understanding the earliest stellar generations. First, the carbon abundance of a star influences the cooling channels and may allow for low-mass star formation (Bromm & Loeb 2003). Secondly, as discussed at length in Beers et al. (2005), Yong et al. (2013a), Bonifacio et al. (2015), Yoon et al. (2016), and Chiaki & Wise (2019), the increase in carbon-enhancement with decreasing metallicity in EMP stars allows us to make a phenomenological taxonomy of ancient stars.

There are two definitions for carbon-enhanced metal-poor (CEMP) stars currently presented in the literature. Beers et al. (2005) propose a definition of CEMP stars as stars with [C/Fe] > +1.0, while Aoki et al. (2007) use [C/Fe] > +0.7 with an additional correction

depending on the luminosity*. These different values do not reflect theoretical studies but still provide a useful quantitative classification. On the other hand, the original critical carbon abundance from Bromm & Loeb (2003) ($[C/H]_{crit} \simeq -3.5 \pm 0.1$) has recently been improved to include the effect of the silicate grains in cooling processes allowing for fragmentation of the proto-stellar clouds (Chiaki et al. 2017). These studies propose three regions in the A(C) − [Fe/H] plane: the carbon dominated area, the silicate dominated area, and an area that is forbidden due to insufficient dust cooling. So far only one star, J1029+1729, belonging to the [Fe/H] < −4.5 regime is clearly carbon-normal (Caffau et al. 2011a) with [C/Fe] < +0.7. J1029+1729 is still the most metal-poor star known but remains in the silicate dominated region well below [C/Fe] = +2.3 line. Discovered by Starkenburg et al. (2018) and included in this work, Pristine_221.8781+9.7844 is the second most metal-poor star also in the silicate dominated region with [C/Fe] < +1.76 and could also potentially be a carbon-normal UMP star. All 11 other stars from the literature with [Fe/H] < −4.5 show a clear enhancement in carbon (see for example Bonifacio et al. 2018b; Yoon et al. 2019, and references therein). Larger samples of extremely metal-poor stars, especially those with robust carbon measurements, are important in order to better understand these trends.

The chapter is organized as follows. In Section 3.2, we summarize the data set, observations, and reduction methods. In Section 3.3, we explain the analysis of the data using the FERRE code. In section 3.4, we present the spectroscopic follow-up catalogue, including a discussion of the updated success rates for finding EMP and VMP stars of the *Pristine* survey. In Section 3.5, we look at the future of *Pristine* and its synergies with other upcoming surveys, and we conclude the chapter in Section 3.6.

3.2 Data and Observations

As discussed in detail in Starkenburg et al. (2017), one of the main aims of the *Pristine* project is to enlarge the number of metal-poor stars currently known in our Galaxy and characterize them to better understand the Galactic halo. Figure 3.1 shows the current *Pristine* footprint which covers a total of ~ 5 000 deg^2 in the Northern Galactic halo. The targets selected for follow-up spectroscopy are shown in cyan, and were selected form a ~ 2 500 deg^2 region of the total footprint.

3.2.1 Observations

The spectroscopic data presented here were collected over a period of 6 semesters, from March 2016 to February 2019. Figure 3.2 shows the V-band magnitude distribution of the spectroscopic follow-up sample, totalling 1 007 stars. Due to the wide range in target brightness, three different facilities were used to conduct follow-up observations of EMP candidates selected from the *Pristine* survey: the Intermediate Dispersion Spectrograph (IDS) on the 2.5-m Isaac Newton Telescope (INT), the Intermediate-dispersion Spectrograph and Imaging System (ISIS, Jorden 1990) on the 4.2-m William Herschel Telescope, and the ESO Faint Object Spectrograph and Camera (EFOSC2, Buzzoni et al. 1984) on the 3.6-m New Technology Telescope (NTT). The selected mode in all cases was *long slit* providing low- and medium-resolution spectroscopy (see Table 3.1 for further technical details).

Fainter targets ($g > 16.2$) were observed with the larger aperture WHT and NTT telescopes, while brighter targets ($g < 16.2$) were observed with the INT. The total number of observing nights were 182 (145 with IDS, 25 with ISIS and 12 with EFOSC). Although the ISIS observations

*More recently Bonifacio et al. (2018b) proposed to establish a fixed A(C) > 5.5 reference value for stars with [Fe/H] < −4.0 to be CEMP stars.

Table 3.2 — FERRE analysis for a sample of well known EMP stars. Uncertainties include both systematic and statistical errors. Values from the literature derived from high-resolution analyses are also shown.

	Values from FERRE					Values from the literature			
Object	T_{eff} [K]	log g	[Fe/H]	S/N	Inst.	T_{eff} [K]	log g	[Fe/H]	Ref.
HE 0057−5959	5333 ± 118	1.72 ± 0.72	-3.69 ± 0.27	39	EFOSC	5257 ± 100	1.72 ± 0.30	-4.08 ± 0.30	1
SDSS J0723+3637	5258 ± 212	2.70 ± 1.33	-3.41 ± 0.21	28	IDS	5150 ± 150	2.20 ± 0.50	-3.32 ± 0.20	2
HD 84937	6379 ± 109	4.75 ± 0.50	-2.19 ± 0.21	181	IDS	6431 ± 100	4.08 ± 0.30	-2.14 ± 0.20	3
SDSS J1004+3442	6002 ± 140	2.84 ± 0.95	-2.83 ± 0.25	13	IDS	6100 ± 150	4.00 ± 0.50	-3.09 ± 0.20	2
SDSS J1036+1212	6052 ± 102	1.26 ± 0.50	-3.24 ± 0.21	34	IDS	5850 ± 150	4.00 ± 0.50	-3.47 ± 0.20	2
SDSS J1108+1747	5930 ± 104	4.89 ± 0.50	-3.07 ± 0.21	35	IDS	6050 ± 150	4.00 ± 0.50	-3.17 ± 0.20	2
SDSS J1128+3841	6416 ± 126	4.61 ± 0.61	-3.28 ± 0.22	39	IDS	6550 ± 150	4.00 ± 0.50	-2.82 ± 0.20	2
HE 1207−3108	5545 ± 156	3.11 ± 0.87	-3.01 ± 0.22	93	EFOSC	5294 ± 100	2.85 ± 0.30	-2.70 ± 0.30	1
HE 1320−2952	5658 ± 123	4.09 ± 0.59	-3.13 ± 0.22	50	EFOSC	5106 ± 100	2.26 ± 0.30	-3.69 ± 0.30	1
HE 1327−2326	6400 ± 109	4.82 ± 0.50	-5.40 ± 0.43	30	IDS	6180 ± 80	4.50 ± 0.50	-5.70 ± 0.20	4
G64−12	6435 ± 105	4.97 ± 0.50	-3.24 ± 0.22	80	IDS	6550 ± 100	4.68 ± 0.30	-3.21 ± 0.20	3
CS 30336−0049	5194 ± 161	2.60 ± 1.14	-3.97 ± 0.22	51	EFOSC	4725 ± 100	1.19 ± 0.30	-4.10 ± 0.30	1
HE 2047−5612	6281 ± 122	4.64 ± 0.55	-2.94 ± 0.22	41	EFOSC	6128 ± 100	3.68 ± 0.30	-3.14 ± 0.30	1
SDSS J2206−0925	5210 ± 100	1.01 ± 0.50	-2.66 ± 0.20	29	IDS	5100 ± 150	2.10 ± 0.50	-3.17 ± 0.20	2
BD+17 4708	6100 ± 106	3.90 ± 0.50	-1.80 ± 0.21	120	IDS	6085 ± 50	4.10 ± 0.10	-1.60 ± 0.10	5
SDSS J2338−0902	5052 ± 101	1.03 ± 0.50	-2.62 ± 0.20	32	IDS	4900 ± 150	1.90 ± 0.50	-3.12 ± 0.20	2

References: 1=Yong et al. (2013a); 2=Aoki et al. (2013); 3=Ishigaki et al. (2012); 4=Frebel et al. (2005); 5=Gratton et al. (2003)

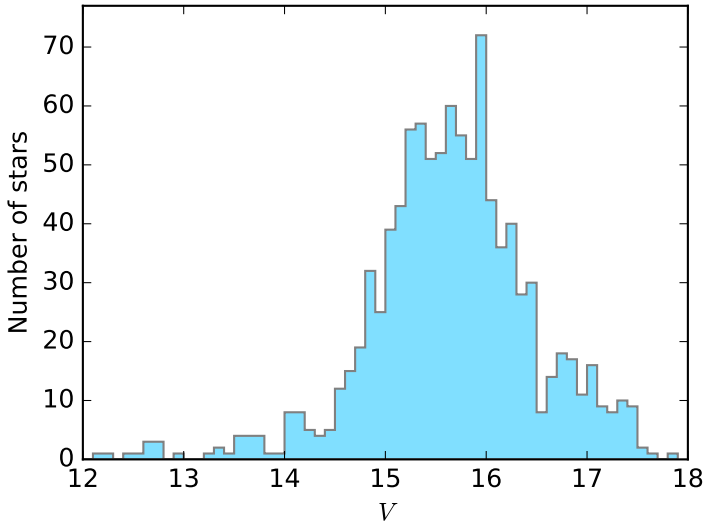

Figure 3.2 — Distribution of V magnitudes for the full follow-up spectroscopic sample of 1 007 *Pristine* stars.

were shared with another program so that the resulting equivalent observing nights came out to ~ 10.

3.2.2 Observational strategy

The minimum desired signal-to-noise (S/N) ratio per pixel for the observations was ~ 15 − 25 in the calcium H & K spectral region (~ 3950 Å), depending on the effective temperature of a given star. Therefore, the average exposure time for a single integration was 1 500 s, 900 s and 1 500 s, for the INT, ISIS and EFOSC observations, respectively. Naturally, exposure times varied slightly for each individual object depending on the target brightness and the visibility conditions. We designed the observational strategy to maximize the ratio between the number of observed candidates and the reliability of the derived parameters. However, stars that were identified as ultra metal-poor (UMP) candidates during an observing run were subsequently followed-up with more exposures to achieve a higher S/N. Stars that still seemed highly interesting at this stage were then followed up with larger telescopes at higher resolution. This observing strategy was designed to maximize the detection of very low-metallicity stars, and has yielded the discovery of Pristine_221.8781+9.7844, an ultra metal-poor sub-giant star with [Fe/H] = -4.66 ± 0.13 and [C/Fe]< 1.76. The detailed analysis of this star with high resolution follow-up with VLT/UVES is described in Starkenburg et al. (2018).

3.2.3 Data reduction

The spectral data reduction included bias substraction, flat-fielding, and wavelength calibration – using CuNe+CuAr lamps for IDS and ISIS, and He+Ar for EFOSC–, and was performed using the ONESPEC package in IRAF (Tody 1993). At the moderate S/N levels required for this program and at medium-resolution, the contribution of the interstellar medium (ISM) in the Ca H & K area is, in general, not resolved (see for example Aguado et al. 2016, 2017a). In order to

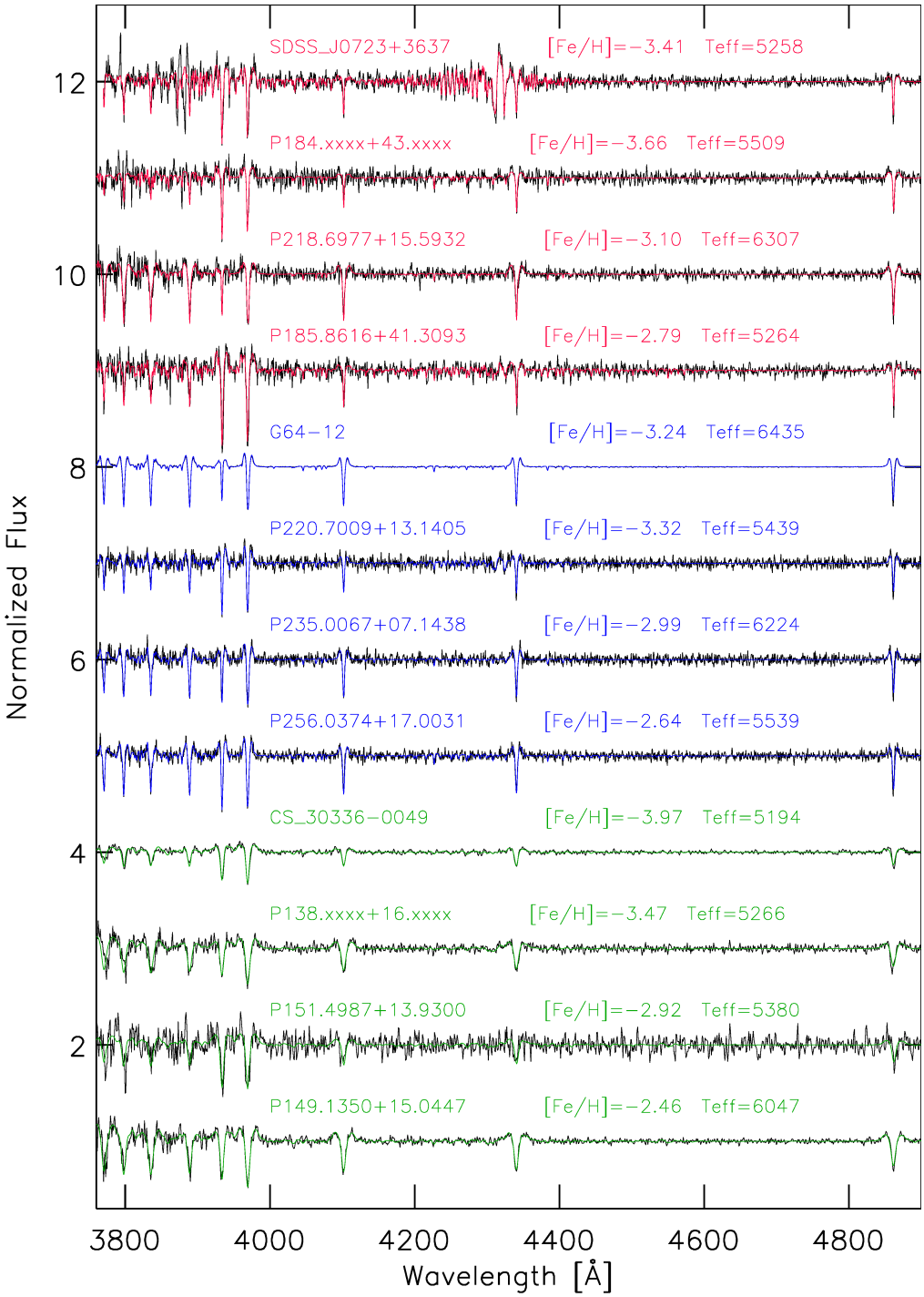

Figure 3.3 — A subsample of the spectra of the *Pristine* candidates observed with IDS/INT (red), ISIS/WHT (blue), and EFOSC/3.6 NNT (Green) together with the best fit derived with FERRE. Three well known metal-poor stars are shown for comparison, SDSS J0723+3637, G64−12, and CS 30336−0049. The main stellar parameters are also displayed.

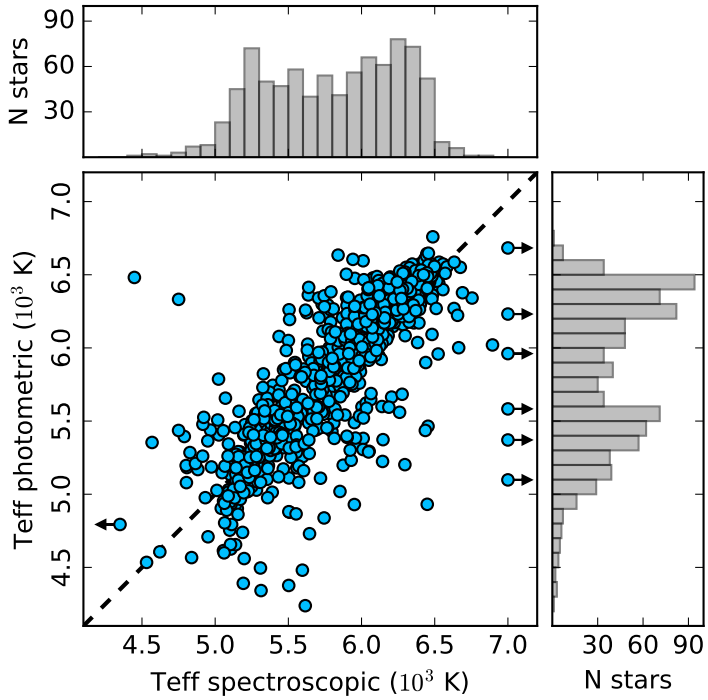

Figure 3.4 — Comparison of the photometric temperatures to those derived spectroscopically with FERRE. Stars that fall outside of the plotted region are plotted on the edge, and marked with an arrow. Histograms of both distributions are shown to indicate the point density.

reduce the uncertainties from the spectral analysis, we remove the bluest part of the spectrum most affected by noise, considering only the region redder than 3700 Å.

3.3 Analysis with FERRE

The entire sample of the spectroscopic data has been analyzed using the grid of synthetic stellar spectra computed with the ASSET code (Koesterke et al. 2008) and published in Aguado et al. (2017b, hereafter DA17). The model atmospheres were computed with the Kurucz codes, and are described in Mészáros et al. (2012). We use the FERRE[†] code (Allende Prieto et al. 2006) to search for the best fit to the observed spectrum by simultaneously deriving the main three stellar atmospheric parameters (effective temperature – $T_{\rm eff}$, surface gravity – $\log g$, metallicity – [Fe/H]), and carbon abundance [C/Fe]. FERRE is able to interpolate between the nodes of the grid and provide a synthetic spectrum for each set of derived parameters. A similar analysis implemented in Chapter 2 used the Powel's truncated Newton algorithm to find the best fit solution. However, for this work we also use a Markov Chain Monte Carlo (MCMC) algorithm based on self-adaptative randomized subspace sampling (Vrug et al. 2009), which provides the added advantage of deriving uncertainties by sampling the probability distribution function. The grid of synthetic spectra spans the space $-6 \leq$ [Fe/H] ≤ -2, $-1 \leq$ [C/Fe] ≤ 5, $4\,750\,{\rm K} \leq T_{\rm eff} \leq 7\,000\,{\rm K}$ and $1.0 \leq \log g \leq 5.0$. Although we targeted objects in the $[{\rm Fe/H}]_{\rm Pristine} < -2$ regime,

[†] FERRE is available from http://github.com/callendeprieto/ferre

there were some stars that were observed with higher metallicities. Those targets were re-analyzed with a more generic grid, suitable for higher metallicities, and described in Allende Prieto et al. (2018).

In order to cross-validate our analysis methods, we observed a number of well-known EMP stars from the literature that have robust stellar parameter determinations from high-resolution analyses. Comparing those stellar parameters with the ones measured in this work, we find a median deviation of 177 K, 0.86, and 0.27 dex for $T_{\rm eff}$, $\log g$, and [Fe/H], respectively. Table 3.2 summarizes the FERRE analysis performed on this sample and demonstrates that our derived metallicities are in very good agreement with those from the literature, thus demonstrating the ability of our method to derive precise metallicities using medium-resolution spectra. More comparisons of stellar parameter determination with the FERRE code and standard stars in the literature can be found in DA17. Figure 3.3 shows a subsample of the observed spectra together with the best fit synthetic spectrum as determined by FERRE for each of the three different instruments as well as three more well-known metal-poor stars.

3.3.1 Stellar parameters

To simultaneously derive $T_{\rm eff}$, $\log g$, [Fe/H], and [C/Fe], we smooth the grid of models and resample them to the appropriate resolving power corresponding to each instrument (see Table 3.1). We then normalize both the synthetic models and the observed spectra using a running-mean filter with a 30-pixel window (see DA17 for further details). Finally, FERRE derives the set of parameters assuming $[\alpha/{\rm Fe}] = +0.4$ and a fixed value of the microturbulence of 2.0 km s^{-1}.

$T_{\rm eff}$ is obtained by fitting the entire spectrum, although the derived $T_{\rm eff}$ is largely influenced by the Balmer lines present in the spectral range (H_β-4861Å, H_γ-4340Å, H_δ-4101Å, H_ϵ-3970Å, H_ζ-3889Å, H_η-3835Å, H_θ-3797Å, H_ι-3770Å, H_κ-3750Å, H_λ-3734Å, H_μ-3721Å). The temperature determination method relies on the broadening theory of the Balmer lines which is described in Barklem et al. (2000). The running mean normalization reduces the dependence on the specific determination of the continuum, allowing improved temperature determinations based on the shape of each H line, even with a moderate S/N ($\sim 15-20$). DA17 consider a systematic uncertainty for deriving temperatures of $\delta T_{\rm eff}$= 100 K, which is then combined quadratically with the statistical error from the MCMC method. Referring back to Table 3.2, the derived effective temperatures are in good agreement with those from previous works. In Figure 3.4, we show the relation between the photometric temperatures derived using the SDSS $(g-i)$-temperature relation[‡] and the temperatures derived from FERRE using the spectroscopic data from IDS, ISIS, and EFOSC.

Measuring $\log g$ values from medium-resolution spectra when no Fe II lines are available is a challenge. Particularly at moderate S/N ($\sim 15-25$), the shape of the Balmer lines alone do not allow for it to be derived precisely. However, a coarse classification between the dwarf/giant regimes is possible with FERRE. Robust $\log g$ determination in metal-poor stars using Gaia data is possible, but good quality parallax measurements are required. Since this is not available for most of our sample, particularly the fainter objects, we use the spectroscopic values from FERRE and assume the same systematic error as were used in DA17 of $\delta \log g$= 0.5.

The deepest metallic absorption in the optical wavelength range is caused by the Ca II H & K resonant lines at 3933 Å and 3968 Å respectively. Using these features as robust indicators to infer metallicities in EMP stars with low/med-resolution spectroscopy is a longstanding method (see for example Beers et al. 1985, 1992; Ryan & Norris 1991; Carney et al. 1996), and is still used today in large spectroscopic surveys such as SEGUE, BOSS and LAMOST (see for example Caffau et al. 2013a; Aguado et al. 2016, 2017a; Li et al. 2015, 2018; François et al. 2018,

[‡]For the equation used to compute the photometric temperatures, see the InfraRed Flux Method (IRFM), `https://www.sdss.org/dr12/spectro/sspp_irfm/`, [Fe/H] = −2.5 was asssumed.

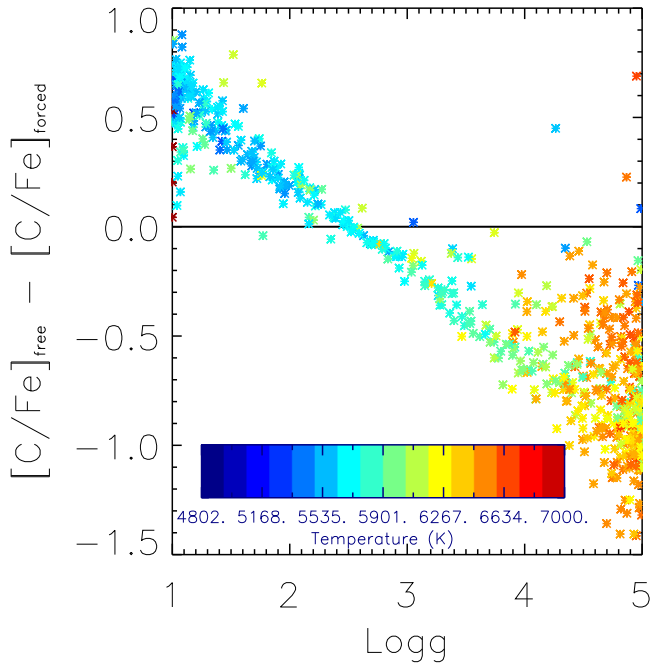

Figure 3.5 — An analysis over the *Pristine* subsample observed with IDS on the INT. The y-axis shows the difference between the original [C/Fe] determination and the one we derive assuming a fixed $\log g = 2.5$ and the x-axis shows the derived $\log g$. Points are colour coded by $T_{\rm eff}$.

and references therein). However, there is additional information present in this spectral range, such as the Mg I b triplet and some weak Fe I and Sr II lines, and these features can also contribute to the derivation of metallicities, provided that the S/N is high enough to resolve them. Reassuringly, we find good agreement between our [Fe/H] values and those from high-resolution analyses even with the relatively low-resolution EFOSC2 instrument (R ∼ 1 000, see Table 3.2 and section 4.1 in DA17). DA17 assumed a systematic uncertainty in metallicity of 0.1 dex. However, due to the significantly lower S/N of the current sample, the ISM contribution to the Ca II H & K absorption lines is largely unresolved for most of the spectra. Therefore, we assume a more conservative value of δ[Fe/H] = 0.2 dex and add it to the derived statistical uncertainty.

3.3.2 Carbon abundance

Due to the lack of spectral features in EMP stars, particularly at higher $T_{\rm eff}$, it is not always possible to derive a reliable carbon abundance, particularly with medium-resolution spectroscopy (e.g. Bonifacio et al. 2015). DA17 provide some reference levels regarding our ability to measure carbon, but as previously discussed, the average S/N in the current sample is significantly lower. With the aim to constrain the confidence levels with which it is possible to derive carbon abundances without important systematic effects, we performed the following theoretical exercise.

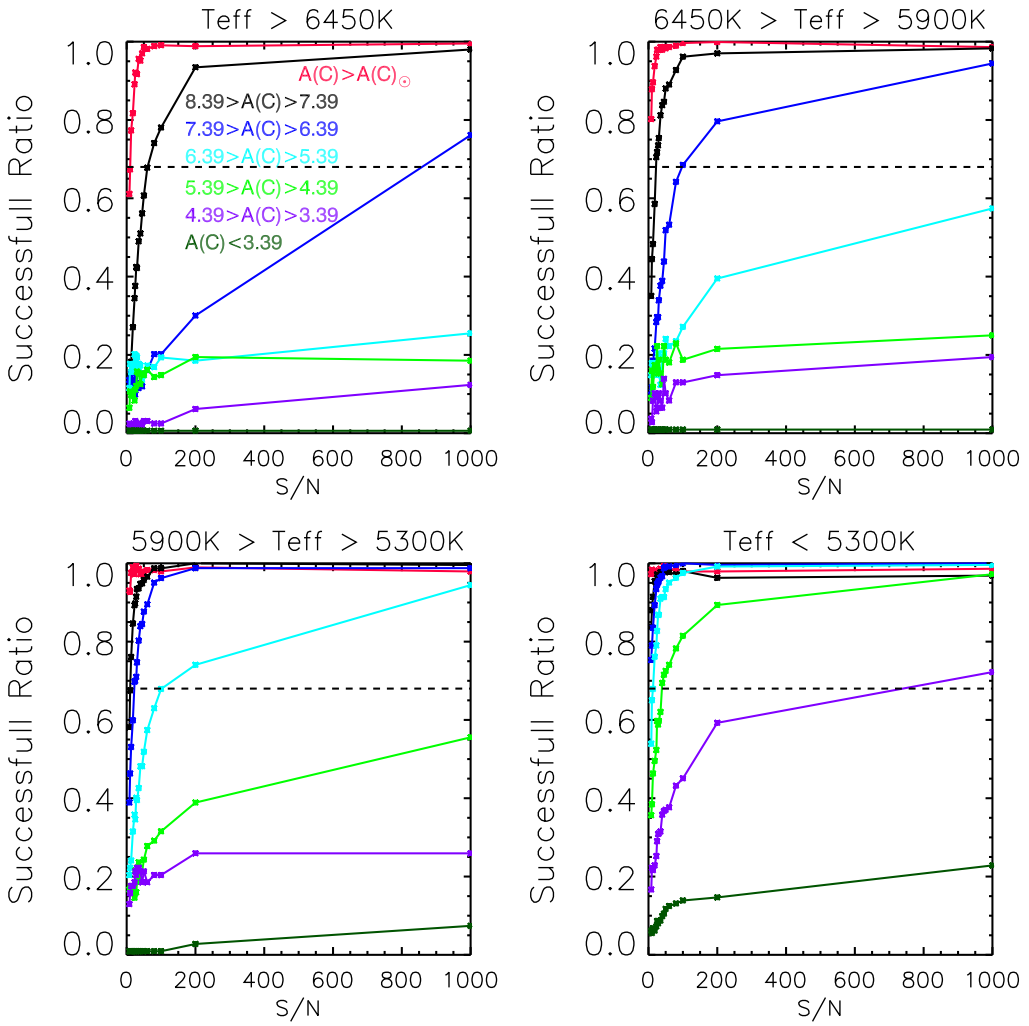

Figure 3.6 — The successful ratio measuring carbon abundance in a set of 30 854 synthetic spectra covering a wide range of effective temperatures, absolute carbon abundance and different levels of signal-to-noise ratio. Dashed line is 68% of successful ratio.

A set of synthetic spectra with the same coverage as our IDS/ISIS/EFOSC data (3600-5200 Å) were computed with ASSET R = 3 000. A total number of 5 670 spectra covering different ranges of T_{eff}, $\log g$, and absolute carbon abundance, A(C), were analyzed using 10 Markov Chains of 1 000 experiments each for different values of S/N ranging from 8 to 200. In total, 30 854 spectra were analyzed with FERRE. We then compared the synthetic absolute carbon abundance A(C) and the corresponding [C/Fe] derived value. We marked a given trial as successful if it was able to recover the theoretical value provided by the synthetic grid, where $|A(C)_{in} - A(C)_{out}| < \epsilon_{A(C)}$, where $\epsilon_{A(C)}$ is the assumed systematic uncertainty of 0.2 dex as estimated in DA17. Figure 3.6 shows all the ratios versus S/N for different effective temperatures and carbon abundances. For this work, we consider the reliable areas of the plot to be those where the correct value is recovered with a frequency that is higher than 68%. For example, at solar ($A(C)$ = 8.39, Asplund et al. 2005) or higher carbon abundance (red line), we are able to measure [C/Fe] at any

Table 3.3 — The minimum S/N values needed to detect the carbon G-band with R~ 3 000 as a function of T_{eff} and A(C).

A(C)	< 5300 K	5300 – 5900 K	5900 – 6450 K	> 6450 K
⩾ ⊙	8	8	8	12
8.4 – 7.4	8	10	25	60
7.4 – 6.4	10	25	100	800
6.4 – 5.4	15	100	–	–
5.4 – 4.4	45	–	–	–
4.4 – 3.4	900	–	–	–
> 3.4	–	–	–	–

temperature with S/N > 10, while for values below $A(C) = 4.4$ it is not likely to be able to detect carbon at this resolution, regardless of the T_{eff}. Table 3.3 summarizes the approximate S/N required to detect the G-band in *Pristine* spectroscopic data. As expected, lower temperatures allow for a better carbon determination due the larger absorption of the G-band. We apply these cuts to the sample and only provide carbon abundance values for the 169 (18%) stars that satisfy these criteria. We note that we are able to measure carbon in 10% of the stars with [Fe/H] < −3, and 24% of the stars with −3 < [Fe/H] < −2.

In order to better understand the systematics involved in the determination of carbon, we assess its correlation with the determination of $\log g$. In Figure 3.5, we compare the derived [C/Fe] with those values we find if we fix $\log g = 2.5$ as a function of $\log g$. The points are also coloured according to T_{eff}. The most relevant part of the plot is the giant regime since this is where the majority of the sample with good [C/Fe] determinations are located, due not only to the lower temperatures (See Table 3.3), but also because at these $\log g$ values carbon is more likely to be overestimated and therefore considered to be reliably determined based on the criteria in Table 3.3. As a result, for the stars for which we derive carbon abundances, we have systematic uncertainties which are large but well delimited, especially at S/N < 25. Therefore, we assume systematic uncertainties from 0.2 dex up to 0.6 dex, depending on the S/N of the spectrum and subsequent reliability of the $\log g$.

3.4 Results

3.4.1 Comparison of the photometric metallicities and spectroscopic metallicities

Photometric metallicities were derived using the *Pristine* narrow-band photometry and the SDSS broad-band photometry. The detailed methods of this procedure are described in Starkenburg et al. (2017).

Figure 3.7 shows the relation between the photometric and spectroscopic metallicities. In the left panel, we show the total parameter space occupied by the data, and in the right panel is a zoomed-in view to better show the details of the plot. We only plot the 863 stars for which there are reliable FERRE and *Pristine* metallicity determinations. For the former, these are stars flagged with "X" in Table 3.4, described in Section 3.4.4. For the later, we have removed stars that exhibit variability, that may be white dwarfs, that are identified as non-point sources in their PSFs, and that are flagged as being problematic in their SDSS *g* or *i* broad band magnitudes (mainly bright sources that show some saturation). These criteria are described in greater detail in the list below Figure 3 in Chapter 2. Here, we have omitted all criteria based

Table 3.4 — Metallicities, temperatures, and carbon abundances of the *Pristine* spectroscopic sample. Uncertainties include systematic and statistical errors. We only include a small sample of 9 of the 1 007 stars observed, to illustrate the structure of the table. The columns are described in more detail in the text. The full table is available online. The adopted solar abundances are those from Asplund et al. (2005).

Name Origin Units	$V^{\dagger}$ SDSS mag	$CaHK$ Pristine mag	[Fe/H] Pristine	[Fe/H] FERRE	Teff FERRE K	log g FERRE	[C/Fe] FERRE	S/N FERRE px^{-1}	Flag Q,C	Previously observed
P138.xxxx+16.xxxx*	16.06	16.90	-2.84	-3.4 ± 0.2	5266 ± 122	4.8 ± 0.5	0.3 ± 0.4	34	X,-1	-
P149.1350+15.0447	16.06	16.38	-2.72	-2.5 ± 0.2	6047 ± 104	1.1 ± 0.5	0.5 ± 1.0	49	X,-1	-
P151.4987+13.9300	16.67	16.90	-2.84	-2.9 ± 0.6	5380 ± 176	1.0 ± 0.5	1.0 ± 0.5	26	X,-1	LAMOST,SEGUE
P184.xxxx+43.xxxx*	15.92	16.71	-2.94	-3.7 ± 0.2	5509 ± 103	4.9 ± 0.5	0.4 ± 0.5	24	X,-1	-
P185.8616+41.3093	15.83	16.71	-2.87	-2.8 ± 0.2	5264 ± 104	1.4 ± 0.6	0.8 ± 0.4	25	X,1	SEGUE
P218.6977+15.5932	15.57	15.89	-2.93	-3.1 ± 0.2	6305 ± 111	4.4 ± 0.5	0.7 ± 0.8	28	X,-1	-
P220.7009+13.1405	16.88	17.49	-3.40	-3.3 ± 0.2	5464 ± 111	3.5 ± 0.5	1.1 ± 0.4	26	X,-1	-
P235.0067+07.1438	16.88	17.12	-3.10	-3.0 ± 0.2	6216 ± 108	4.9 ± 0.5	0.0 ± 0.7	27	X,-1	SDSS
P256.0374+17.0031	16.51	17.12	-2.97	-2.6 ± 0.2	5552 ± 117	3.6 ± 0.5	0.3 ± 0.2	45	X,-1	-

*Coordinates of select stars have been removed as they are the subject of an ongoing high-resolution follow-up study (Kietly et al., in prep.)
†derived using SDSS g and r according to Lupton (2005), `https://www.sdss3.org/dr8/algorithms/sdssUBVRITransform.php`

on metallicity but keep all criteria pertaining to photometric quality. Many of the removed stars were observed early on in the follow-up campaign as we were improving our selection of targets. These are, however, still included in the full table for completeness since the derived spectroscopic metallicities are not affected by the problematic photometry.

In Figure 3.7, most of the stars are clustered at $-3.5 <$ [Fe/H] < -2.0 due to our follow-up strategy of the best metal-poor candidates first. Since there are more metal-rich stars than metal-poor stars, the metal-rich stars will scatter into the metal-poor regime with a higher frequency than the other way around, and the relative contamination will be higher at the metal-poor end. The combination of this effect and the photometric selection function from the follow-up strategy produces the offset from the 1 to 1 relation (black-dotted line). The right panel also shows a fairly significant dispersion, but given that the uncertainties are on the order of ~ 0.2 dex for both the vertical and horizontal axes, it is not surprising to see a dispersion of $\sim$0.5 dex, although the scatter seems more severe due to the small range in metallicities covered by the data, and the outliers at $[Fe/H]_{FERRE} < -2$. We also note that it is not crucial to have a tight relation in this space, because a coarse differentiation of stars as EMP or VMP is enough to identify promising candidates for follow-up, as well as for much of the interesting ancillary science cases.

There is a distinct population of stars for which the photometric metallicities from *Pristine* are highly discrepant from the spectroscopic metallicities. These are seen in the left panel of Figure 3.7 as the tail of stars extending to $[Fe/H]_{FERRE} > -2$. The criteria for selecting stars for spectroscopic follow-up was investigated and summarized in detail in Chapter 2. Despite ensuring good quality photometry, cleaning white dwarfs (cutting all stars with $(u_0 - g_0) < 0.6$) and variable stars, there are still 12% of stars predicted to have $[Fe/H]_{Pristine} < -2.5$ that have $[Fe/H]_{FERRE} > -2$. This number rises to 18% for $[Fe/H]_{Pristine} < -3$ (see Table 3.5). Many of these stars have a large temperature discrepancy between spectroscopy and photometry ($|\Delta T_{eff}| >$ 500 K for $\sim$ 40% of these stars), which probably indicates problems with the SDSS broad-band photometry for these stars. This would, in turn, affect the colour, and thus the measured photometric metallicity. In addition, some of this contamination may be attributable to long-period variable stars that were not detected in the Pan-STARRS1 variability catalogue, non-stellar objects, or chromospherically active stars with Ca H & K in emission (although we note that only 9 such objects with peculiar spectra were identified in the follow-up spectroscopy).

At the lowest metallicities of $[Fe/H]_{Pristine} < -3.5$, the percentage of stars with spectroscopic [Fe/H] > -2 rises to 57%. This clearly indicates an increasing contamination fraction with decreasing metallicity. Although the slope in this region of the Milky Way metallicity distribution function is not well constrained, it is known to be quite steep, such that stars at these metallicities are incredibly rare with respect to stars of higher metallicity. As a result, even a small number of interloping higher metallicity stars can dominate the candidate sample at these low metallicities.

3.4.2 Updated purity and success rates of the *Pristine* survey

The success rates of the *Pristine* survey were reported after the first year of spectroscopic follow-up using a sample of 205 stars observed at medium-resolution at the WHT and INT (see Chapter 2). Due to the small size of that sample, the success rates for finding metal-poor stars computed from them were preliminary estimates. Now that we have a larger follow-up sample of nearly 5 times as many stars, we can update these numbers with better statistics. In order to remain consistent and to allow for an easy comparison, we will use the same metrics to quantify the purity and success rates as were used in Chapter 2, namely:

$$\text{success rate}\,\% = \frac{[Fe/H]_{FERRE} < X}{[Fe/H]_{Pristine} < X} \times 100,$$

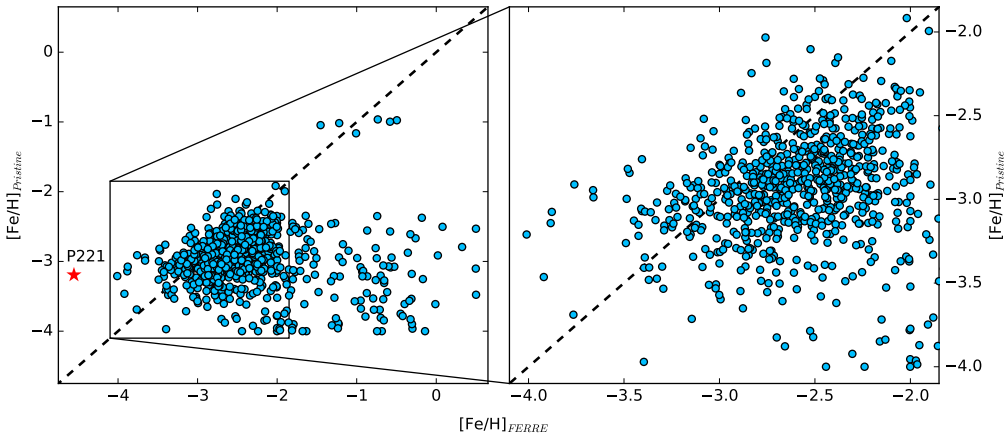

Figure 3.7 — Photometric metallicities derived with *Pristine* ([Fe/H]$_{Pristine}$) versus spectroscopic metallicities derived with FERRE ([Fe/H]$_{\rm FERRE}$) for the total sample. The red star represents Pristine_221.8781+9.7844 from Starkenburg et al. (2018). The left panel shows the full metallicity space covered by the follow-up sample, and the right panel shows a zoom in of the highest density region around $-4 <$ [Fe/H] < -2.

where FERRE refers to the spectroscopically derived [Fe/H], *Pristine* to the photometric prediction by *Pristine* and X the metallicity range of interest.

For all of the stars included in Table 3.4, we did not make a cut in S/N, but rather checked by eye the goodness of the fit for the synthetic spectrum by FERRE. The reason for this is because stars that are cooler and more metal-rich have larger absorption lines, and are therefore easier to classify at lower S/N than stars that are hotter and more metal-poor. As a result, we successfully determine the metallicities for more stars at higher metallicities ([Fe/H] > -2) with low S/N values (S/N < 15), rather than to cut these stars out with a S/N cut. However, for the calculation of the success rates, this would bias our sample with more metal-rich stars and fewer metal-poor stars. Therefore, we compute the success rates using only stars with S/N $>$ 25, the regime in which we can reliably measure metallicities, even at [Fe/H] < -3. Taking this sample, we find a success rate of 23% for finding stars with [Fe/H] < -3.0, and 56% for finding stars with [Fe/H] < -2.5. In Chapter 2, we reported a success rate of 22% for [Fe/H] < -3.0, and 70% for [Fe/H] < -2.5. This discrepancy can be attributed to the cut at S/N > 25. If we make the same cuts as we made in Chapter 2, this decreases the success rates to 20% and 58% for [Fe/H] < -3 and [Fe/H] < -2.5, respectively, meaning that those values are fully compatible with what we find in this work. In Chapter 2, we did not originally make a cut at S/N > 25 when computing the success rates as this would have reduced the sample from 205 down to 62 stars, leading to uncertainties of low number statistics. In the current work, making this cut still leaves 331 stars, and still allows for a robust determination of the success rates.

We therefore update the success rates of the *Pristine* survey to 23% for [Fe/H] < -3.0, and 56% for [Fe/H] < -2.5. These values, along with other diagnostics, such as the contamination rate (fraction of stars with [Fe/H] > -2) are summarized in Table 3.5.

3.4.3 The carbon-enhancement present in the sample

Figure 3.8 shows the distribution of absolute carbon (A(C); bottom panel) and [C/Fe] abundances as a function of metallicity for the 169 stars for which we are able to make a reliable

carbon determination (Section 3.3.2). Both CEMP reference lines at [C/Fe] = 1.0 (Beers et al. 2005) and [C/Fe] = +0.7 (Aoki et al. 2007) are plotted as solid and dashed lines, respectively. The high-resolution carbon abundance value for Pristine_221.8781+9.7844 is also included as the red star.

To compute the CEMP fractions, we first draw a new sample of values for the [C/Fe], and [Fe/H] measurements, taking into account both the statistical and systematic uncertainties of each. We then compute the fraction of stars with [C/Fe] above the two limits of [C/Fe] = +1.0 and +0.7, and repeat this exercise 10^6 times in a Monte Carlo fashion. The resulting distributions are approximately Gaussian in shape, and are therefore reasonably well described by a mean and standard deviation. For the [Fe/H] < −3 sample, we compute CEMP fractions of 58 ± 14% and 43 ± 13% for the [C/Fe] > +0.7 and +1.0, respectively. For the −3 < [Fe/H] < −2 sample, we compute CEMP fractions of 41 ± 4% and 23 ± 3% for [C/Fe] > +0.7 and +1.0, respectively.

Placco et al. (2014) find 43% of stars to have [C/Fe] ≥ +0.7 and −4 < [Fe/H] < −3, a value which differs at the 1σ level compared to the 58 ± 14% derived in this work, and is therefore not statistically significant. Furthermore, Norris & Yong (2019) perform a rigorous analysis of the 3D and NLTE corrections relevant for the carbon abundance determinations, and demonstrate a significant decrease in the carbon content for a number of CEMP stars from the literature when full 3D-NLTE corrections are taken into account. The CEMP-no group are stars that do not show significant enrichment in neutron-capture elements (s- and r-process), and are the most numerous subgroup among CEMP stars. As demonstrated in Norris & Yong (2019), applying the 3D NLTE corrections can result in a significant number of CEMP-no stars to become carbon-normal. We do not know the fraction of CEMP-no stars in our current sample, but if we consider that a typical fraction of them could belong to this category, it is likely that the computed CEMP fractions would decrease considerably. It is therefore difficult to draw firm conclusions from this current sample of CEMP stars, but further, more detailed follow-up –particularly targeting carbon and the neutron capture elements in the EMP stars– could provide a very useful sample with which to investigate this further.

3.4.4 The full sample

In this chapter we present a full catalogue from three years of follow-up spectroscopy of *Pristine* candidates. The full table, consisting of 1 007 stars is available online and can be accessed from Aguado et al. (2019). An abbreviated version of the full table showing the provided columns as well as a sample of 9 rows is shown in Table 3.4. The column $CaHK$ is the magnitude obtained from the *Pristine* narrow-band filter, the column $[Fe/H]_{Pristine}$ is the photometric metallicity determined using the $(g_0 - i_0)$ SDSS colours and *Pristine* photometry (described in Starkenburg et al. 2017, Section 3.2). The next two columns are the spectroscopic metallicities, effective temperatures and surface gravities derived from FERRE and their associated uncertainties. Column S/N is the signal-to-noise ratio of the analyzed spectrum. We also provide a Q-flag, representing the reliability of the spectroscopic metallicity determination. An entry of "X" indicates that the synthetic spectral fit was reliable and that the given $[Fe/H]_{FERRE}$ value can be trusted to within the provided uncertainties (93% of the sample have this flag). In order to provide as much information as possible, we also provide tentative metallicity values for stars for which the S/N is too low for a robust determination of stellar parameters, but that still have some information in the observed spectrum. These stars are given a flag of T (6% of the sample), and are good candidates to be re-observed with higher S/N and at higher resolution. The C-flag shows if the carbon determination is reliable (1) or not (−1), and was derived based on S/N and temperature criteria described in Section 3.3.2. The last column indicates whether the object was already spectroscopically observed by other surveys. Finally, the object coordinates are

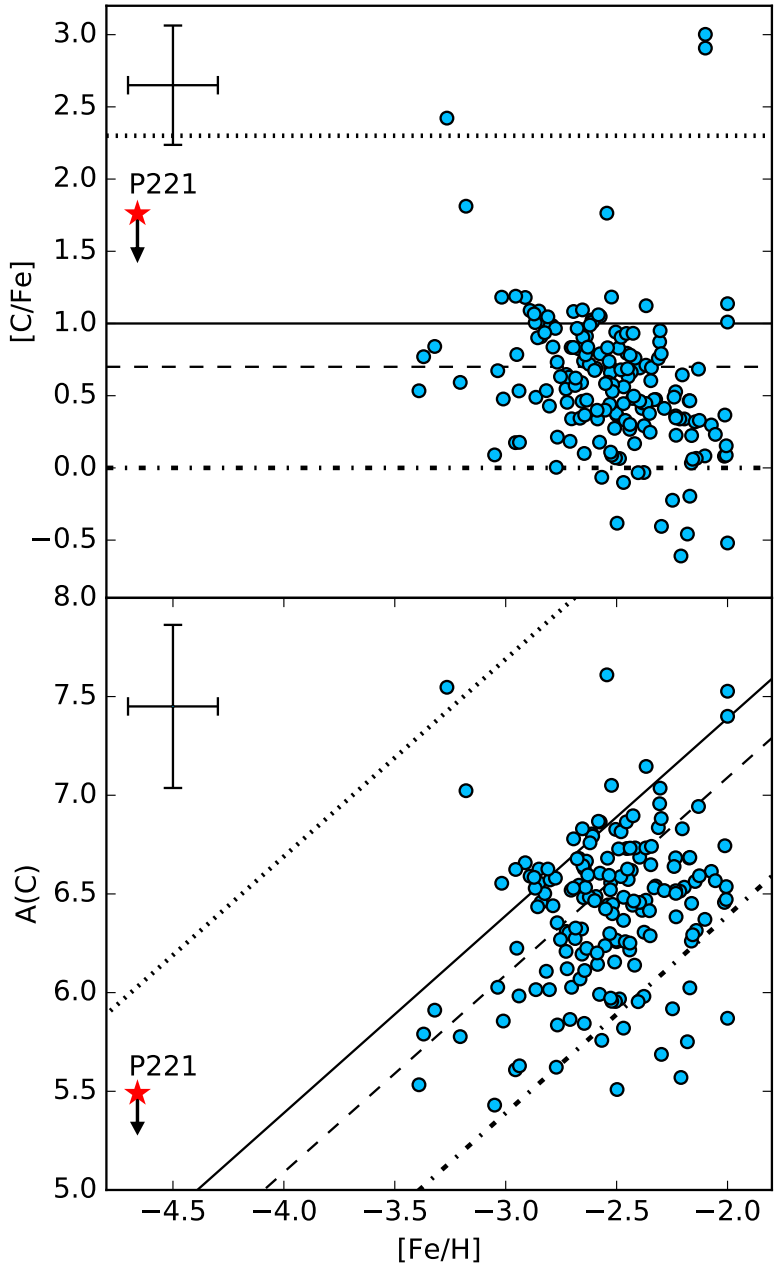

Figure 3.8 — Carbon vs iron for [C/Fe] (top) and absolute carbon (bottom). The red star represents Pristine_221.8781+9.7844, with parameters derived from the analysis of a high-resolution UVES spectrum (Starkenburg et al. 2018). The dashed-dotted line at [C/Fe] = 0 shows the solar carbon abundance, the dashed and solid lines at [C/Fe] = +0.7 and +1.0 show the thresholds for carbon enhancement defined in Aoki et al. (2007) and Beers et al. (2005), respectively, and the dotted line at [C/Fe] = 2.3 shows the boundary of the silicate dominated region, as described in Chiaki et al. (2017). The error bars shown in the top left of each panel represent the median errors of the sample.

Table 3.5 — Numbers of stars with photometric predictions $[Fe/H]_{Pristine}$ below −2.5 and −3.0, the numbers of stars that are spectroscopically confirmed below those metallicities, and the success rates (bolded), given for all stars with S/N > 25, with the selection criteria applied (described partially in Section 3.4.1, and in detail in Chapter 2), and the sample of stars with $[Fe/H]_{Pristine} \leq -3.0$.

	All stars S/N > 25	Selection criteria S/N > 25	$[Fe/H]_{Pristine} \leq -3.0$ S/N > 25
Total number	344	331	129
$[Fe/H]_{Pristine} \leq -2.5$	325/344 (94%)	315/331 (95%)	129/129 (100%)
$[Fe/H]_{Pristine} \leq -3.0$	132/344 (38%)	129/331 (39%)	129/129 (100%)
$[Fe/H]_{FERRE} \leq -2.5$	184/344 (53%)	180/331 (54%)	76/129 (59%)
$[Fe/H]_{FERRE} \leq -3.0$	48/344 (14%)	47/331 (14%)	30/129 (23%)
$[Fe/H]_{FERRE} \geq -2.0$	45/344 (13%)	39/331 (12%)	23/129 (18%)
success $[Fe/H] \leq -2.5$	178/325 (55%)	**175/315 (56%)**	-
success $[Fe/H] \leq -3.0$	30/132 (23%)	30/129 (23%)	**30/129 (23%)**

Table 3.6 — Number of candidate stars in different magnitude bins and metallicity ranges. The first number in each cell is the number of stars followed up with spectroscopy from the sample in this chapter, and the second is the total number of candidates as of the time of publication over the ∼ 2 500 deg^2 of the *Pristine* footprint used to select candidates that are the focus of this chapter. [Fe/H] values shown are photometric *Pristine* metallicities.

	# Candidates	[Fe/H]≤-2.5	[Fe/H]≤-3.0
$V < 15$	169/509	139/293	66/92
$15 < V < 16$	535/1 809	475/989	160/206
$16 < V < 17$	246/5 423	238/2 785	148/540
$17 < V < 18$	57/14 682	56/7 321	43/1 393
$18 < V < 19$	0/35 036	0/16 887	0/3 977
Total	1 007/57 459	908/28 275	417/6 208

contained in the name, but we provide these explicitly and with greater precision on the online version of the table.

There are a small number of stars for which the *Pristine* metallicity classification fails, meaning that from photometry the object was expected to be a metal-poor star, but from spectroscopy it was determined to be some other type of object. These could be stars with CaHK in emission, non-stellar objects, or various other objects with unusual behaviour in the CaHK region. However, this only occurs for 9 of the observed objects (< 1% of the sample), indicating that the sample is well cleaned. We remove these 9 objects from the catalogue since both their photometric and spectroscopic metallicities are unreliable, but consider them in the sample when computing the success rates since they do contribute to the contamination.

3.5 Future of the survey

In addition to hunting for the most metal-poor stars in the Galaxy, the photometric metallicities that are produced by the narrow-band photometry of the *Pristine* survey can be used for several other interesting science cases. For instance, Longeard et al. (2018) conducted an in depth study of the metallicity distribution and velocity dispersion of the faint Milky Way satellite *Draco II* using $CaHK$ photometry, and work is ongoing on a similar analysis to characterize the properties of many other nearby satellites (Longeard et al. 2019). In Chapter 5 we demonstrate

the powerful capabilities of the *Pristine* narrow-band filter to identify blue horizontal branch (BHB) stars and disentangle them from the contaminating blue straggler (BS) population, providing a uniquely clean sample of distance indicators with which to study the outer reaches of the Galactic halo. Finally, Arentsen et al. (2020) are studying the metal-poor component of the Galactic bulge with the Pristine Inner Galaxy Survey (PIGS).

3.5.1 Pristine and Gaia

The highly anticipated Gaia data have initiated a revolution in the study of Galactic archaeology and it is changing our understanding of the Galaxy (Gaia Collaboration et al. 2018). The latest data release provided high-precision astrometric measurements and 3-filter photometry for over 1.3 billion stars. The range of possibilities for using Gaia photometry together with more than 5 million *Pristine* metallicity determinations are broad, and open the door to an unprecedented mapping of the Galaxy using the full six-dimensional phase-space plus metallicity information. For example, work is ongoing using Gaia and *Pristine* to study the substructures present in and around our Galaxy and their dependence on metallicity (see Chapter 6), as well as an analysis of the metallicity distribution function of the halo at the lowest metallicities (see Chapter 4). On the other hand, the exquisite Gaia parallaxes, proper motions and photometry allow us to derive surface gravities, effective temperatures and orbits for EMP stars (see for example Bonifacio et al. 2018a; Frebel et al. 2019; Sestito et al. 2019a). The dynamics of the most ancient stars of the Milky Way could be a crucial piece of information for understanding the formation and evolution of the Galactic halo. For example, recent work by Sestito et al. (2019a), demonstrated that an important fraction of the known UMP stars seem to have orbits that are confined to Galactic plane, suggesting interesting new scenarios for their origins. In a follow-up study, Sestito et al. (2019b) substantiated these findings using a larger sample of EMP stars based on the sample presented in this chapter. Finally, Bonifacio et al. (2019) combined Gaia parallaxes and *Pristine* photometry to derive photometric metallicities, effective temperatures and surface gravities. These authors also studied the chemical composition and ages of 40 metal-poor stars with the SOPHIE high-resolution spectrograph (Bonifacio et al. 2019).

3.5.2 Pristine and WEAVE

The impending arrival of upcoming large spectroscopic surveys will nicely complement the still ongoing Gaia project. A new, deeper view –not only kinematically but also chemically– of the Milky Way halo, will shed light on the formation and evolution of the Galaxy. This unprecedented amount of high quality data will greatly expand the capabilities of Galactic archaeology studies thanks to surveys like 4MOST, DESI or WEAVE. The success rates presented in Table 3.5 demonstrate that the Pristine filter is one of the best ways to pre-select EMP candidates to observe in those surveys. In particular, the WEAVE project will devote up to twenty fibers per WEAVE 3.14 deg^2 field of view to Pristine-selected EMP candidates in the magnitude range $15 < G < 19$, in the low-resolution Galactic archaeology survey of high Galactic latitudes (Jin et al. 2019 in preparation). Over the planned $\sim 8\,500$ deg^2 of the survey, of which we anticipate $\geq 5\,000$ deg^2 will be in common with the Pristine footprint at the time they are observed in WEAVE, this adds up to $\sim 30\,000$ candidate EMP stars, of which according to Table 3.5, $\sim 5\,000 - 7\,000$ should turn out to be [Fe/H] < -3. This would increase the number of spectroscopically confirmed EMP and UMP stars with known chemical signatures by one order of magnitude. After five years of observing we we expect to have measured the chemical abundances such as C, Na, Mg, Al, Si, Ca, Ti, Fe, for about $\sim 3\,000$ stars with [Fe/H] < -3.0 and $\sim 150 - 200$ stars with [Fe/H] < -4.0, including $\sim 5 - 10$ hyper metal-poor stars ([Fe/H] < -5.0),

doubling the samples currently available from several decades of efforts. Additionally, WEAVE Galactic archaeology high resolution survey (HR) will be able to measure the full suite of chemical signatures for the brightest part of the Pristine sample ($g \leq 15.5$) where it overlaps with the WEAVE HR survey, although the density of such bright targets will be much lower.

3.6 Conclusions

Expanding upon the previous work conducted in Starkenburg et al. (2017) and Chapter 2, we have presented a sample consisting of 1 007 stars, representing three years of follow-up of medium- and low-resolution spectroscopy of EMP candidates from the *Pristine* survey. The number of stars followed-up spectroscopically has increased by a factor of 5, allowing for the success rate of stars with [Fe/H] < -2.5 and [Fe/H] < -3.0 to be updated to 56% and 23%, respectively. This is a relevant milestone in the field of Galactic archaeology, demonstrating the utility of the *Pristine* filter to select EMP candidates for the next generation of spectroscopic surveys such as WEAVE. The recent discovery of Pristine_221.8781+9.7844 (Starkenburg et al. 2018), the second most metal-poor star yet discovered, shows that *Pristine* photometry is also effective in finding UMP stars in the most interesting and poorly populated regime of [Fe/H] < -4.5. In addition, we demonstrated that the FERRE code is capable of deriving stellar parameters even at relatively low-resolution, namely with the stars observed with EFOSC2. Furthermore, we show for the first time in the *Pristine* project that we are able to provide individual carbon abundances from measurements of the G-band with moderate S/N in medium-resolution spectra. We derive carbon for 169 stars, or $\sim$ 20% of the total sample, although the lower average S/N as compared to DA17 results in higher overall uncertainties in the carbon measurements than previously achieved. With this medium-resolution follow-up sample (along with the previous analysis of Chapter 2), we have thoroughly characterized the photometric selection and success rates of the *Pristine* survey in this magnitude range. Further observations with low- and medium-resolution spectroscopic facilities of metal-poor candidates selected from *Pristine* are highly desirable to continue increasing the number of ultra/hyper metal-poor stars and also to provide a larger sample of CEMP and carbon-normal EMP stars. Future follow-up is planned to mostly be done with MOS facilities such as WEAVE.

Acknowledgements

We thank the reviewer, Tim Beers, for his thorough review and highly appreciate the comments and suggestions, which significantly contributed to improving the quality of the publication and the definitive shape of the online material. We gratefully acknowledge the Isaac Newton Group (ING) staff, in particular the support astronomers and staff at the INT/WHT for their expertise and help with observations. We also thank the staff at ESO for helping during EFOSC observations, and the CFHT staff for performing the observations in queue mode. DA thanks the Leverhulme Trust for financial support. DA acknowledges the Spanish Ministry of Economy and Competitiveness (MINECO) for the financial support received in the form of a Severo-Ochoa PhD fellowship, within the Severo-Ochoa International PhD Program. DA, CAP, and JIGH and CAP also acknowledge the Spanish ministry project MINECO AYA2017-86389-P. JIGH acknowledges financial support from the Spanish Ministry of Science, Innovation and Universities (MICIU) under the 2013 Ramón y Cajal program MICIU RYC-2013-14875, and also from the Spanish ministry project MICIU AYA2017-86389-P. ES, KY and AA gratefully acknowledge funding by the Emmy Noether program from the Deutsche Forschungsgemeinschaft (DFG). This work has been published under the framework of the IdEx Unistra and benefits from a funding from the state managed by the French National

Research Agency as part of the investments for the future program. NFM, RI, NL, PB, EC, VH, CK, and PS gratefully acknowledge support from the French National Research Agency (ANR) funded project “Pristine” (ANR-18-CE31-0017) along with funding from CNRS/INSU through the Programme National Galaxies et Cosmologie and through the CNRS grant PICS07708. The authors benefited from the International Space Science Institute (ISSI) in Bern, CH, thanks to the funding of the Teams “The Formation and Evolution of the Galactic Halo” and “Pristine”. The French co-authors acknowledge support from the Agence National de la Recherche (ANR), through contract N. 183787. CL acknowledges financial support from the Swiss National Science Foundation (Ambizione grant PZ00P2_168065). DA thanks Fátima Mesa-Herrera from Laboratory of Membrane Physiology and Biophysics, University of La Laguna, for those beautiful nights observing the sky at La Palma in December 2017.

4 The metallicity distribution function of the Milky Way halo down to the extremely metal-poor regime

K. Youakim, E. Starkenburg, N. F. Martin, G. Matijevič, D. S. Aguado, C. Allende Prieto, A. Arentsen, P. Bonifacio, R. G. Carlberg, J. I. González Hernández, V. Hill, G. Kordopatis, C. Lardo, J. F. Navarro, P. Jablonka, R. Sánchez-Janssen, F. Sestito, G. F. Thomas K. Venn

— This chapter has been submitted as a paper to MNRAS —

ABSTRACT

The *Pristine* survey uses narrow-band photometry to derive precise metallicities down to the extremely metal-poor regime ([Fe/H] < -3), and currently consists of over 4 million FGK-type stars over a sky area of $\sim 2\,500\,\mathrm{deg}^2$. We focus our analysis on a subsample of $\sim 80\,000$ main sequence turnoff stars with heliocentric distances between 6 and 20 kpc, which we take to be a representative sample of the inner halo. The resulting metallicity distribution function (MDF) has a peak at [Fe/H] $= -1.6$, and a slope of $\Delta(\mathrm{LogN})/\Delta[\mathrm{Fe/H}] = 1.0 \pm 0.1$ in the metallicity range of $-3.4 <$ [Fe/H] < -2.5. This agrees well with a simple closed-box chemical enrichment model in this range, but is shallower than previous spectroscopic MDFs presented in the literature, suggesting that there may be a larger proportion of metal-poor stars in the inner halo than previously reported. We identify the Monoceros/TriAnd/ACS/EBS/A13 structure in metallicity space in a low latitude field in the anticenter direction, and also discuss the possibility that the inner halo is dominated by a single, large merger event, but cannot strongly support or refute this idea with the current data. Finally, based on the MDF of field stars, we estimate the number of expected metal-poor globular clusters in the Milky Way to be 10 for [Fe/H] $<$ -2.5 and 2.8 for [Fe/H] < -3, suggesting that the lack of low metallicity globular clusters in the Milky Way is not due simply to statistical undersampling.

4.1 Introduction

The extended stellar halo of our Galaxy is diffuse and full of stellar substructures, continually being discovered as overdensities and kinematical and/or chemical substructures by on-going large surveys, especially now in the era of the Gaia mission (e.g., Helmi et al. 2018; Haywood et al. 2018; Belokurov et al. 2018; Malhan et al. 2018; Koppelman et al. 2018; Myeong et al. 2018b). Together they relate a complex history in which the halo is built up through many merger events of smaller systems. While the majority and largest events are thought to have happened in the distant past, some are still continuing in the present day (with the most obvious example being the Sagittarius dwarf galaxy; Ibata et al. 1994). It is a complicated task to characterize such a complex structure, especially taking into account our vantage point of observing the halo from within the stellar disc. We know that the halo is highly structured, that its average metallicity is much lower than that of the disc(s) or inner Galaxy and that its stellar population has a very low net rotation (see for reviews on the stellar halo Helmi 2008; Belokurov 2013; Bland-Hawthorn & Gerhard 2016). It has further been suggested that the character of the stellar halo is dual in nature, consisting of two broadly overlapping components in density, orbits, and chemistry (Carollo et al. 2007; Beers et al. 2012), but this view has also been challenged (Schönrich et al. 2011, 2014) and is complicated further by the discovery of large substructures in the Gaia DR2 data, standing out in their (partly retrograde) orbits and chemistry in different ways (e.g., Helmi et al. 2018; Haywood et al. 2018; Belokurov et al. 2018; Myeong et al. 2019). Furthermore, simulations of merging galaxy systems have shown that in some cases a significant part of the inner halo can consist of in situ material from the original progenitor (as suggested by Abadi et al. 2006), rather than accreted material from merging satellites.

Robust observational results on the metallicity structure of the Galactic halo can help us understand not only its structure – as stellar substructures often stand out in metallicity space – but also the processes involved in its formation. The extremely low-metallicity tail of the metallicity distribution function (MDF), for instance, constrains the pace of chemical evolution at its earliest epochs (for a review on this topic see Frebel & Norris 2015). Metal-rich halo stars, on the other hand, are more likely to be deposited by larger rather than smaller galaxies (or building blocks). Simulations in a Λ-CDM framework demonstrate that the median metallicity of a stellar halo is primarily determined by the mass of the most massive building block (Robertson et al. 2005; Deason et al. 2016; Oman et al. 2017; D'Souza & Bell 2018b). An interesting pair of galaxies in this respect are the Milky Way and its close neighbour Andromeda. While the Milky Way has many features in common with our neighbouring galaxy Andromeda, one of the more striking differences is the metallicity of their stellar halos, namely that the Andromeda stellar halo is significantly more metal-rich out to large radii (e.g., Ibata et al. 2014). This could be linked to the presence of stellar substructure on vast scales throughout its large stellar halo and to M32 and its Giant Southern Stream (e.g., McConnachie et al. 2018; D'Souza & Bell 2018a). Additionally, a stellar halo mass-metallicity relation is also seen in stellar haloes of galaxies outside the Local Group (Harmsen et al. 2017).

To understand the merging history of the Galaxy in detail, it is thus of clear importance to measure the MDF of the Milky Way stellar halo at different radii and through different lines of sight, minimizing selection biases as much as possible. Such measurements will also help to understand the globular cluster population with respect to the field stars, and constrain chemical evolution in the early Milky Way. Unfortunately, the metal-poor end of the MDF in the Galactic halo remains especially poorly constrained. Several previous studies have used large spectroscopic surveys for this purpose, selecting a halo sample either kinematically to build up a local halo sample, by metallicity to study the metal-poor tail, or by stellar type and magnitude to study the halo at different distances (e.g., Ryan & Norris 1991; Schörck et al. 2009; Allende

Prieto et al. 2014). These different approaches result in very different distributions, especially in the very metal-poor regime. One difficulty is to understand all of the selection biases that affect the spectroscopic survey selection and correct for them appropriately. Especially further out and at lower metallicities, spectroscopic samples quickly become sparser. Additionally, distances are difficult to constrain – even in the era of the Gaia mission – as the Galactic halo stretches out far beyond the regime where parallax information is reliable.

Until recently, observational studies constraining the metal-poor tail of the halo MDF have mostly been done using spectroscopic samples, since photometric metallicity calibrations have been limited to [Fe/H] > -2 (e.g. Ivezić et al. 2008). However, several recent studies have shown that it is possible to provide reliable photometric metallicities using SDSS broad-band photometry down to [Fe/H] $= -2.5$, if deeper u-band photometry is used (An et al. 2013, 2015; Ibata et al. 2017b; Zuo et al. 2017). Subsequently, this metallicity limit has been pushed even further with the implementation of narrow-band filters targeting the Ca II H & K absorption lines in surveys such as *Pristine* and *SkyMapper* (Keller et al. 2007; Casagrande et al. 2019), which have shown that with this technique it is possible to derive metallicities down to the regime of [Fe/H] < -3 (Starkenburg et al. 2017; Da Costa et al. 2019).

In this chapter, we present MDFs using the *Pristine* survey (Starkenburg et al. 2017). *Pristine* uses a narrow-band filter to obtain photometric metallicities for millions of stars over large regions of the norther Galactic halo. We convolve its results with spectroscopic follow-up information from thousands of stars to remove spurious features that may compromise the photometric metallicity determinations. Because of its photometric nature, the *Pristine* survey is (almost) completely unbiased in its selection of stars, although its metallicity determination is most reliable in the metal-poor regime ([Fe/H] < -1). We therefore focus our work on the derivation of the MDF at low metallicities for a sample of stars that has been selected to be consistent with halo ages and distances.

We describe the photometric survey, as well as the spectroscopic samples used to rescale the MDF in Section 4.2. In Section 4.3, we describe the selection of main sequence turnoff (MSTO) stars and a halo sample, and show the resulting MDFs for different distance ranges. In Section 4.4, we apply corrections to the photometric MDF, to account for uncertainties in the photometric metallicities and metallicity biases induced by the colour cut used to select the halo sample. We investigate substructures seen in the MDF as a function of observational direction in Section 4.5. We discuss the potential remaining biases from the halo sample selection in Section 4.6, and compare the metal-poor halo MDF to those derived in other works. In this section, we also discuss the possibility of the inner halo being dominated by a single, large accretion event, the duality of the stellar halo, and the implications that the halo MDF has for the lack of metal-poor globular clusters (GCs) in the Milky Way. In Section 4.7, we summarize and conclude the chapter.

4.2 Description of the data

4.2.1 Photometric sample

The *Pristine* survey and the derivation of photometric metallicities are described in detail in Starkenburg et al. (2017). In summary, *Pristine* utilizes a narrow-band filter that is nearly top-hat in shape and centered on the Ca II H & K absorption lines at 3900 – 4000 Å. These absorption lines are some of the strongest features in a stellar spectrum and their size is directly dependent on the metallicity of the star. The coverage of the survey is $\sim$ 2 500 deg^2 in the northern hemisphere and goes to a maximum depth of $g_0 \sim 21.5$ at S/N = 10, although to avoid differential depths of the observed fields and the resulting selection effects, we limit the sample to $g_0 < 20$ in this work. All detected sources from the *Pristine* survey are then

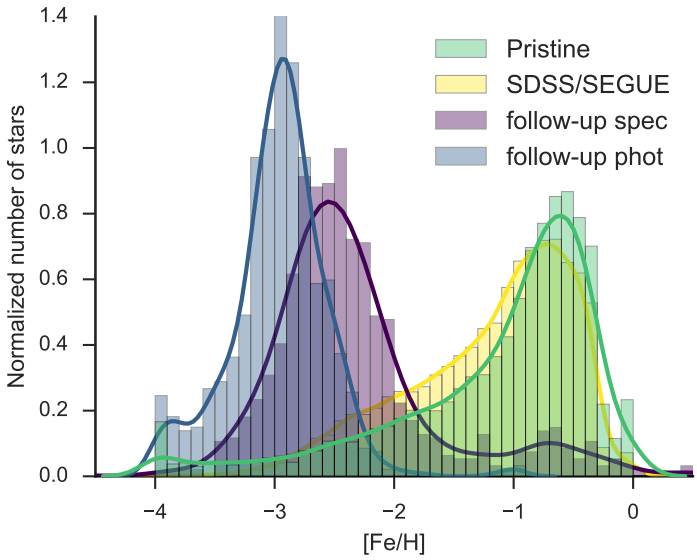

Figure 4.1 — Normalized distributions of metallicities from *Pristine* and the spectroscopic samples. The distribution of the *Pristine* spectroscopic follow-up sample is shown with photometric metallicities (blue) and spectroscopic metallicities (purple), along with the distribution of *Pristine* photometric metallicities for the matched SDSS/SEGUE sample (yellow) and the distribution of the photometric metallicities for the whole *Pristine* sample (green).

matched to stars in the Sloan Digital Sky Survey (SDSS, York et al. 2000). The metallicity of each star is determined using a colour-colour space combining the narrow-band information and broad-band information to solve for the temperature of the star (in Starkenburg et al. 2017 and in this work we use SDSS broad-band colours, but a similar principle can be followed using Gaia G_{BP} and G_{RP} colours, as demonstrated in Bonifacio et al. 2019). Our method is calibrated through the use of overlapping SDSS/SEGUE spectra as well as some stars in the footprint that were additionally observed with high-resolution spectroscopy. Overall, we find that we can determine metallicities for F, G, and K type stars with 4 000 K < $T_{\rm eff}$ < 7 000 K, 0.15 < $(g_0 - i_0)$ < 1.5) and −3.0 < [Fe/H] < −0.5 with ∼ 0.2 dex precision (Starkenburg et al. 2017). Spectroscopic follow-up has demonstrated that *Pristine* is also successful at identifying stars with [Fe/H] < −3.0 (see Chapters 2 and 3). The raw photometric MDF for all stars in the current *Pristine* footprint is shown as the green histogram in Figure 4.1.

4.2.2 Spectroscopic sample

In addition to the photometric component, the *Pristine* survey includes a large spectroscopic follow-up campaign targeting metal-poor star candidates (Caffau et al. 2017; Bonifacio et al. 2019; Venn et al. 2019, Chapter 3). Although the efficiency of the *Pristine* survey to discover extremely metal-poor stars is very good (for more details see Chapters 2 and 3) and we take precautions to clean the sample of outliers such as variable stars (using Hernitschek et al. 2016) and white dwarfs (through additional colour cuts), the metal-poor tail of the MDF is still, to some degree, contaminated by more metal-rich stars. This is to be expected, as the metal-poor population is much less numerous than the more metal-rich stars so that stars of higher metallicity are more likely to scatter down to lower metallicities than the other way

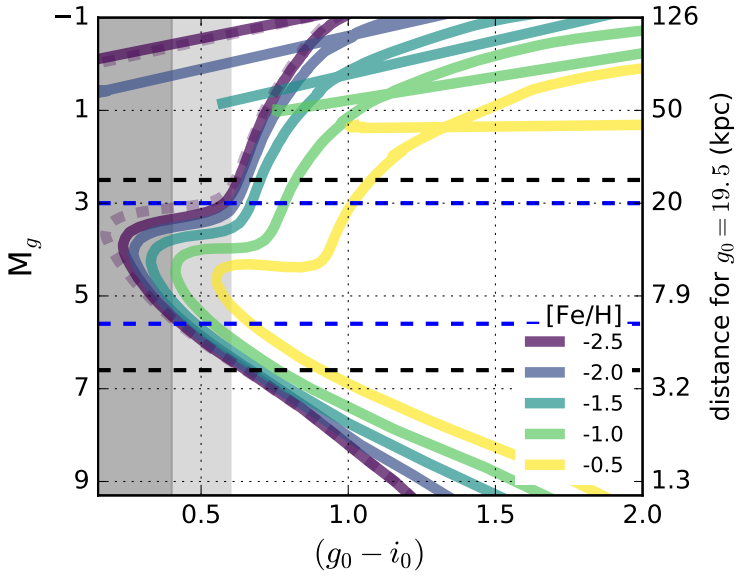

Figure 4.2 — The MSTO colour selection and corresponding distances. PARSEC isochrones of decreasing metallicity and an age of 12 Gyr are plotted as solid lines, and the [Fe/H] = −2.5 isochrone with an age of 10 Gyr is shown by the dotted line. The absolute SDSS g magnitude is shown on the left axis, and the corresponding distances at an apparent magnitude of $g_0 = 19.5$ on the right axis. The grey shaded region shows the colour cuts used to select the MSTO stars and the dark shaded grey region shows the narrower colour cut used to select the halo sample. The black and blue dashed lines show the corresponding magnitude ranges spanned by the MSTO and halo selections, respectively.

around, assuming similar uncertainties. In Chapter 3, we estimate this contamination to be ∼ 18% for stars that are predicted to be extremely metal-poor ([Fe/H] < −3) with photometry, but that are found with spectroscopy to be metal-rich ([Fe/H] > −2). It might also be the case that some unidentified contamination remains, if stars misclassified by photometry are also misclassified by the medium-resolution follow-up spectroscopy, although these cases are likely to be rare. Figure 4.1 shows the raw MDFs, where the blue and purple histograms show the spectroscopic follow-up sample of *Pristine* metal-poor candidates with their predicted photometric metallicities and derived spectroscopic metallicities, respectively. The purple histogram is shifted, showing that the spectroscopic metallicities are more metal-rich than the predicted photometry on average (this effect has been previously described in Starkenburg et al. 2017), and also has a tail that extends to high metallicities. For the purpose of building a representative MDF down to the lowest metallicities from photometry, it is important to take these misclassifications into account and correct for them.

4.3 Selection of the halo sample

The *Pristine* survey is a photometric survey, and as such observes every star in the footprint that is in the right magnitude range (excluding holes caused by chip gaps and stars occluded by bright foreground stars, etc.). However, for each line-of-sight, the total sample will consist of a mixture of stars in different Galactic components, and of different evolutionary stages that probe a large variety of distances at the same magnitude. As it is our goal to produce an MDF

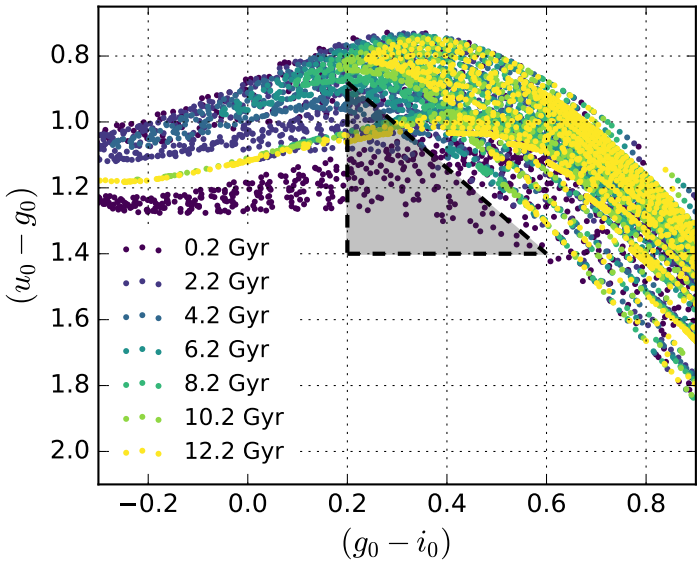

Figure 4.3 — Colour-colour plot of PARSEC isochrones sampling ages of 0.2−12.2 Gyr and a metallicities of $-2.5 <$ [Fe/H] < -0.5. The isochrones are coloured by age, but if instead they were coloured by metallicity, the metal-poor populations would be at the top of the plot and the metal-rich ones at the bottom. The grey shaded triangle shows the region removed (see text for details), which targets young disc stars and old horizontal branch stars.

of the Galactic halo, we need a way to separate out halo stars from the far more numerous nearby disc stars. This can in principle be done kinematically, chemically, or spatially. In this work, we choose to select a halo sample based primarily on distance from the Galactic plane. The selection of these distance samples and the Galactic environments which they probe are discussed in this section.

4.3.1 Distance slices into the Galaxy using MSTO stars

To obtain heliocentric distances for our sample, we make a colour selection to target MSTO stars. Figure 4.2 shows isochrones of different metallicities and a fixed age of 12 Gyr, produced using the PAdova and TRieste Stellar Evolution Code (PARSEC, Bressan et al. 2012). We convert [M/H] to [Fe/H] for these isochrones assuming [α/Fe] = +0.4. The dotted purple line shows the isochrone with [Fe/H] = −2.5 and an age of 10 Gyr, to show how the turnoff shifts as a function of age. We therefore make a colour selection such that we select the full MSTO from the tip of the turnoff up to the sub-giant branch for the most metal-poor isochrone, giving a colour cut of $0.15 < (g_0 - i_0) < 0.6$, shown by the combined dark and light grey regions in Figure 4.2. The absolute magnitude range spanned by this selection is shown by the black dotted lines, and corresponds to a heliocentric distance range of $\sim 4 - 25$ kpc for a magnitude of $g_0 = 19.5$ (the middle of the faintest magnitude bin $19 < g_0 < 20$). By selecting different magnitude bins, we separate the Galaxy into different distance ranges, allowing us to target different Galactic environments.

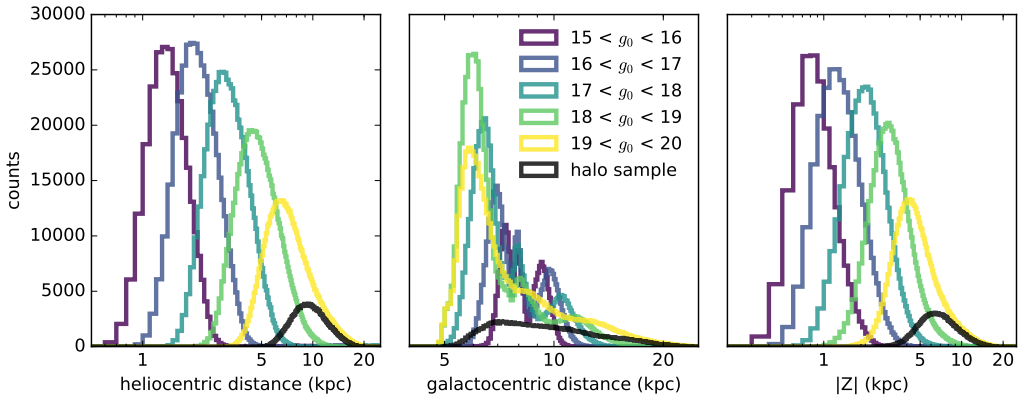

Figure 4.4 — Distance distributions for the MSTO cuts with $0.15 < (g_0 - i_0) < 0.6$ applied to the Galaxia mock catalogues. The left panel shows the heliocentric distances, the middle panel the Galactocentric distances, and the right panel the height above/below the disc, |Z|. Histograms are coloured according to the magnitude range shown, and the black histogram shows the distance distribution for the halo sample with $0.15 < (g_0 - i_0) < 0.4$ and $19 < g_0 < 20$.

4.3.2 The halo sample

Although the faintest MSTO sample selected with $0.15 < (g_0 - i_0) < 0.6$ and $19 < g_0 < 20$ is likely composed of mostly halo stars, it still probes a wide distance range of $\sim 4-25$ kpc, and thus a wide range of Galactic environments. In order to narrow this range, we impose an additional colour cut at $0.15 < (g_0 - i_0) < 0.4$ (depicted by the dark grey region in Figure 4.2) to more precisely target the tip of the MSTO. The blue dotted lines show the narrower magnitude range spanned by this sample, which corresponds to a narrower distance range of $\sim 6-20$ kpc. Therefore, this sample does not probe as far out into the halo as the faintest MSTO sample, but has a larger minimum distance and thus contains less of a contribution from solar neighbourhood and disc stars.

In a further attempt to clean the sample of young disc stars, we identify a region in the $(u_0 - g_0)$ versus $(g_0 - i_0)$ diagram that is dominated by young stars. Figure 4.3 shows this colour-colour space with PARSEC isochrones of metallicities ranging from $-2.5 <$ [Fe/H] < -0.5 and ages from $0.2-12.2$ Gyr. The isochrones are colour coded by age, but colour coding them instead by metallicity would show that the more metal-poor isochrones would be at the top of the diagram and the more metal-rich ones would be at the bottom. The grey triangle shows the region that is almost solely populated by young, metal-rich stars. In the upper left of this triangle there are also some older populations, but these are horizontal branch stars, which constitute a much smaller population compared to the more numerous MSTO stars. We therefore decide to remove all stars that fall within this triangular region from the analysis in both the spectroscopic and photometric data sets to clean the young disc dwarfs from the sample. The underlying assumption is that there are far more young stars in the disc than young metal-poor stars in the halo. This cut removes $\sim 10\%$ of the halo sample selected with *Pristine* stars and only $\sim 3\%$ of the spectroscopic sample.

4.3.3 Distance ranges of the samples

To further investigate the distance ranges probed by these samples, we generate a simulated data set using the Galaxia code (Sharma et al. 2011), which implements the analytical Besançon

Table 4.1 — Heliocentric distance ranges probed by the MSTO samples selected with $0.15 < (g_0 - i_0) < 0.6$ and the shown magnitude cuts. The bottom row shows the heliocentric distance range of the halo sample, selected with $0.15 < (g_0 - i_0) < 0.4$ and $19 < g_0 < 20$. Distance ranges are computed both by using isochrones as well as the distances from the Galaxia mock data. The bold numbers indicate the peaks of the distributions in the left panel of Figure 4.4.

Magnitude range	Isochrone distance (kpc)	Galaxia distance (kpc)
$15 < g_0 < 16$	[0.6 - 4]	[0.8 - **1.5** - 3]
$16 < g_0 < 17$	[1 - 6.3]	[1 - **2** - 4.5]
$17 < g_0 < 18$	[1.5 - 10]	[1.6 - **3** - 7]
$18 < g_0 < 19$	[2.4 - 15.9]	[2.5 - **4.5** - 11.2]
$19 < g_0 < 20$	[3.8 - 25]	[3.5 - **6.5** - 18]
halo sample	[6 - 20]	[6 - **9.5** - 16]

model (Robin et al. 2003) of the Galaxy. We run Galaxia using the default parameters, and selecting the regions of the *Pristine* footprint with a magnitude limit of $14 < g_0 < 22$, an oversampling factor of 3.0, and disc flare and warp turned on. We mimic the survey depth and observational selection function of the *Pristine* survey by selecting a subset of stars from the Galaxia catalogue to match the stellar density of *Pristine*. This is done using SDSS as a reference sample, under the assumption that it is complete at the magnitude limit of the *Pristine* survey ($V < 21.5$). We then bin the stars spatially (by each 1 x 1 deg^2 *Pristine* field), by colour (bins of 0.15 in $(g_0 - i_0)$) and by magnitude (bins of 0.2 in g_0) and count the number of stars in each bin. A random subset of Galaxia stars is selected from a given bin, such that the ratio of the subset to the total is equal to the ratio of *Pristine* stars to SDSS stars in that given bin (i.e. $N_{select}/N_{Galaxia} = N_{Pristine}/N_{SDSS}$ for each bin in RA and Dec, colour, and magnitude). Since the *Pristine* sources are cross-matched to SDSS, there are no bins in which $N_{Pristine} > N_{SDSS}$. In addition, the oversampling of Galaxia with fsample = 3 minimizes the cases where $N_{select} > N_{Galaxia}$, and results in very few bins in which this is the case. When this does happen, we simply take all of the Galaxia stars in that bin. The result of this is a Galaxia generated mock sample which has the same coverage as the *Pristine* footprint, and mimics the depth, magnitude and colour distributions of *Pristine* .

We then apply the same selections for the MSTO and halo samples and compute the distance distributions for each in the Galaxia mock catalogue. Figure 4.4 shows the heliocentric, Galactocentric, and height above the disc distance distributions for each of the samples, and Table 4.1 summarizes the specific heliocentric distance ranges probed, computed with isochrones and with the Galaxia catalogue. Both methods yield similar distances, particularly at the lower limit of the distance range, but the isochrone technique has slightly higher upper limits due to the extension of the beginning of the sub-giant branch, which in Galaxia only represents a small number of stars in the tail of the distributions. Using heliocentric distance ranges can sometimes make it difficult to interpret exactly which Galactic environments are being probed, however, as can be seen in the middle and right panels of Figure 4.4, the high Galactic latitudes ($|b| > 20°$) and limited depth of the *Pristine* sample result in a distance distribution that still nicely samples the inner Galactic halo when converted to the reference frame of the Galactic center. Most importantly, the halo sample probes a distance range of $\sim 5-20$ kpc from the Galactic center, and a minimum height above the disc of $|Z| > 4$ kpc, avoiding in large part the stellar populations of the bulge and disc.

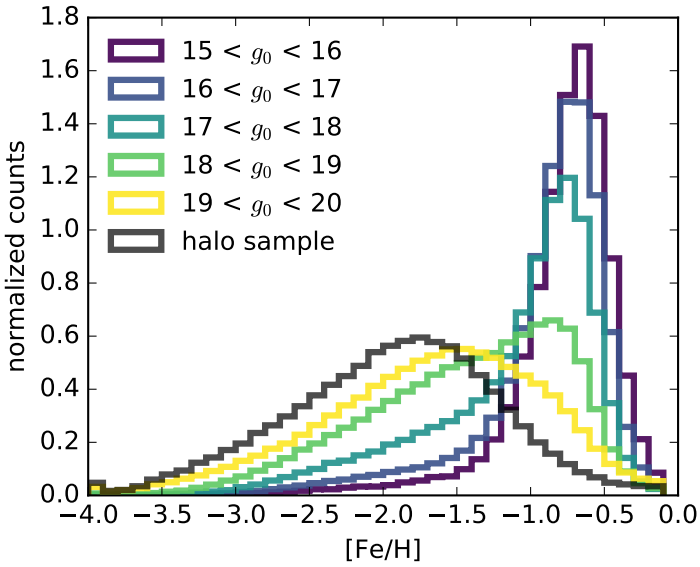

Figure 4.5 — MDFs of MSTO samples selected with a colour cut of $0.15 < (g_0 - i_0) < 0.6$, separated into bins of magnitude. The black histogram shows the halo sample with $0.15 < (g_0 - i_0) < 0.4$ and $19 < g_0 < 20$.

4.3.4 The MDF at different distances

Figure 4.5 shows the resulting MDFs for the MSTO samples at different magnitude ranges. These are raw photometric MDFs from *Pristine* data without any corrections applied, to qualitatively show the decreasing metallicity gradient with increasing heliocentric distance. The coloured histograms are produced with the $0.15 < (g_0 - i_0) < 0.6$ colour cut and limiting the samples to the shown magnitude ranges. The black histogram shows the halo sample with $0.15 < (g_0 - i_0) < 0.4$ and $19 < g_0 < 20$. There is a clear trend that the MDFs shift to lower metallicities with increasing distance. There is also a clear transition from the strong peak at [Fe/H] ~ -0.7 in the brighter samples where the disc dominates, and the metal-poor peak at [Fe/H] ~ -1.6 in the fainter samples where the halo population dominates. The green and yellow histograms in the magnitude range of $18 < g_0 < 20$ show the transition, where both populations are visible.

4.4 Correcting the MDF

For the current data set, there are two main corrections that we take into account in order to account for the metallicity biases present in the MDF. The first comes from determining metallicities photometrically and the inherent uncertainties associated with this procedure, particularly in the metal-poor regime, and the second are the biases introduced by making the colour cut of $0.15 < (g_0 - i_0) < 0.4$ to select the distance limited halo sample. The effect of these biases, as well as the methods used to correct them, are discussed in detail in this section.

4.4.1 Corrections for the photometric metallicities

Despite cleaning the photometric metallicity sample of contaminating objects such as white dwarfs and variable stars, some degree of contamination still persists, particularly at the metal-poor end. This is mostly due to metal-rich stars preferentially scattering into the metal-poor regime, but can also be caused by stars with unusual colours (e.g. blue stragglers, or stars with long-period variability not picked up in Hernitschek et al. 2016) or peculiar behaviour in their Ca II H & K lines (e.g. binaries, or stars with Ca II H & K in emission). Fortunately, most of these stars that are misclassified photometrically can be identified with the spectroscopic follow-up sample, and can subsequently be corrected for using a statistical rescaling of the photometric MDFs. The left panel of Figure 4.6 shows the relation between the *Pristine* photometric metallicities and the derived spectroscopic metallicities for the spectroscopic sample, which we assume to represent the true metallicity of the star. This sample consists of ~ 1 000 stars from *Pristine* spectroscopic follow-up, which predominantly cover the metallicity range of $-4 <$ [Fe/H] < -2.5 and are shown by the purple points (for more details on this sample see Chapter 3). To supplement the sample at [Fe/H] > -2.5, we add ~ 17 000 cross-matched SDSS/SEGUE stars over the *Pristine* footprint that have reliable metallicity determinations from the SEGUE Stellar Parameter Pipeline (SSPP, our selection of this sample is described in detail in Starkenburg et al. 2017), shown by the teal points in the Figure. For the 59 stars in common between the two samples, we find no systematic shift between the metallicities derived in the *Pristine* follow-up analysis and the SSPP, and a dispersion of ~ 0.3 dex. The middle panel shows the same samples after the application of the halo selection, with $0.15 < (g_0 - i_0) < 0.4$. We additionally apply a cut of $g_0 > 17$ to the SDSS/SEGUE sample to remove the strong disc feature at [Fe/H] > -1, and to allow for the distribution to have a larger spread to compensate for the larger uncertainties at faint magnitudes. We do not apply a magnitude cut of $19 < g_0 < 20$ since there are too few stars in the spectroscopic sample at these faint magnitudes. Similarly, we do not make a magnitude cut on the *Pristine* spectroscopic sample, since this sample is restricted to an even brighter magnitude range. Therefore, the implicit assumption that we make is that the relation between photometric and spectroscopic metallicity at [Fe/H] < -2.5 does not vary significantly between stars with $14 < g_0 < 18$ and $18 < g_0 < 20$. We also apply the $(u_0 - g_0)$ vs $(g_0 - i_0)$ cut to remove young disc stars as described in Section 4.3.2, but this only removes ~ 3% of the stars from the spectroscopic samples. The sizes of the samples after the cuts are applied are 1163 and 340 for the SDSS/SEGUE and *Pristine* samples, respectively.

To perform the rescaling, we build a probability density function (PDF) to describe the relation between the photometric and spectroscopic metallicities. We choose to use a Gaussian mixture model (GMM) for this purpose, because Gaussians are simple functions that are well suited to fitting smooth, continuous data.

We make the GMM using the python package *sklearn.mixture.GaussianMixture* (Pedregosa et al. 2011), which implements the expectation-maximization algorithm for fitting distributions composed of several Gaussian components, the number of which are provided as an input parameter. We use the default parameters, with the exception of increasing the maximum number of iterations to 500 to ensure that the solution converges.

To determine the number of components to use to best fit the data we use the Bayesian information criterion (BIC). We compute GMMs for each of the samples using between $1 - 10$ components, and then compute the BIC for each using the built-in attribute of the sklearn mixture model package. We repeat this 1 000 times, taking the mean of the resulting distribution of test values, and then choose the number of components with the lowest computed BIC. For the halo sample, this gives a solution with 3-components, and the resulting GMM is shown in the right panel of Figure 4.6. All three of the components have similar weights,

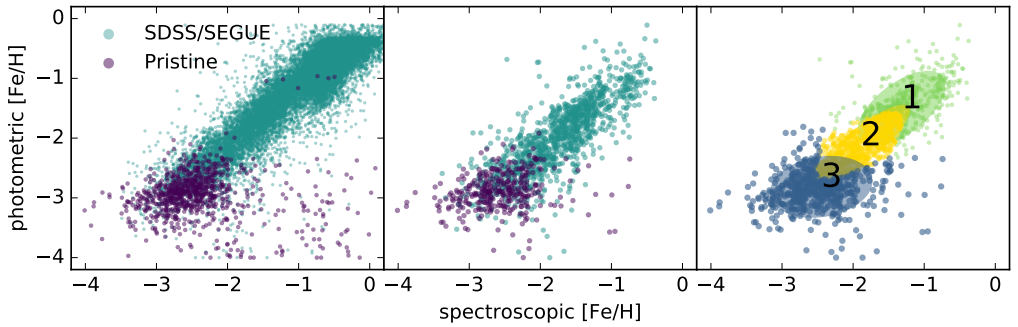

Figure 4.6 — The spectroscopic sample used to rescale the *Pristine* photometric MDF for the halo sample. The left panel shows the full spectroscopic samples from *Pristine* follow-up (purple) and SDSS/SEGUE (teal) and the middle panels shows the samples after applying the colour cut of $0.15 < g_0 < 0.4$ and a magnitude cut of $17 > g_0$ to select the halo sample. The right panel shows the corresponding 3-component GMM used to fit the data. The points are coloured according to which of the Gaussian components they belong, and the 1-σ contour of each is shown by the over-plotted oval of the same colour. The varied point sizes are simply to enhance the clarity in crowded regions.

such that the GMM is not dominated by any one component. Table 4.8 in the Appendix shows the best fit parameters for each of the Gaussian components, including the means, weights, and covariance matrices.

To check that the GMM fits the data well, we use a 2D Kolmogorov-Smirnov (KS) test, which begins with the null hypothesis, H_0, that two samples are drawn from the same parent distribution. First, we draw a sample from the GMM and compute the KS statistic and p-value between this sample and the original spectroscopic one using a two-sample KS test [*]. We repeat this procedure 1 000 times and take the mean p-value for the set of iterations. We find $p > 0.1$ for a GMM with 3 components, meaning that we cannot reject H_0 and the two samples are consistent with being drawn from the same distribution. If we use 2 components, then the resulting p-value is marginal and inconclusive, with $0.1 > p > 0.01$. Using only 1 component gives $0.01 >> p$, and H_0 is rejected. Taken together, these tests confirm that the GMM is an appropriate model for fitting a PDF to the spectroscopic sample, as long as the right number of components are used.

To apply the GMM correction, we take each photometric metallicity from the *Pristine* sample, and simply draw from the corresponding 1-D PDF made by sampling along a horizontal line at that given metallicity in Figure 4.6.

4.4.2 Corrections for the halo sample selection

For the colour cut used to select the halo sample, there is a metallicity bias that preferentially selects metal-poor stars over metal-rich stars, since the turnoff point of the MSTO is significantly redder for metal-rich populations than for metal-poor ones. This can be clearly seen by the shifting isochrones in the grey region of Figure 4.2, as a larger fraction of the isochrones fall outside of this region with increasing metallicity. In addition, younger populations will have a bluer turnoff than older ones. In order to investigate this effect and its influence on the MDF, we use PARSEC simulated stellar populations with a total mass of 10 000 $M_\odot$ across a range of metallicities to compute the mass fraction selected by these colour cuts to the total mass of the

[*]The code used for the 2D KS test was modified from `https://github.com/syrte/ndtest`

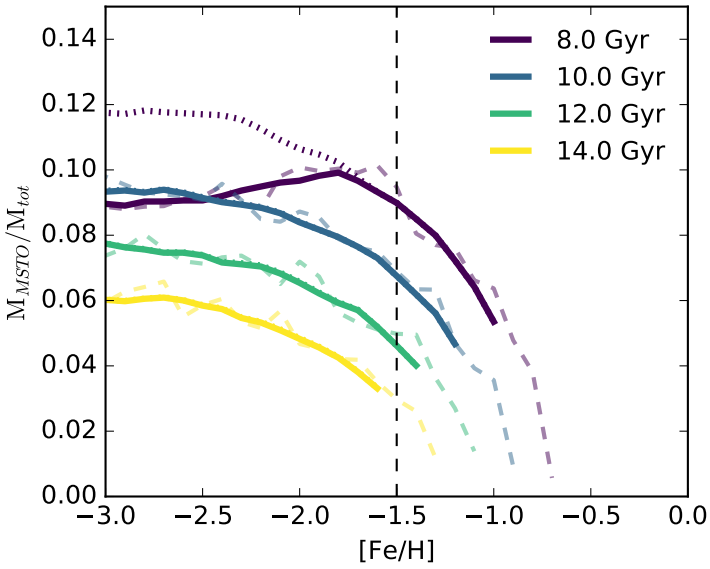

Figure 4.7 — Fractional contribution of the mass from the colour cut of $0.15 < (g_0 - i_0) < 0.4$ to the total mass of the simulated stellar population. The darker lines show the smoothed contours and the lighter dashed lines show the true values in each 0.1 dex metallicity bin. The dotted lines show the mass ratio for a colour cut of $0 < (g_0 - i_0) < 0.4$, and shows the degree of truncation for the MSTO as it extends below $(g_0 - i_0) = 0.15$.

stellar population. Figure 4.7 shows the metallicity dependence for the $0.15 < (g_0 - i_0) < 0.4$ colour cut, for a series of ages. The lighter dashed lines show the true values for each bin, and the darker lines are smoothed with a mean kernel and a width of 7 bins. The dotted lines (only visible for the 8 Gyr population) represent a colour cut of $0 < (g_0 - i_0) < 0.4$, and show the result when the entire MSTO is included at all ages.

For the 8 Gyr simulated population we see the effect of a the tip of the MSTO falling outside of our selected colour range ($g_0 - i_0 < 0.15$), shown by the deviation of the solid purple line from the dotted line at [Fe/H] < -1.7 in Figure 4.7. For older ages the MSTO is not truncated and the dotted lines lie underneath the solid lines. Therefore, even though we are not able to reliably assign metallicities to stars bluer than $(g_0 - i_0) = 0.15$ with *Pristine* , this will only result in a (small) bias at low metallicities even if there is a significant population of the halo that is younger than 10 Gyr. Furthermore, at all ages the shape of the curves in Figure 4.7 show that the relative mass selected decreases with increasing metallicity, as this colour selection preferentially selects the (blue) MSTO stars of metal-poor populations. At the most metal-poor end the trend flattens, as the isochrones begin to pile up on each other, i.e. the colour selection is largely unbiased at [Fe/H] < -3. On the right side of the plot, the metal-rich MSTO populations are increasingly suppressed as they fall outside of the colour selection, i.e. $0.4 < (g_0 - i_0)$. Table 4.8 in the Appendix gives correction factors for this effect at each metallicity bin. The factors are computed for each age separately by scaling each bin up to the value at [Fe/H] $= -3$. The correction we apply to the halo sample is the mean of the 10 and 12 Gyr populations, and is given in the rightmost column of Table 4.8.

When making the MSTO cuts we also investigate whether using redder broad-band colours to make the MSTO selections could reduce the correction factors, since isochrones of different metallicity do not show as strong of a colour difference as they do with bluer colours.

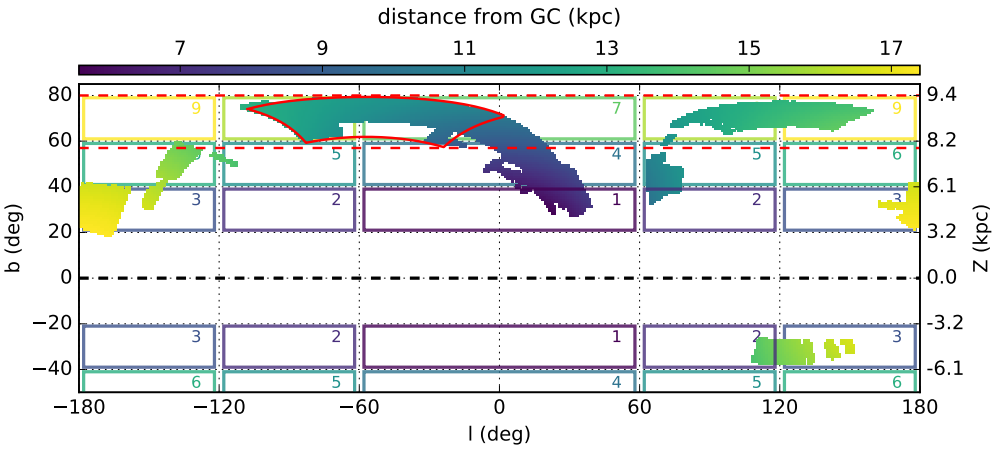

Figure 4.8 — The ~ 2 500 deg^2 *Pristine* footprint, colour coded by the distance to the Galactic center for the halo sample, assuming a heliocentric distance of 9.5 kpc. The coloured squares and the corresponding numbers separate the footprint by latitude and longitude, and each coloured region corresponds to the MDF of the same colour in Figure 4.10. The region outlined in red shows the region selected to contain the Virgo overdensity, and the red dashed lines are the equal latitude comparison sample.

Recomputing Figure 4.7 for $(i_0 - z_0)$, we find that the curves are significantly flatter in the metal-rich regime, indicating a decrease in the metallicity bias when using these colours instead of $(g_0 - i_0)$ colours. However, at the limiting magnitude of $g_0 = 20$, uncertainties in the g magnitudes are not larger than 0.03, while uncertainties in the z magnitudes in SDSS can be as large as 0.15. Given that in $(i_0 - z_0)$ space the MSTO is more compressed, these uncertainties are equal in size to the entire width of the MSTO, and are therefore too large to make a reliable MSTO selection. For this reason, we choose to use the $(g_0 - i_0)$ colours to select the MSTO.

Finally, this colour range also includes horizontal branch (HB) stars (shown in the top part of Figure 4.2), which are significantly brighter and more distant than the turnoff stars. Using the PARSEC simulated populations, we calculate the mass contribution of HB stars compared to the mass at the rest of the turnoff, and find that in the colour range of $0.15 < g_0 < 0.4$, the relative contribution of HB stars is at most ~ 0.2% for the 8 Gyr population, and less than this for the older populations. Therefore, we conclude that the HB stars constitute only a very small amount of the sample and subsequently that distance cross contamination between bins is small, resulting in no noticeable effect on the computed MDFs.

4.5 Substructure in the halo MDF

4.5.1 The MDF in different directions

We would like to further assess the impact of observational direction on the resulting MDF of the sample. Figure 4.8 shows the footprint of the *Pristine* survey on the sky, coloured by distance to the Galactic center. On the left hand y-axis is the Galactic latitude, and on the right hand axis the corresponding distance above or below the Galactic plane (Z = d $\cdot sin$(b)). The distances shown are computed to the approximate peak of the halo sample distance distribution of 9.5 kpc, as shown by the black histogram in the left panel of Figure 4.4. To convert from heliocentric distances to Galactocentric, we use a solar position of $(X, Y, Z) = (8.2, 0, 0.011)$ kpc. We separate the sky into 9 regions, made by the intersection of

3 regions in longitude and 3 regions in latitude. The longitude regions are each 60 degrees, and are mirrored across the Galactic center, and the latitude regions are each 20 degrees and are mirrored across the Galactic plane. The 9 regions are shown by the coloured squares and labeled by number in the Figure. Figure 4.10 shows the corresponding MDFs for the halo sample in each of these regions, coloured according to the corresponding region in Figure 4.8 and with the number of stars making up the sample also listed in the top left corner. These are plotted over the grey, filled histogram representing the full halo sample for comparison. We see that for most locations in the footprint, the MDF is invariant, meaning that the relative contribution of the disc and halo does not change with varying latitude or longitude, suggesting that the halo cuts were effective in selecting a pure halo sample. Additionally, it shows that the halo is relatively well-mixed in metallicity at these distances. Regions 2, 5, 6 and 9 show some small scale deviations from the total distribution, but these are all regions which have relatively small numbers of stars and thus these features may be a result of insufficient statistical sampling rather than real substructures. In region 3, however, we see a significantly stronger contribution at the metal-rich end, suggesting that at low galactic latitudes there is an additional contribution from a higher metallicity population.

4.5.2 The Monoceros ring

Region 3 is at relatively low Galactic latitude and in the direction of the Galactic anticenter. In this observational direction, a substantial amount of substructure has been previously observed, including the Monoceros Ring (Newberg et al. 2002), A13 (Sharma et al. 2010), TriAnd (Majewski et al. 2004; Rocha-Pinto et al. 2004), the Anti-Center Stream (ACS) and the Easter Banded Structure (EBS) (Grillmair 2006; Grillmair et al. 2008; Li et al. 2012). Although these structures were originally thought to be disrupted debris from accreted satellite galaxies, some recent studies have shown that some or all of these structures may be the result of kicked-up disc material, given that their chemistry and kinematics are consistent with a disc origin (e.g. Li et al. 2017; Bergemann et al. 2018; Sheffield et al. 2018). The peak of the distance distribution of the halo sample coincides with the distance to these structures (d $\sim$ 10 kpc), suggesting that these are indeed responsible for the observed metal-rich component in the MDF. Despite being able to differentiate this structure from the rest of the halo based on its metallicity distribution, we cannot from the MDF alone constrain whether this indeed is consistent with kicked-up disc material or debris from an accreted satellite.

To further investigate this metal-rich feature, we separate region 3 into high latitude ($|b| > 30°$) and low latitude ($|b| < 30°$) stars. The MDFs of each of these latitude ranges are shown in the bottom left panel of Figure 4.10. From this, we see that for stars with $|b| > 30°$, the metal-rich component predominantly disappears, and the MDF is consistent with that of the rest of the footprint regions. We therefore choose to cut out this region from the halo sample and keep only the stars with $|b| > 30°$, to minimize potential contamination of metal-rich stars from the disc and its associated structures.

4.5.3 The contribution of the disc

We investigate the possible degree of disc contamination in the halo sample further by using the Galaxia mock catalogue described in Section 4.3.3. With this model, we can identify which component of the Galaxy a given star belongs to. Figure 4.9 shows the relative contributions of the thin disc, thick disc, and halo for each of the MSTO selected samples. For the brighter samples with $g_0 < 19$, the sample is dominated by disc stars, and for the $19 < g_0 < 20$ sample there are slightly more halo stars than disc stars. For the halo sample, the contamination from the disc decreases significantly as compared to the faintest MSTO sample. This is due to the

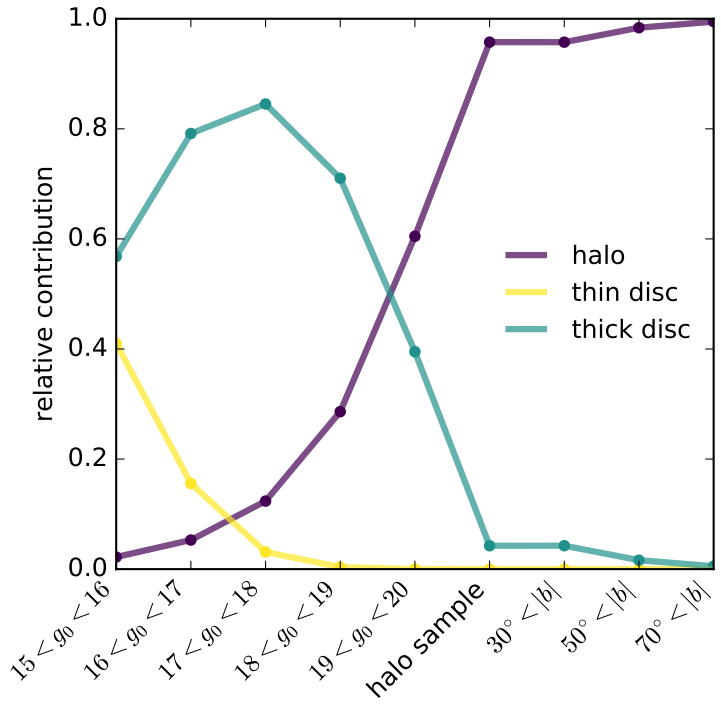

Figure 4.9 — Relative contribution of the thin disc, thick disc and halo for the different distance samples when selected in the Galaxia mock catalogue. The magnitude ranges refer to the MSTO samples, with the colour cut $0.15 < (g_0 - i_0) < 0.6$, and the halo sample refers to the sample made with $0.15 < (g_0 - i_0) < 0.4$ and $19 < g_0 < 20$, as described in the text. The latitude cuts are applied to the halo sample.

greater minimum distance and also due to the removal of metal-rich stars with the narrower colour cuts when selecting the MSTO. Adding latitude cuts further increases the purity of the halo selection until at a latitude of $70° < |b|$ the sample consists of almost entirely halo stars. Although this model represents a simplified version of the Galaxy, with an axisymmetric disc devoid of perturbations and substructures from interactions with accreted satellites, this still offers insight into the relative contributions of the selections to the purity of the sample. Therefore, due to the distance range probed by the selected halo sample, the contribution of foreground metal-rich disc stars is likely to be low, although there may still be some component of high latitude, kicked-up disc material, as discussed in Section 4.5.2.

4.5.4 The Virgo overdensity

In addition to investigating the effect of direction in Figures 4.8 and 4.10, we also take a closer look at the largest Galactic substructure in our footprint that is probed by our distance range: the Virgo overdensity. Although the specific nature and origin of this structure is still not well understood, it is known that the Virgo overdensity is located at a distance of 8 - 25 kpc (e.g., Lokhorst et al. 2016), and that it is likely comprised of overdense regions of halo stars. Located at approximately RA = 190° and Dec = -5° and spanning a large region of sky of over 1 000 deg^2, a significant region of the overdensity is located within the *Pristine* footprint. The current data set is therefore well suited to exploring the MDF of this structure. Although the Sagittarius

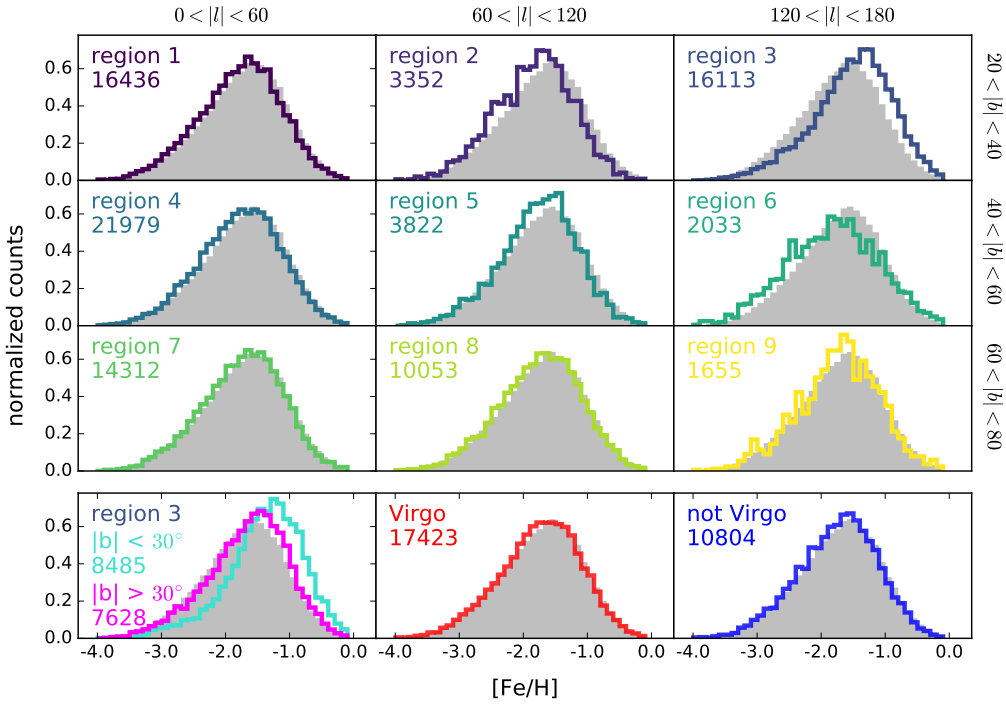

Figure 4.10 — MDFs for the 9 different regions described in the text, with colours that correspond to the numbered regions depicted in Figure 4.8. The grey filled histogram is the MDF of the total, GMM corrected halo sample for comparison. The bottom left panel shows the MDFs for region 3, separated into a high and low latitude sample (fuchsia and cyan, respectively), and the other two panels in the bottom row show the region selected to contain Virgo in red (the selected region is outlined in Figure 4.8), and the rest of the stars at the same Galactic latitude as a comparison are shown in blue.

stream is also prominent over our footprint, its vast leading and trailing arms are too distant to be sufficiently seen in the sample presented in this work (see for example Law & Majewski 2010). We therefore select a region around the Virgo overdensity, and compare its metallicity distribution to another region of the *Pristine* footprint at the same Galactic latitude that does not include Virgo. The region selected to contain Virgo is 180° < RA < 210° and -1° < Dec < 20°, and is demarcated in red in Figure 4.8, and the comparison sample is shown by the red dashed lines. The MDFs of these regions are shown in the bottom middle and bottom right panels of Figure 4.10. Although there are minor fluctuations from the underlying halo sample, these are not significant enough to draw any conclusions. Therefore, this indicates that the Virgo overdensity metallicity is not very different from that of the general stellar halo at these Galactic latitudes.

4.6 Discussion

So far, we have produced an MDF for a carefully selected halo sample, with $0.15 < (g_0 - i_0) < 0.4$, $19 < g_0 < 20$, $|b| > 30°$ and a cut in $(u_0 - g_0)$ vs $(g_0 - i_0)$ space to remove young disc stars. We calculated with isochrones and a Galaxia model that this sample spans $\sim 6-20$ kpc, a distance range which nicely probes the inner Galactic halo. This sample is not free of selection biases, as

we discussed in Section 4.4.2, but these are mostly present at the metal-rich end, and constitute moderate corrections that we have applied at [Fe/H] < -1.5.

The slope of the metal-poor regime, particularly at [Fe/H] < -1.5, is important in that it constrains the relative ratios of the most metal-poor stars in the Galaxy, which in turn encode the rate of chemical enrichment at early times. This is useful for building chemical enrichment models and for constraining the physical processes driving them. This also helps to better understand the role of metal-poor stellar populations in the formation of the Galactic halo, such as accreted dwarf galaxies and globular clusters.

4.6.1 Biases in the photometric metallicities

The *Pristine* survey is fairly unbiased in the stars that it observes. However, there are still various inter-dependencies on the colour, magnitude, and derived metallicities. We list the most important effects below, and discuss their possible effects on the MDF in the following section.

- *Dependence of the metallicity sensitivity on colour:* Figure 4.11 shows the overlapping sample of SDSS/SEGUE stars with the *Pristine* sample (using the same quality cuts on the *Pristine* photometry and the SDSS/SEGUE spectra as described in Section 3 of Starkenburg et al. 2017). The colour-coding by metallicity nicely illustrates the metallicity sensitivity of the survey. This figure also shows that at bluer colours the range of metallicities span a narrower range in $(CaHK_0 - g_0)$ than at redder colors, meaning that for the same uncertainties in photometry there will be a larger uncertainty in metallicity, and consequently a larger contamination across metallicity bins. In addition, the histogram in the bottom panel of Figure 4.11 shows that most of the stars are concentrated at $(g_0 - i_0) < 1.0$. The population redder than that is mostly dominated by metal-rich disc dwarfs.

- *Dependence on the depth of the survey on metallicity:* Due to the nature of the narrow-band *Pristine* CaHK filter, stars that are metal-poor will have smaller absorption lines and thus appear brighter in the photometry. This results in a slight difference in brightness for the most metal-poor stars as compared to the metal-rich ones. The middle panel of Figure 4.11 shows that on average, the more metal-poor stars will be $0.2 - 0.5$ mag brighter in the *Pristine* narrow-band filter than metal-rich stars. This will lead to the metal-poor stars probing a slightly larger volume. However, imposing the colour cut of $0.15 < (g_0 - i_0) < 0.4$ used to select the halo sample reduces this metallicity dependant magnitude discrepancy to $0.1 - 0.2$ mag, which is on the order of the uncertainty inherent in determining these distances with isochrones.

- *Dependence of measurement uncertainties on magnitude:* Another effect results from the uncertainties on the photometric flux being larger at fainter magnitudes. At the bright end, the uncertainties in both the CaHK and the SDSS broad-band magnitudes are small, but at the faint end $\sigma_{CaHK} >> \sigma_{SDSS}$, such that it is largely the CaHK uncertainty that drives the overall metallicity uncertainty. The photometric uncertainties as a function of magnitude for the sample are summarized in Table 4.2. The larger uncertainties at the faint end are partially mitigated with the GMM rescaling, especially since we construct the GMM based on a fainter sample, resulting in a model that takes into account the larger uncertainties in the spectroscopic data set.

- *Dependence of contamination fraction as a function of metallicity:* As previously mentioned in the introduction, there are far more metal-rich stars than metal-poor stars in

Table 4.2 — Uncertainties as a function of magnitude for the *Pristine* CaHK and SDSS broad-band magnitudes.

	median uncertainties		
Magnitude	g	i	CaHK
g = 15	0.003	0.004	0.004
g = 16	0.004	0.004	0.006
g = 17	0.005	0.005	0.01
g = 18	0.006	0.006	0.02
g = 19	0.01	0.01	0.05
g = 20	0.02	0.02	0.1

the halo MDF, meaning that more metal-rich stars will scatter into the metal-poor regime than the other way around, resulting in a relatively higher contamination fraction at low metallicities. This is largely corrected for by the GMM rescaling, but only in so far as metal-rich stars that scatter into the metal-poor end can be identified by medium-resolution spectroscopy.

4.6.2 Previous determinations of the halo MDF

The MDF of the halo produced from *Pristine* data is shown in Figure 4.12. The grey histogram is the raw *Pristine* MDF, the cyan is after rescalling with the GMM and the black histrogram is corrected with both the GMM and the correction factors for the colour cut from Table 4.8. Application of the GMM correction to the halo sample has two main effects on the MDF, namely that it shifts the peak of the distribution to slightly higher metallicities, and removes contaminants from the metal-poor tail and shifts them to higher metallicities. The corrections for the colour cut do not significantly change the slope of the metal-poor tail of the MDF ([Fe/H] < -2), but modestly enhance the relative numbers of stars in the bins at higher metallicities.

The right panel of Figure 4.12 shows a comparison of several MDFs from the literature that are presented in this chapter, and the left panel shows the same but with a log scale to better emphasize the differences in the slopes of the metal-poor regime. Each distribution is scaled to the *Pristine* distribution at [Fe/H] $= -2$, to allow for an easier comparison of the slopes. The purple, yellow and green histograms are the MDFs presented in An et al. (2015), Allende Prieto et al. (2014) and Schörck et al. (2009), respectively, and the blue dashed line is the MDF expected from the simple chemical enrichment model presented in Hartwick (1976), assuming an effective yield of $y_{\rm eff}$ = -1.6. The MDF of Ryan & Norris (1991) is not shown in this Figure, but is in good agreement with the Hartwick (1976) model. In this section, we first describe each MDF and how it was produced, and then discuss how it compares to our results.

The chemical evolution model of Hartwick (1976)

Using a sample of 60 halo globular clusters, Hartwick (1976) modifies a simple chemical evolution model (Searle & Sargent 1972; Pagel & Patchett 1975) to describe the MDF. Starting from a closed-box of primordial gas, this model is based on a simple parametrization of the rate at which the ISM is enriched in heavy elements by exploding stars. The model assumes instantaneous recycling and a constant IMF, and can be described using a single free parameter known as the effective yield, $y_{\rm eff}$. This parameter represents the ratio of the mass of heavy

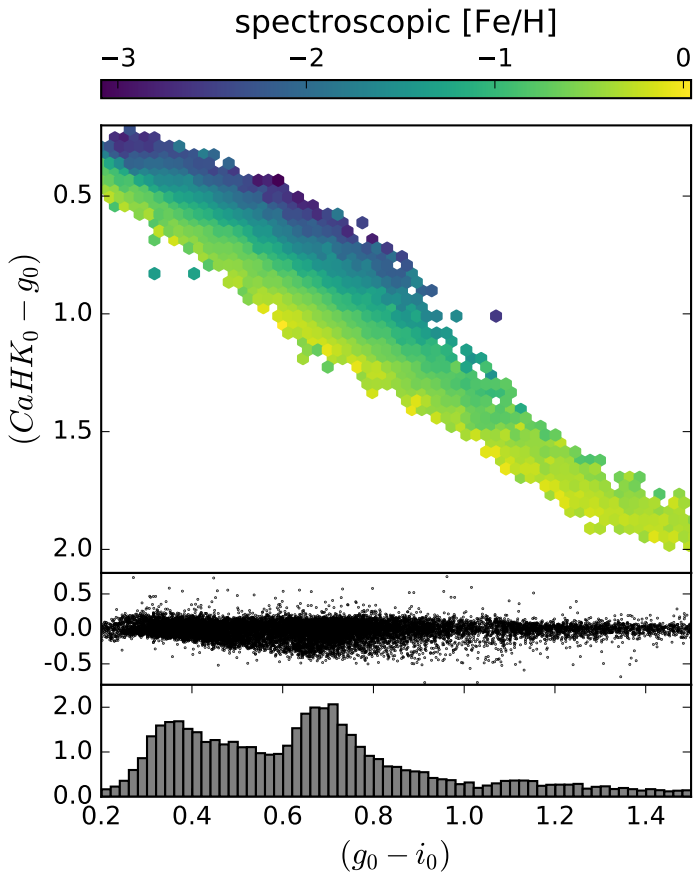

Figure 4.11 — Top panel: Colour-colour plot of the SEGUE spectroscopic sample, colour-coded by SEGUE metallicities. Middle panel: $(CaHK_0 - g_0)$ colour subtracted by the median value per $(g_0 - i_0)$ colour bin. Bottom panel: histogram of the distribution of the top panel.

elements ejected, to the mass locked up in long-lived stars or remnants, and also sets the location of the peak of the distribution. Following Ryan & Norris (1991) we choose $y_{\rm eff}$ = -1.6. Under the assumption that the halo globular cluster population traces the field star population, we can use this model to approximate the idealized MDF of the Milky Way halo.

SDSS sample (Allende Prieto et al. 2014)

Allende Prieto et al. (2014) use a sample of spectroscopically selected F-type stars from the SDSS survey (in particular those targeted in the BOSS survey as flux calibrators) to determine a metallicity distribution of the Milky Way halo. The MDF shown by the yellow histogram in Figure 4.12 is produced using a sample of ~ 16 000 F-type MSTO stars selected with $2.5 < \log g < 4.4$ and $g_0 > 17$. Most of the observed stars are at high Galactic latitudes, and they estimate them to probe a distance range from 5 kpc out to a few tens of kpc. They note that this sample is biased due to colour cuts used to select the BOSS spectrophotometric stars (summarized in their Table 1). To correct for these, they compute photometric colours for a grid of spectral energy distribution models covering a range of stellar parameters and identify which ones

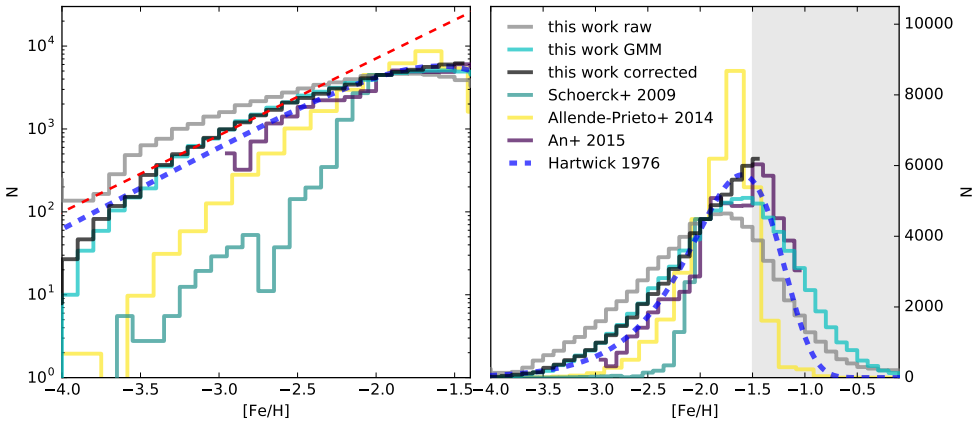

Figure 4.12 — Comparison of the *Pristine* MDF to previous works. The right panel shows the MDFs over the full metallicity range to illustrate the global shapes and location of the peaks, while the left panel shows only the metal-poor end and is plotted with a log scale to emphasize the differences in the metal-poor slopes. The grey shaded region shows where the corrections to the *Pristine* MDF become large and uncertain, at [Fe/H] > −1.5. The raw MDF of the *Pristine* halo sample is shown in grey, the GMM corrected in cyan, and the GMM/colour corrected in black. The MDFs from An et al. (2015), Allende Prieto et al. (2014) and Schörck et al. (2009) are shown in purple, yellow, and green respectively, and the Hartwick (1976) model is shown as the blue dashed line. The red dashed line is an exponential fit to the *Pristine* corrected sample in the metallicity range of −3.4 < [Fe/H] < −2.5.

make it through the colour cut and are in the temperature range of 5 600 < T_{eff} < 6 500 K. They then rescale the MDF to account for the missing populations, making the simplifying assumption that all stars are evenly distributed over the stellar parameter space. Although this slightly boosts the number of stars at the peak of the distribution at [Fe/H] = −1.6, it has very little effect on the tails of the distribution. The corrections are expected to be most significant for higher metallicities, as at these metallicities few stars with typical turnoff temperatures and gravities are selected in BOSS (as is also apparent from the results in Table 1 of Allende Prieto et al. 2014).

HES sample (Schörck et al. 2009)

Schörck et al. (2009) estimate the slope of the MDF at the metal-poor end by correcting the observed MDF from the HES survey for selection effects. For the correction, they create an MDF following Hartwick (1976), replace all the [Fe/H] values with pairs in their observables KP (an index on the Ca K absorption line following Beers et al. 1999) and color $(B-V)_0$, convolve them with Gaussian uncertainties, and assess the completeness in [Fe/H] vs $(B-V)_0$ space after applying the HES selection cuts. The result of this exercise is that they find HES to be essentially complete below [Fe/H] = −3.5, but the corrections become quite significant at higher metallicities. For example, at [Fe/H] = −2.05 they apply corrections of a factor of 10 − 80 (depending on colour). They build the MDF based on a sample of 1638 HES stars with available spectroscopic follow-up, and scale it up to include all 3 439 stars from the total candidate sample.

The sample of Ryan & Norris (1991)

Ryan & Norris (1991) present the metallicity distribution of the local halo from 372 subdwarfs selected in proper motion studies with space velocities relative to the LSR > 250 km s^{-1}. They find a very good fit to the Hartwick (1976) model using $y_{\rm eff} = -1.6$. The excellent fit does not include the 10% most metal-rich stars in the sample, but they speculate that this might be contamination from the disc. The model would predict 13 stars with [Fe/H] < −3, whereas 6 are observed, this difference is however too small to be significant for their sample size.

The sample of An et al. (2015)

An et al. (2015) derive a halo MDF photometrically from main-sequence turnoff stars observed in the extra deep SDSS stripe 82, and estimate that their sample covers a distance range of 5−10 kpc. They find excellent agreement with the work of Ryan & Norris (1991) mentioned above. In addition to a single Hartwick (1976) or Gaussian model fit, they also provide a double Gaussian fit to their derived MDF, finding a peak at [Fe/H] = −1.4 and a second one at [Fe/H] = −1.9. The extra Gaussian component is motivated by a difference in kinematics for the stars in both populations, where the second population shows a larger number of retrograde orbiting stars.

4.6.3 Comparison to the literature

The number of metal-poor stars

Figure 4.12 shows that there are distinct differences between the various MDFs in the relative numbers of stars they contain at [Fe/H] < −2. Most notably, the *Pristine* halo sample shows the largest number of stars, even more than the Hartwick (1976) model, in the range −3.4 < [Fe/H] < −2.2. The photometric sample of An et al. (2015) shows a pronounced bimodality, with the metal-poor component showing a similar abundance of stars to *Pristine* for [Fe/H] < −2.3, but it begins to drop off at [Fe/H] < −2.6, presumably as the metallicity calibration becomes less reliable. Both of these samples are in reasonably good agreement with the model. The SDSS/BOSS spectroscopic sample of Allende Prieto et al. (2014) shows a smaller contribution with respect to the photometric samples at [Fe/H] < −2, and the HES sample of Schörck et al. (2009) shows the lowest relative contribution of metal-poor stars of all of the shown MDFs.

From the left panel of Figure 4.12, we can see that the shape of the *Pristine* halo sample MDF is not linear in log space, and therefore is not well described by a single exponential curve. Instead, the observed distribution is much more suited to a fit with 3 separate exponentials, in the metallicity ranges of [Fe/H] < −3.5, −3.4 < [Fe/H] < −2.5, and −2.5 < [Fe/H] < −1.5. The best fit line in the metallicity range of −3.4 < [Fe/H] < −2.5 is shown as the red dashed line in the left panel of Figure 4.12, and has a slope of $\Delta(\mathrm{LogN})/\Delta[\mathrm{Fe/H}] = 1.0 \pm 0.1$. Below this metallicity range the slope becomes substantially steeper with a value of 2.0 ± 0.2.

Using a spectroscopic sample of metal-poor halo giants selected from the *SkyMapper* survey, Da Costa et al. (2019) derive a slope of the MDF of $\Delta(\mathrm{LogN})/\Delta[\mathrm{Fe/H}] = 1.5 \pm 0.1$ for a metallicity range of −2.75 < [Fe/H] < −4. The MDF they derive is well fit by a single exponential for the entire metallicity range, but drops off steeply at [Fe/H] = −4. Fitting a single exponential in the same way naively to the *Pristine* MDF also yields a slope of 1.5 ± 0.1, but as discussed above, this does a poor job of describing the distribution. Therefore, the slope in the metal-poor regime from the *Pristine* halo sample is somewhat flatter, consists of more components and has a drop off at a higher metallicity ([Fe/H] < −3.5) than the MDF described in Da Costa et al. (2019). The slopes of both of the spectroscopic samples are also steeper than the slope of the *Pristine* halo sample for [Fe/H] < −2, particularly for the Schörck et al. (2009) sample which shows the steepest slope in the metallicity range of −2.5 < [Fe/H] < −2 of all the MDFs.

Interestingly, the slope of the *Pristine* MDF between $-3.4 <$ [Fe/H] < -2.5 matches very well to the slope of the Hartwick (1976) simple chemical evolution model at [Fe/H] < -2. This breaks down at [Fe/H] < -3.5, however, where the *Pristine* MDF drops steeply. This also happens to be the metallicity below which the *Pristine* photometric metallicity determination becomes less reliable, and where there are very few stars in the spectroscopic sample with which to fit the GMM correction. Nevertheless, this steep decrease is seen in the other MDFs as well, which all show a drop between $-4 <$ [Fe/H] < -3.5, whereas the model does not. At these very low metallicities – and at the earliest epoch of star formation – chemical enrichment is dominated by a small number of supernovae explosions. Therefore, star formation is likely to proceed in a very inhomogeneous and stochastic way, deviating more strongly from the assumptions of the model that the ISM is well-mixed and homogeneous. Additionally, the model does not account for channels of enrichment from external sources, which could very quickly enrich the interstellar medium and result in a steeper slope at the lowest metallicities. In Schörck et al. (2009), they note that despite the discrepancy between the slope of their metal-poor MDF and the Hartwick (1976) model, their MDF better matches theoretical models which include a critical metallicity for low-mass star formation at $Z_{cr} = 10^{-3.4} Z_{\odot}$ (Salvadori et al. 2007). It should be noted, however, that Schörck et al. (2009) rely on rescaling the models to match their selection function. Nevertheless, this motivates the exploration of further modifications of these simplified models by incorporating more complex physics (e.g. gas accretion, outflow and inhomogeneous mixing) which may improve the agreement between the models and the data.

Selection effects and metallicity bias

Each of the MDFs discussed above are selected in different ways and from different samples. Ryan & Norris (1991) and Hartwick (1976) describe the local halo, whereas An et al. (2015), Schörck et al. (2009), Allende Prieto et al. (2014) and the current work all target the "distant" halo, by selecting stars at larger distances from the Galactic plane, albeit each exploring a slightly different distance range and thus different Galactic environments. More importantly than just the differences of environments probed, is that each of these samples come with their own selection functions and biases. Despite all of these works applying careful corrections to the MDFs, they still do not agree fully with one another, particularly at the metal-poor end.

As previously discussed, the greatest difficulty in producing an MDF is to correct for metallicity dependant biases and selection effects incurred from the selection of a halo sample. Spectroscopic samples are typically more affected by completeness issues, since stars are specifically selected for follow-up observations, whereas photometric samples observe stars more ubiquitously. In addition, spectroscopic samples may be designed to target metal-poor stars, and thus have an implicit bias towards stars of low metallicity which must be corrected (such as is the case for HES for example). This correction process is highly non-trivial, and must be based on some models or assumed knowledge of the underlying stellar populations, and an over correction could lead to an underestimate of the number of metal-poor stars. In contrast, a photometric approach is likely much more complete (in so far as it is complete within the probed magnitude range), but is more susceptible to contamination due to the larger uncertainties associated with photometric metallicity determinations. These uncertainties may cause a broadening of the distribution, and an increase in the number of stars in the wings (i.e. at the metal-poor and metal-rich ends). We have made an effort to correct for this contamination by applying a statistical rescaling of the MDF with a GMM fit to a spectroscopic sample, but this is only effective in correcting for erroneous photometric metallicity determinations that are caught by the spectroscopy. Stars that have peculiar features in their spectra which affect their metallicities but that cannot be detected with

low/medium-resolution spectroscopy (e.g. certain binaries, blue stragglers) will go undetected and thus uncorrected by this approach, and result in an over estimation of the number of metal-poor stars. We have discussed the potential influence of these contaminants in Section 4.6.1, and although their contribution is likely to be small, this remains a caveat of the current work.

Nevertheless, a photometric sample has other advantages, such as larger sample sizes. Despite the many cuts imposed to make the *Pristine* halo sample, it still consists of ~ 80 000 stars. The next largest sample from the ones discussed above is from Allende Prieto et al. (2014), and consists of ~ 16 000 stars, although the $\log g$ cuts imposed for that sample are more comparable to the faintest MSTO sample presented in this chapter, which consists of ~ 230 000 stars. The narrower $\log g$ sample from Allende Prieto et al. (2014) that is comparable to our halo sample has < 1 000 stars. This naturally allows for an MDF that is statistically more robust than those made with the somewhat smaller spectroscopic samples.

Ultimately, none of the MDFs presented here are completely devoid of bias or selection effects. It is only through multiple, repeated efforts using a diversity of independent samples and methods, each with their own selection effects and biases, that we may hope to converge on a true determination of the halo MDF. This is particularly promising as we continue to collect large amounts of spectra with large scale surveys, and continue to improve the quality of photometric metallicity samples and distance determinations.

4.6.4 Comparison with Gaia-Enceladus

The MDF presented for the halo sample, with its peak [Fe/H] = −1.6 is qualitatively consistent with the MDF presented for the Gaia Enceladus structure in Helmi et al. (2018) (histogram in their Figure 2), although the MDF presented in this work has a much stronger tail of metal-poor stars then the one presented for Gaia Enceladus. It should be noted that the MDF presented in Helmi et al. (2018) is produced using a sample of APOGEE stars, and no attempt is made to correct for selection effects. Another study conducted by Gallart et al. (2019), select halo stars by taking all stars with large tangential velocities (> 200 km/s) from a local sample ($\lessapprox$ 2 kpc). They further separate these stars based on their metallicity distributions into a blue sequence that is consistent with the Gaia Enceladus population, and a red sequence that they suggest is representative of an "in situ" halo population. They determine the metallicity and age distributions for these population using a linear combination of synthetic isochrones, and find that they are both consistent with very old stellar populations. The photometric MDFs for these two sequences match reasonably well with spectroscopic MDFs produced by cross-matching with the LAMOST and GALAH spectroscopic surveys. The MDF of the blue sequence alone has a peak of [Fe/H] ~ −1.4 and loosely follows the distribution provided in Helmi et al. (2018), but the MDF of the red sequence is much more metal-rich with a peak at [Fe/H] ~ −0.7. If we consider these two populations together to make up the halo, then there is a much stronger metal-rich population than is shown in the MDF of the halo sample that we present in this work. However, this may simply indicate differences that result from selection of halo samples in both works as well as the limitations at high metallicity in this work. If the inner halo is indeed dominated by the stars from a single massive accretion event, then the MDF presented here with the distance range of our halo sample should consist mostly of those stars. However, despite qualitative similarities between the distributions, we cannot with the current data corroborate or reject this hypothesis.

4.6.5 The duality of the stellar halo

It remains an open question in the field of Galactic archaeology of whether the Galactic halo is composed of a single or multiple components. As discussed in Section 4.6.2, An et al. (2015)

derive a halo MDF that is consistent with having two distinct components, and suggest that this supports the dual halo hypothesis. The purple histogram in Figure 4.12 shows a fairly clear bimodality, to which An et al. (2015) fit two Gaussians with peaks at [Fe/H] = −1.4 and −1.9. They go on to estimate that the local population of halo stars is made up of 35 − 55% outer halo stars. The *Pristine* MDF does not show such a clear bimodality, but is also clearly not well fit by a single exponential in the metal-poor regime, which may indicate some more complex underlying populations. Similarly, neither of the spectroscopic MDFs from Schörck et al. (2009) or Allende Prieto et al. (2014) display an obvious bimodality, and neither of those authors discuss a potential bimodality in the MDFs as evidence for a dual halo. Interestingly, however, the halo MDF produced in Zuo et al. (2017) with SDSS photometry also does not show a clear bimodality, but they nevertheless decompose it into 3 underlying Gaussian components, one quite metal-rich and consistent with a disc population, and two that they identify as the inner and outer halo. They show this for two samples, one with $-8 < Z < -4$ and $10 < R < 14$ that shows peaks of [Fe/H] = −1.43 and −1.92, and a more distant sample with $-12 < Z < -8$ and $12 < R < 16$ for which they derive peaks of [Fe/H] = −1.51 and −2.25. With our current data, we do not have precise enough distances, or a clean enough selection of halo giant samples to go further out into the halo and see if a well defined break radius between these two halo populations can be identified, although this should be the topic of future work. Therefore, we do not draw a firm conclusion as to whether the MDF presented in this work is consistent with a multi or single component halo.

4.6.6 Implications for the formation of globular clusters

In the Milky Way there is an apparent paucity of globular clusters (GCs) at low metallicities. More specifically, of the ∼ 160 GCs observed to date, there are none with [Fe/H] < −2.5 (Harris 2010), which is seemingly at odds with the old ages derived for many GCs. In order to gain insight into this discrepancy, we can compare the metal-poor MDF of Milky Way GCs to the MDF of halo field stars to differentiate whether this lack of observed GCs could be the result of a real physical process or simply due to statistical undersampling. From the catalogue of Harris (2010), there are 56 GCs with [Fe/H] ≤ −1.5, 13 with [Fe/H] ≤ −2, and 0 with [Fe/H] ≤ −2.5. Using the relative number of stars at these metallicities from the scaled MDF of this work, we calculate an expected number of GCs to be 56, 26, 10, and 2.8 for [Fe/H] ≤ −1.5, −2, −2.5 and −3, respectively. Taking only the GCs in the Galactocentric distance range of the selected halo sample ($d_{GC} \sim 5 - 20$ kpc), the counts from the Harris (2010) catalogue are 30, 10, 0, and 0, and from *Pristine* 30, 14, 5.4, and 1.5 for [Fe/H] ≤ −1.5, −2, −2.5, −3, respectively. In both cases, we see that if the GC population were to follow the same MDF as the halo field stars, there is a significant discrepancy between the observed and expected number of GCs. This is suggestive that the perceived lack of metal-poor GCs is due to a real truncation of the GC MDF, or alternatively that there are underlying differences in formation mechanisms of GCs with respect to the field stars. A more robust statistical analysis to put these results on a firmer footing, as well as a more in-depth discussion will be presented in Gieles et al., in prep.

4.6.7 Future work

The halo MDF we present here relies on a spectroscopic sample to rescale the photometric metallicities. The *Pristine* spectroscopic follow-up sample, totalling ∼ 1 000 stars, was obtained over a three year period using a total of ∼ 165 nights of observing time on single-slit spectrographs mounted on 2−4m class telescopes. Although this approach was successful at obtaining a sample with which to explore the data set, it is unfeasable to follow-up the full number of *Pristine* candidates, particularly at fainter magnitudes. The William Herschel Telescope

Enhanced Area Velocity Explorer (Dalton et al. 2018, WEAVE) is a multi-object spectrograph (MOS), and starting this year, will begin collecting a large number of spectra to survey the northern sky. In a Memorandum of Understanding between the surveys, it is agreed that WEAVE will allocate ~ 6 fibres per deg^2 to follow-up *Pristine* metal-poor star candidates as part of its Galactic archaeology program. As the *Pristine* survey footprint currently spans over ~ 5 000 deg^2 and is still growing, we estimate that the WEAVE survey will observe a total of ~ 30 000 *Pristine* candidates during its operation. Apart from providing an excellent sample of metal-poor stars in the halo with which to conduct a multitude of Galactic archaeology studies, this will also provide a much larger spectroscopic sample with which to extend the current work. This will allow for an even more detailed investigation of the MDF in various galactic environments as we will have much more information on the dependence of contamination in the samples as a function of colour and magnitude.

Other upcoming MOS facilities and large spectroscopic surveys will continue to provide large numbers of spectra of halo stars that can be used to determine MDFs when their selection functions are not too complex, or with which to correct photometric surveys as in this work. In combination with the next data releases from the Gaia mission including more precise and accurate parallaxes, they will provide an excellent opportunity to study the metallicity distribution function extending further out into the halo and in narrower distance ranges.

4.7 Conclusions

In this chapter, we built an MDF of the Galactic halo using a carefully selected halo sample from the *Pristine* survey. The *Pristine* survey currently covers an area of ~ 2 500 deg^2 and provides photometric metallicities to ~ 4 million FGK type stars. From this sample, we applied colour and magnitude cuts to select MSTO samples spanning various heliocentric distance ranges between $0.6-25$ kpc. We found that there is a gradient showing decreasing metallicity with increasing distance, with a clear transition from a strong disc component at [Fe/H] ~ -0.7 in the brighter samples to a metal-poor peak at [Fe/H] ~ -1.6 where the halo dominates in the fainter samples. We then selected a halo sample consisting of ~ 80 000 stars, selected with $0.15 < (g_0 - i_0) < 0.4$, $19 < g_0 < 20$, $|b| > 30°$, and a cut in $(u_0 - g_0)$ vs $(g_0 - i_0)$ space to remove young foreground disc stars. This produced a sample covering a heliocentric distance range of $\sim 6-20$ kpc, which we confirmed to have a Galactocentric radius of $5 < R_{GC} < 20$ kpc and height above the disc $|Z| > 4$ kpc using a mock *Pristine* sample generated with the Galaxia code. We applied a statistical rescaling based on a sample of *Pristine* follow-up spectroscopy and SDSS/SEGUE stars, to account for contamination of metal-rich stars at the metal-poor end, and a correction to account for the preferential selection of metal-poor stars by the colour cut used to select the distance limited halo sample. At low galactic latitudes ($20° < |b| < 30°$) and in the direction of the Galactic anticenter, we identified a metal-rich population at the disc-halo interface, which we concluded to correspond to the Monoceros/TriAnd/ACS/EBS/A13 structure. We also selected a region containing the Virgo overdensity but found no difference in the MDF of this region when compared to other regions of equal Galactic latitude. Comparing the MDF from the current work with others derived in the literature, we find a larger fraction of stars at [Fe/H] < -2 than in previously published halo MDFs. In particular, we found a slope of $\Delta(\mathrm{LogN})/\Delta[\mathrm{Fe/H}] = 1.0 \pm 0.1$ in the metallicity range of $-3.4 <$ [Fe/H] < -2.5, which is substantially shallower than previous observational MDFs from spectroscopic samples, but agrees well with the simple closed-box chemical enrichment model of Hartwick (1976), although the model overestimates the number of stars at [Fe/H] < -3.5. We compare the MDF to that provided in Helmi et al. (2018) and Gallart et al. (2019) for the Gaia Enceladus structure, but despite some qualitative similarities we cannot conclude whether or not the halo MDF

we present is consistent with a single large merger event dominating the stellar content in the inner halo. We see no clear evidence of duality in the halo MDF, but conclude that we need better distances and cleaner samples of distant giants to probe the inner and outer halo in order to properly address this. Finally, we compare the MDF for Milky Way GCs and the MDF for the halo field stars presented in this chapter, and find that if these two populations follow the same metallicity distribution, the expected number of observed GCs is 10 at [Fe/H] < -2.5 and 2.8 at [Fe/H] < -3. We therefore suggest that the lack of GCs at [Fe/H] < -2.5 in the Milky Way is due to physical processes, rather than statistical undersampling.

Acknowledgements

We would like to thank Søren Larsen and Mark Gieles for useful discussions which motivated the development of this work. We gratefully thank the CFHT staff for performing the observations in queue mode, for their reactivity in adapting the schedule, and for answering our questions during the data-reduction process. We thank Nina Hernitschek for granting us access to the catalogue of Pan-STARRS variability catalogue.

KY, ES and AA gratefully acknowledge funding by the Emmy Noether program from the Deutsche Forschungsgemeinschaft (DFG). This work has been published under the framework of the IdEx Unistra and benefits from a funding from the state managed by the French National Research Agency as part of the investments for the future program. NFM, gratefully acknowledges support from the French National Research Agency (ANR) funded project "Pristine" (ANR-18-CE31-0017) along with funding from CNRS/INSU through the Programme National Galaxies et Cosmologie and through the CNRS grant PICS07708. KY, ES, AA, PJ, VH and NFM benefited from the International Space Science Institute (ISSI) in Bern, CH, thanks to the funding of the Teams "The Formation and Evolution of the Galactic Halo" and "Pristine". JIGH acknowledges financial support from the Spanish Ministry of Science, Innovation and Universities (MICIU) under the 2013 Ramón y Cajal program MICIU RYC-2013-14875, and also from the Spanish ministry project MICIU AYA2017-86389-P. CL thanks the Swiss National Science Foundation for supporting this research through the Ambizione grant number PZ00P2_168065.

Based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique of France, and the University of Hawaii. SDSS-IV is managed by the Astrophysical Research Consortium for the The Pan-STARRS1 Surveys (PS1) have been made possible through contributions of the Institute for Astronomy, the University of Hawaii, the Pan-STARRS Project Office, the Max-Planck Society and its participating institutes, the Max Planck Institute for Astronomy, Heidelberg and the Max Planck Institute for Extraterrestrial Physics, Garching, The Johns Hopkins University, Durham University, the University of Edinburgh, Queen's University Belfast, the Harvard-Smithsonian Center for Astrophysics, the Las Cumbres Observatory Global Telescope Network Incorporated, the National Central University of Taiwan, the Space Telescope Science Institute, the National Aeronautics and Space Administration under Grant No. NNX08AR22G issued through the Planetary Science Division of the NASA Science Mission Directorate, the National Science Foundation under Grant No. AST-1238877, the University of Maryland, and Eotvos Lorand University (ELTE).

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance

Computing at the University of Utah. The SDSS web site is www.sdss.org. SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatário Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group,Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

4.8 Appendix - Tables of the corrections and MDFs

Table 4.3 — Best fit parameters for the GMM shown in the right panel of Figure 4.6.

Component	Mean (x,y)	Weight	Covariance matrix $\begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_x^2 \end{bmatrix}$
1	(-1.320, -1.501)	0.28	$\begin{bmatrix} 0.183 & 0.113 \\ 0.113 & 0.188 \end{bmatrix}$
2	(-1.893, -2.121)	0.39	$\begin{bmatrix} 0.232 & 0.197 \\ 0.197 & 0.222 \end{bmatrix}$
3	(-2.448, -2.815)	0.33	$\begin{bmatrix} 0.251 & 0.033 \\ 0.033 & 0.127 \end{bmatrix}$

Table 4.4 — MDFs for the raw *Pristine* halo sample, the sample corrected with the GMM, and the sample corrected with the GMM and colour cuts, plotted in Figure 4.12 as the grey, cyan, and black histograms, respectively.

[Fe/H]	raw *Pristine*	GMM	corrected
-3.95	395	10	10
-3.85	143	34	34
-3.75	170	59	59
-3.65	294	104	104
-3.55	500	149	149
-3.45	657	192	192
-3.35	787	353	353
-3.25	1037	466	466
-3.15	1189	625	628
-3.05	1458	765	767
-2.95	1650	984	989
-2.85	1999	1241	1255
-2.75	2235	1491	1520
-2.65	2516	1808	1847
-2.55	2845	2098	2154
-2.45	3186	2531	2634
-2.35	3521	2816	2996
-2.25	3770	3142	3366
-2.15	4116	3596	3892
-2.05	4503	3910	4342
-1.95	4635	4486	5174
-1.85	4778	4730	5680
-1.75	4798	4951	6178
-1.65	4650	5054	6557
-1.55	4409	5086	7092
-1.45	4011	4925	7589
-1.35	3415	4529	7841
-1.25	2865	4124	7889
-1.15	2259	3588	8787
-1.05	1816	3101	12421
-0.95	1326	2477	-
-0.85	1051	1875	-
-0.75	763	1426	-
-0.65	538	997	-
-0.55	415	779	-
-0.45	310	483	-
-0.35	264	331	-
-0.25	217	200	-
-0.15	218	130	-
-0.05	-	63	-

Table 4.5 — Correction factor applied for the colour cut to select the halo sample. The rightmost columns is the correction factor applied to produce the black histogram in Figure 4.12 and is the mean of the 10 and 12 Gyr columns.

[Fe/H]	8 Gyr	10 Gyr	12 Gyr	14 Gyr	Correction
-3.95	1.000	1.000	1.000	1.000	1.000
-3.85	1.000	1.000	1.000	1.000	1.000
-3.75	1.000	1.000	1.000	1.000	1.000
-3.65	1.000	1.000	1.000	1.000	1.000
-3.55	1.000	1.000	1.000	1.000	1.000
-3.45	1.000	1.000	1.000	1.000	1.000
-3.35	1.000	1.000	1.000	1.000	1.000
-3.25	1.000	1.000	1.000	1.000	1.000
-3.15	1.009	1.000	1.011	1.031	1.006
-3.05	1.007	1.003	1.003	1.027	1.003
-2.95	1.006	1.012	1.000	1.011	1.006
-2.85	1.009	1.009	1.015	1.018	1.012
-2.75	1.000	1.016	1.024	1.006	1.020
-2.65	1.005	1.007	1.037	1.000	1.022
-2.55	1.007	1.017	1.037	1.014	1.027
-2.45	1.010	1.034	1.048	1.041	1.041
-2.35	1.012	1.049	1.079	1.059	1.064
-2.25	1.024	1.058	1.085	1.112	1.072
-2.15	1.050	1.069	1.096	1.141	1.083
-2.05	1.083	1.090	1.131	1.191	1.111
-1.95	1.111	1.127	1.180	1.258	1.154
-1.85	1.129	1.160	1.242	1.330	1.201
-1.75	1.156	1.193	1.303	1.429	1.248
-1.65	1.212	1.239	1.356	1.594	1.298
-1.55	1.267	1.297	1.492	1.835	1.395
-1.45	1.315	1.403	1.679	2.054	1.541
-1.35	1.391	1.531	1.932	2.338	1.732
-1.25	1.479	1.689	2.137	5.174	1.913
-1.15	1.635	2.029	2.869	-	2.449
-1.05	1.841	2.411	5.600	-	4.006
-0.95	2.209	2.658	-	-	-
-0.85	2.471	10.235	-	-	-
-0.75	3.101	-	-	-	-
-0.65	20.823	-	-	-	-
-0.55	-	-	-	-	-
-0.45	-	-	-	-	-
-0.35	-	-	-	-	-
-0.25	-	-	-	-	-
-0.15	-	-	-	-	-
-0.05	-	-	-	-	-

5 A uniquely clean view of the Galactic outer halo using blue horizontal branch stars

E. Starkenburg, K. Youakim, N. F. Martin, G. Thomas, D. S. Aguado, A. Arentsen, R. G. Carlberg, J. I . González Hernández, R. Ibata, N. Longeard, A. McConnachie, J. Navaro, R. Sánchez Janssen, and K. Venn

— Published as E. Starkenburg, K. Youakim et al., 2019, MNRAS 490, 5757 —

ABSTRACT

We use the *Pristine* survey $CaHK$ narrow-band photometry, combined with the SDSS ugr photometry, to provide a cleaner sample of blue horizontal branch stars in the Galactic halo out to large distances. We demonstrate a completeness of 91% and a purity of 93% with respect to available spectroscopic classifications. We subsequently use our new clean sample of these standard candles to investigate the substructure in the Galactic halo over the *Pristine* footprint. Among other features, this allows for a careful tracing of multiple parts of the Sagittarius stream, providing a measurement independent from other tracers used and reaching larger distances. Moreover, we demonstrate with this clean and complete sample that the halo follows a density profile with a negative power-law slope of $3.5-4.0$. As the relatively shallow SDSS u-band is the limiting factor in this technique, we foresee large potential for combining Pristine survey photometry with the much deeper u-band photometry from the Canada-France-Imaging Survey.

5.1 Introduction

In order to measure the 3D structure of our local Universe, we have always been very much dependent on the existence of the various standard candles that, together, form the components of the distance ladder. Standard candles that are suited for the study of the outer regions of our Galaxy include several types of intrinsically variable stars such as RR Lyrae, but also (with lesser precision) M-giants, carbon stars, and blue horizontal branch (BHB) stars. For several decades now, they have been used as important tracers to study the total enclosed mass of the Milky Way galaxy at different distances (e.g., Sommer-Larsen et al. 1989; Norris & Hawkins 1991) and to study various (kinematical) substructures in the outer Galactic halo (e.g., Arnold & Gilmore 1992; Kinman et al. 1994).

As the most striking example of substructure in our Milky Way, the long tidal features originating from the Sagittarius dwarf galaxy (first discovered by Ibata et al. 1994) have received much interest (see Law & Majewski 2016 for a review). Because the stream wraps around the Galaxy and traces a large range of distances in the Galactic halo, it provides an excellent opportunity to constrain the dark matter potential of our Galaxy and even its shape (e.g., Ibata et al. 2001; Helmi 2004; Johnston et al. 2005; Law et al. 2005). Over the years it has been mapped using a variety of tracers, from M-giants (e.g., Majewski et al. 2003; Koposov et al. 2015), main-sequence turn-off stars (e.g., Koposov et al. 2013; Pila-Díez et al. 2014; Lokhorst et al. 2016), red horizontal branch stars (Shi et al. 2012), carbon-rich long-period variables (Huxor & Grebel 2015), RR-Lyrae (e.g., Hernitschek et al. 2017; Cohen et al. 2017; Sesar et al. 2017a), K-giants (Liu et al. 2014; Janesh et al. 2016), to BHB stars (e.g., Fukushima et al. 2018). Belokurov et al. (2014) use a mixture of different tracers to follow both the leading and trailing tails of the stream and find that their apocenters lie at $R^L = 47.8 \pm 0.5$ kpc and $R^T = 102.5 \pm 2.5$ kpc respectively. Using RR Lyrae, Sesar et al. (2017a) follow the stream beyond this distance and discover a plume of stars 10 kpc beyond the apocenter of the leading arm, and, even more spectacularly, a "spur" extending to 130 kpc at the trailing arm apocenter, thus reaching almost 30 kpc beyond the furthest distance at which it had previously been mapped. Whereas the first feature is confirmed using very deep broad-band data selecting BHB stars (Fukushima et al. 2018), the second feature has up until now only been detected using RR Lyrae stars and possibly one M-giant member (Li et al. 2019). These measurements at large radii are particularly constraining for the modelling of streams in the Galactic halo throughout its kinematical history (Dierickx & Loeb 2017; Fardal et al. 2019). Better constraints on the different features of the Sagittarius stream clearly enhance our understanding of the formation and evolution of the Milky Way halo.

In this work, we use the *Pristine* survey (Starkenburg et al. 2017), which employs a narrow-band filter around the Ca II H & K absorption features to select BHB stars and, in particular, to distinguish them from the intrinsically fainter population of blue straggler (BS) stars at similar temperatures. The narrow-band filter used by the *Pristine* survey was initially designed for excellent metallicity sensitivity in FGK-type stars and major science cases of the survey include the search for rare extemely metal-poor stars, the study of very metal-poor satellite systems, and the characterisation of metallicity structures in the Milky Way halo (Starkenburg et al. 2017; Youakim et al. 2017; Caffau et al. 2017; Starkenburg et al. 2018; Longeard et al. 2018, 2019; Bonifacio et al. 2019; Aguado et al. 2019). However, in this work, we instead make use of the excellent gravity sensitivity of the narrow-band filter for A stars, in particular when it is combined with u-band photometry from, for instance, the Sloan Digital Sky Survey (hereafter SDSS; York et al. 2000). Using broad-band information alone, success rates for separating BHB and BS stars have reported completenesses of 50 to 70% and contamination rates up to 30% (e.g., Bell et al. 2010; Vickers et al. 2012; Fukushima et al. 2018; Thomas et al. 2018b). Naturally, these success rates drop as fainter stars are investigated and the uncertainties on the

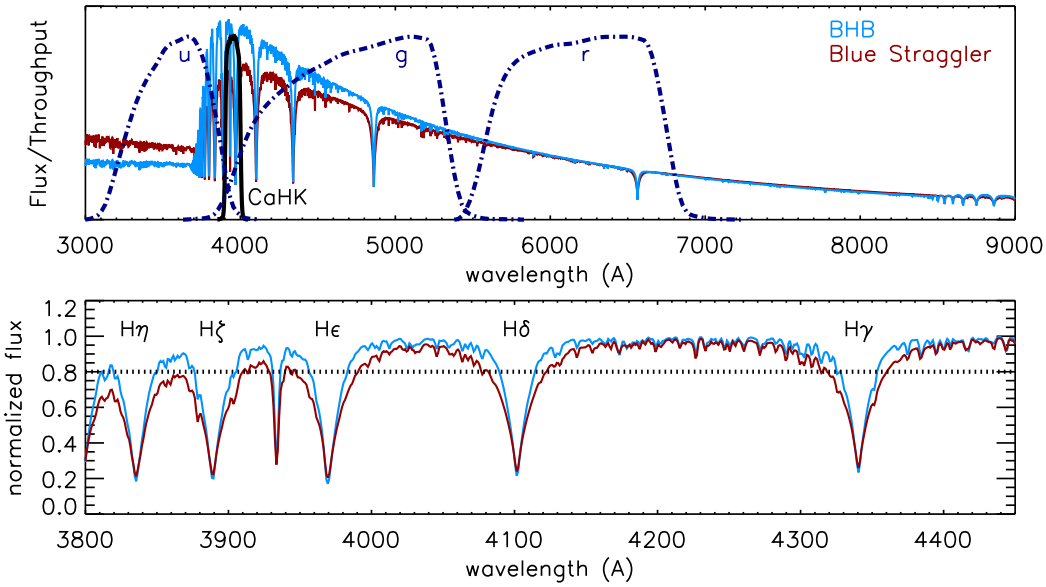

Figure 5.1 — Top panel: Synthetic spectra from a library of synthetic stellar spectra by Munari et al. (2005) for stars with stellar parameters T_{eff}= 8 500 K and [Fe/H] = −0.5. One of them has $\log g = 3.0$ (light blue, typical parameters for a BHB star) and the other one $\log g = 5.0$ (dark red, typical for a BS star) overplotted with the throughput curves for the SDSS ugr filters (dash-dotted dark blue lines) and the filter curve for the *Pristine* survey CaHK filter (thicker black line). Bottom panel: The same spectra, but now for a narrower wavelength range such that the Balmer series of hydrogen lines is plainly visible. The horizontal dotted line indicates a level of 0.8 in the normalized spectra where the widths of the hydrogen lines are traditionally measured to determine if the star is a BHB or BS star (see text for details).

photometry increase. Deason et al. (2012) report that out of 38 SDSS photometrically selected stars with spectroscopic follow up in the $20 < g < 22$ magnitude regime, only 7 bona fide BHB stars were found.

We demonstrate in this work how the combination of *Pristine* survey and SDSS photometry leads to an unprecedentedly clean and complete sample of BHB stars in Section 5.2. In Section 5.3 we subsequently focus on characterizing the outer halo using this new BHB sample in the *Pristine* survey footprint. We investigate its spatial coverage in several ways, focusing on the halo profile and quantifying its clumpiness. In addition, we use this new sample to map the Sagittarius stream in our footprint, and tentatively trace the "spur" feature further out than previously possible. Finally, we conclude our findings and present a future outlook in Section 5.4.

5.2 Selection of BHB stars

The first selection criterion for BHB stars is a measure of their overall temperature to select A-type stars, as given, for instance, by broad-band SDSS colours. In this region of colour space, there are not many contaminants except for very blue (hot and young) main-sequence stars, which are rarely found in old populations, and blue straggler stars. Careful colour cuts can avoid most contamination from other hot objects such as QSOs (Navarrete et al. 2018), white dwarfs (Gentile Fusillo et al. 2015), and hot subdwarfs (Geier et al. 2019). In contrast to genuine

hot main-sequence stars, blue straggler stars are found in almost all types of populations. In the colour-magnitude diagram, they reside above the main-sequence. The physical difference between them and BHB stars is that they have higher gravities, originating from hydrogen burning stars on the main sequence whereas the BHB stars are giant helium-burning stars. While BHB stars typically have gravities $\log g = 2.8-3.75$, BS stars show gravities in the range of $\log g = 3.75-5.0$ (Vickers et al. 2012). Because the latter are intrinsically fainter than BHB stars, it is important to clean any BHB samples of these contaminants before using their standard candle distances to investigate the 3D structure of the Milky Way halo.

This difference in gravity between the two populations in turn leads to differences in their spectra and their broad-band colours. In Figure 5.1, two synthetic spectra with typical parameters for a BHB star and a BS star of the same temperature are shown, both spectra are taken from a library of synthetic stellar spectra by Munari et al. (2005). From this library, we use the models calculated with the new ODF atmospheres (Castelli & Kurucz 2003), no rotation, alpha-enhancement of +0.4, a micro-turbulent velocity of 2 km s^{-1}, and with 1Å/pix dispersion. The importance of the u-band for distinguishing these two stars stands out clearly, mainly due to the very gravity-sensitive Balmer jump feature which is much steeper for stars of lower gravity (see also Lenz et al. 1998). The g-band has some gravity sensitivity as well. But at redder wavelengths, where the SDSS g-band filter is most efficient, the total flux is dominated by the temperature-sensitive (rather than gravity-sensitive) blackbody curve. Further gravity sensitivity is found to a lesser extent in the Paschen lines in the z-band around 8700 Å. The *Pristine* survey CaHK filter (black solid filter curve) covers a very interesting wavelength range where the difference between the two spectra is at its maximum, but in reversed order to the difference observed in the u-band. There is thus reason to expect that a combination of these two filters will provide an excellent diagnostic to discriminate between these two types of stars, even though the *Pristine* survey filter was not specifically designed for this science case.

In addition to photometric measurements at the blue end of the spectrum, BHB and BS stars can be separated by studying their Balmer lines as is illustrated for the same synthetic spectra in the bottom panel of Figure 5.1. The wings of the Balmer lines are gravity sensitive and the width of these lines (often measured at 80% – or alternatively 85% – of the total flux level as illustrated by the horizontal dotted line) indicates the gravity of the star (e.g., Rodgers et al. 1981; Clewley et al. 2002; Sirko et al. 2004; Xue et al. 2008). Alternatively, the Balmer lines are fitted by Sérsic profiles (Sérsic 1963) and the line shape parameters are used to distinguish the different types of stars. Such methods to separate BHB, BS, and main-sequence (MS) stars using the spectra for A-type stars (as pioneered by Pier 1983) are for instance applied to large spectral libraries of SDSS/SEGUE A-type stars (Yanny et al. 2000; Xue et al. 2008; Deason et al. 2011; Xue et al. 2011), or more targeted efforts studying the substructure in the Milky Way outer halo (e.g., Navarrete et al. 2018). Combinations of spectroscopic and photometric measurements are also reported in the literature, such as for instance in Kinman et al. (1994); Wilhelm et al. (1999); Clewley et al. (2002).

In this section we attempt to select a sample that is as clean and pure as possible using either SDSS ugr colours alone (Section 5.2.1), or SDSS ugr colours supplemented with the $CaHK$ magnitudes from the *Pristine* survey (Section 5.2.2). We thus focus on characterizing this large and more complete population of BHB stars, for which there is not yet spectroscopic follow-up. Due to their use as standard candles, a clean selection is very important for an accurate characterization of 3D structures in the outer Milky Way halo.

5.2.1 Towards a clean and pure BHB sample using SDSS

In order to select BHB stars from SDSS colours we closely follow the method outlined by Deason et al. (2011). To start, all objects from the star catalogue of SDSS DR14 (Abolfathi

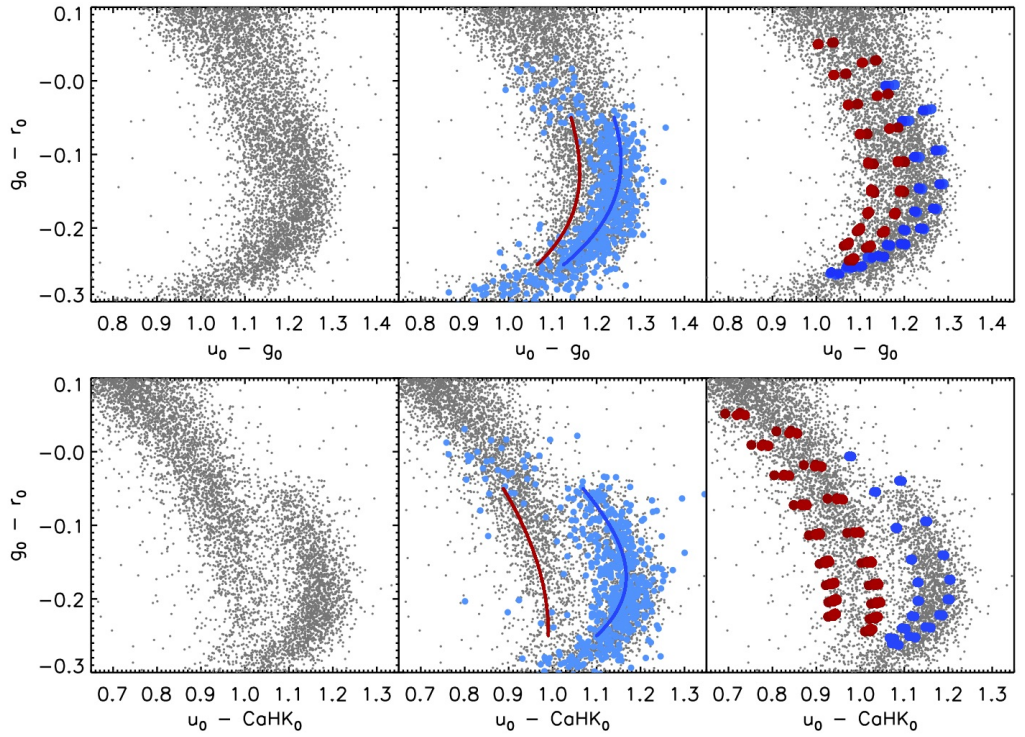

Figure 5.2 — Top panels: colour-colour space using SDSS broad-band data for a selected sample of clean photometry for hot stars (see text for details). In the middle, the ridgelines defined in Deason et al. (2011) for BHB and BS stars (blue and red lines respectively) are plotted on top of the BHB sample of Xue et al. (2011) (light blue filled circles). In the right panel, the results from synthetic colours for spectra from the Munari et al. (2005) library with typical BHB (blue points; $-2.5 < [\mathrm{Fe/H}] < -1.5$, $7\,500 < T_{\mathrm{eff}} < 9\,500$ K, $3.0 < \log g < 3.5$) and BS (red points; $-1.0 < [\mathrm{Fe/H}] < -0.5$, $7\,500 < T_{\mathrm{eff}} < 9\,500$ K, $4.0 < \log g < 4.5$) stellar parameters are overplotted. Bottom panels: The same stellar sample as in the top panels, but now for a colour-colour space combining the SDSS and the *Pristine* survey photometry. The ridgelines are defined in this work and described in the text.

et al. 2018) are selected to fulfill the colour cuts: $0.9 < u_0 - g_0 < 1.4$ and $-0.25 < g_0 - r_0 < 0.0$, corresponding to the colour-colour space for A-stars. In both these colour cuts the DR14 dereddened magnitudes are used. As can been seen in Figure 4 of Deason et al. (2014), these broad band cuts ensure that the white dwarf contamination becomes negligible and cut out most QSO contaminants. We subsequently check that the stars are not flagged as variable in Pan-STARRS1 (hereafter PS1, Chambers et al. 2016). The variability criteria by Hernitschek et al. (2016) provide an independent cleaning by which most QSO sources can be identified (see their Figure 3.). The sample is further cleaned of any star that has not passed the SDSS selection flag for clean photometry, or are flagged in SDSS u, g, or r photometry separately to be saturated, have de-blending or interpolation problems, are suspicious detections, or are close to the edge of a frame.

The distribution of these A-type stars in $u_0 - g_0$ and $g_0 - r_0$ colour-colour space is shown in the top-left panel of Figure 5.2, where only stars with u-band uncertainties less than 0.02 are selected to avoid clutter from low S/N observations. We furthermore restrict ourselves to stars

that are also in the *Pristine* survey to facilitate a fairer comparison with the results of the next section. With some effort, two sequences of stars can be distinguished in this panel, although they seem to touch and partly overlap. If the final overlapping sample of spectroscopically confirmed BHB stars from Xue et al. (2011) is overplotted (see middle panel), it becomes clear that the rightmost of these sequences corresponds to BHB stars. Any remaining QSOs in the sample are much more likely to contaminate the BS sample rather than the BHB sample, as they predominantly occupy the regions of lower $u_0 - g_0$ (Deason et al. 2014).

Deason et al. (2011) have defined two ridgelines to separate the two populations using their classifications from stellar spectra. We here adopt the same ridgelines, but we shift them by 0.04 magnitudes in the u-band, since there have been some shifts in the calibration of the dereddened magnitudes in the various SDSS data releases (Finkbeiner et al. 2016) and this is the mean shift for these stars between DR8 and DR14. The lines are overplotted in blue and red, respectively, in the middle top panel of Figure 5.2, and described by the following equations:

$$\begin{aligned}(u_0 - g_0)^0_{\rm BHB} =& 1.167 - 0.775(g_0 - r_0) - 1.934(g_0 - r_0)^2 \\ &+ 9.936(g_0 - r_0)^3 + 0.04,\end{aligned} \tag{5.1}$$

$$\begin{aligned}(u_0 - g_0)^0_{\rm BS} =& 1.078 - 0.489(g_0 - r_0) + 0.556(g_0 - r_0)^2 \\ &+ 13.444(g_0 - r_0)^3 + 0.04\end{aligned} \tag{5.2}$$

Finally, in the top-right panel of Figure 5.2, the values for synthetic models from the Munari stellar library (Munari et al. 2005, using the same sets of synthetic spectra as illustrated in Figure 5.1) are overplotted, as integrated under the SDSS filter curves. For the synthetic spectra, a shift of 0.04 mag was needed in the u-band to correct for the offset between the SDSS u-band and its AB magnitude as integrated under the filter curve*. As in the previous panel, we have additionally added another 0.04 shift since this is the mean shift between DR8 and DR14 photometry in the u-band for these stars. Both sets of models are run with effective temperatures ranging from 7 500 to 9 500 K. Furthermore, the synthetic BHB stars (shown in blue) are selected to have $3.0 \leq \log g \leq 3.5$ and $-2.5 \leq$[Fe/H]≤ -1.5, in accord with what is typical for this population (Xue et al. 2008). The blue straggler synthetic spectra are selected to have $4.0 \leq \log g \leq 4.5$ and $-1.0 \leq$[Fe/H]≤ -0.5. Indeed, the synthetic models with these parameters seem to fit the observed locii of the sample well.

To calculate the probability that a star is either a BHB or a BS star, we again follow the method presented by Deason et al. (2011) and, given their measured standard deviations for the BHB and BS populations ($\sigma_{\rm BHB,0}(u_0 - g_0) = 0.04$ and $\sigma_{\rm BS,0}(u_0 - g_0) = 0.045$), we use

$$\begin{aligned}p(ugr \mid {\rm BHB}) &\propto \exp\left(-\frac{\left[(u_0-g_0) - (u_0-g_0)^0_{\rm BHB}\right]^2}{2\sigma^2_{\rm BHB}}\right), \\ p(ugr \mid {\rm BS}) &\propto \exp\left(-\frac{\left[(u_0-g_0) - (u_0-g_0)^0_{\rm BS}\right]^2}{2\sigma^2_{\rm BS}}\right).\end{aligned} \tag{5.3}$$

Here,

$$\begin{aligned}\sigma_{\rm BHB} &= \sqrt{\sigma^2_{\rm BHB,0} + \sigma^2_{\rm BHB,(u_0-g_0)}} \\ \text{and } \sigma_{\rm BS} &= \sqrt{\sigma^2_{\rm BS,0} + \sigma^2_{\rm BS,(u_0-g_0)}},\end{aligned} \tag{5.4}$$

*see https://www.sdss.org/dr14/algorithms/fluxcal/

thus folding in both the intrinsic width of the populations and the uncertainty on the colour measurement. The colour-based posterior probabilities of class membership are then described as:

$$P(\mathrm{BHB} \mid ugr) = \frac{p(ugr \mid \mathrm{BHB})\ N_{\mathrm{BHB}}}{p(ugr \mid \mathrm{BHB})\ N_{\mathrm{BHB}} + p(ugr \mid \mathrm{BS})\ N_{\mathrm{BS}}},$$
$$P(\mathrm{BS} \mid ugr) = \frac{p(ugr \mid \mathrm{BS})\ N_{\mathrm{BS}}}{p(ugr \mid \mathrm{BHB})\ N_{\mathrm{BHB}} + p(ugr \mid \mathrm{BS})\ N_{\mathrm{BS}}}. \tag{5.5}$$

Where the total numbers of stars N_{BHB} and N_{BS} in a given colour range are found iteratively and are described in Table 1 of Deason et al. (2011).

5.2.2 Towards a clean and pure BHB sample using SDSS and *Pristine*

Here we repeat the selection steps as outlined in the last section. However, instead of using $u_0 - g_0$ and $g_0 - r_0$ as our colour-colour selection space, we replace the former with $u_0 - CaHK_0$; thus taking into account the *Pristine* narrow-band filter. This is illustrated in the bottom panels of Figure 5.2. The selected stars, shown here as grey points, are identical to the stars selected for the upper panels. It is clear that the combination of u-band and $CaHK$-band photometry allows for a much cleaner selection of BS and BHB stars, as expected from the synthetic spectra in Figure 5.1. Now, clearly, two different sequences can be distinguished. When the spectroscopically confirmed BHB stars from Xue et al. (2011) are overplotted (in the bottom-middle panel), it is clear that the BHB stars are on the sequence on the right. For the coolest stars, with the highest $g_0 - r_0$ values, we do see that some spectroscopically classified BHB stars fall onto the bluer sequence. However, taking into account that the spectroscopic classification in this stellar parameter regime becomes more difficult (see Figure 5 of Xue et al. 2008), we actually take this as likely evidence that these stars are spectroscopically misclassified. We note that, independently, a similar conclusion is reached by Lancaster et al. (2019b), who remove stars with $u_0 - g_0 < 1.15$ and $g_0 - r_0 > -0.07$ from the sample based on their high tangential velocities if a BHB star distance is adopted (indicating they are probably misclassified BS stars). For the remainder of this work, we restrict ourselves to $g_0 - r_0 < -0.05$.

As in the previous section, we find that the synthetic colours derived from synthetic spectra integrated under the filter curves do a good job in following the two sequences. We therefore let the synthetic predictions guide the definition of our ridgelines in this new colour-colour space, by fitting the synthetically predicted points in the right bottom panel with polyomials. The ridgelines we obtain are shown in the bottom middle panel and are described by

$$(u_0 - CaHK_0)_{\mathrm{BHB}} = 0.997 - 1.465(g_0 - r_0) + 0.411(g_0 - r_0)^2 + 18.531(g_0 - r_0)^3, \tag{5.6}$$

$$(u_0 - CaHK_0)_{\mathrm{BS}} = 0.832 - 1.222(g_0 - r_0) - 2.094(g_0 - r_0)^2 + 1.046(g_0 - r_0)^3. \tag{5.7}$$

We remove from the sample any objects that are beyond 3σ from any of these two ridgelines (taking into account the dispersion and photometric uncertainties as outlined in Equation 5.4) to remove spurious objects and contamination. This only removes $\sim 1\%$ of all objects. Subsequently, we again follow the procedure as described in Equations 5.3 until 5.5, with the difference that we replace $u_0 - g_0$ with $u_0 - CaHK_0$ and that we set both $\mathrm{N_{BHB}}$ and $\mathrm{N_{BS}}$ to 1, since

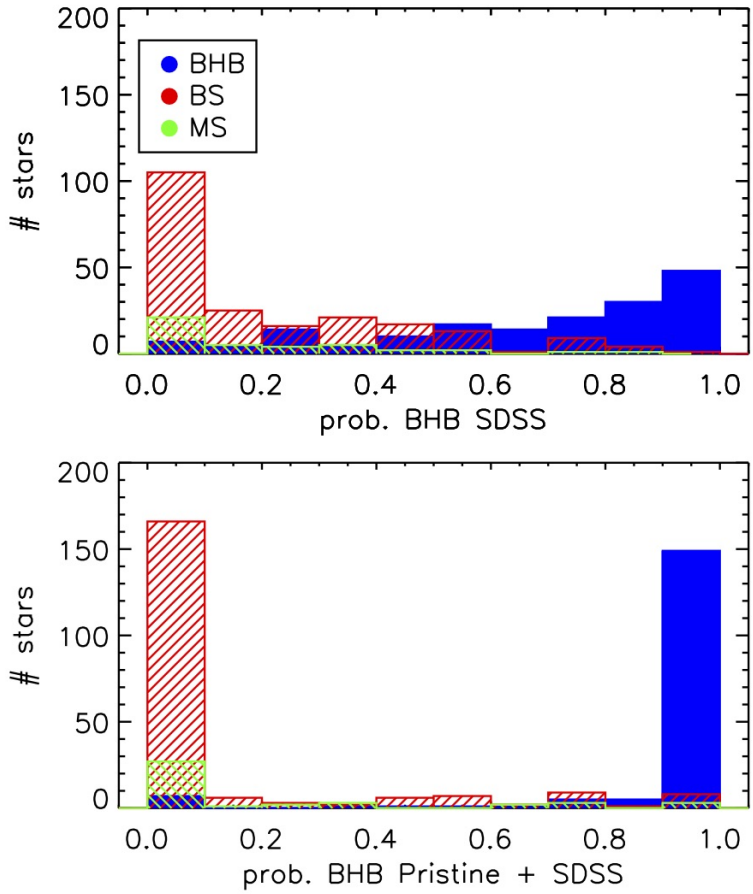

Figure 5.3 — The probability that stars from the Xue et al. (2008) spectroscopically followed up sample are BHB stars according to the methods presented in Sections 5.2.1 and 5.2.2. The stars from Xue et al. (2008) are held against the same quality criteria (see text for details) and are separated in BHB (blue filled histogram), BS (red histogram), and MS (green histogram) stars according to the Xue et al. (2008) classification. It is clear that the addition of $CaHK$ photometry adds to both the purity and completeness of a photometrically selected BHB sample.

the sequences are so well separated that no strong prior on the ratios of BS and BHB stars as a function of colour is needed. This allows us to be more agnostic about the fraction of BS and BHB stars as a function of colour and distance.

5.2.3 Comparison between the two samples

In this section we compare the performance of both the samples defined above in both purity and completeness of their BHB selection. For this purpose, we use the spectroscopically defined samples of Xue et al. (2008), which includes both BS and BHB stars, as well as genuine MS stars, that are spectroscopically classified. This sample covers the magnitude range $14.0 < g < 19.2$. As mentioned above, a cut is imposed on $g_0 - r_0$ where we trust the spectroscopic analysis at < -0.05. Furthermore, the following quality cuts are applied to both samples to remove sources with bad photometry in either SDSS or the *Pristine* survey: $0.7 <$

$u_0 - CaHK_0 < 1.3$, CASU photometry flag for *Pristine* photometry equals -1 (object has a stellar point-like point spread function), the star shows no sign of variability in PS1 photometry, its SDSS photometry is clean, and the uncertainty on $CaHK < 0.1$. The last criterion only removes very few stars, as the overlap with the spectroscopic sample of Xue et al. (2008) contains mostly relatively bright stars and the mean uncertainty on $CaHK$ within this sample is 0.013. We also note that in this sample overlapping with spectroscopic observations the number of BS/MS stars is only slightly larger than the number of BHB stars. The natural expectation is that this ratio will vary with height above the Galactic plane (in this sample, the Galactic latitude is always greater than 20 degrees).

Figure 5.3 shows for both samples the resulting distribution of the Xue et al. (2008) sample as a function of the derived $P(\text{BHB} \mid ugr)$ or $P(\text{BHB} \mid ugrCaHK)$. This figure confirms and quantifies the qualitative result from Figure 5.2. With the addition of the *Pristine* $CaHK$ photometry, the selection of BHB stars becomes both purer and more complete. Were one to select stars with $P(\text{BHB}) > 0.8$ in both samples, the purity of BHB stars in the SDSS sample is 93% (7% is still contamination), while the completeness is 46% (54% of the spectroscopically confirmed BHB stars are not selected). These numbers are very similar, or even a bit better, to those quoted in the literature (Sirko et al. 2004; Bell et al. 2010; Vickers et al. 2012), also when the z-band is used, instead or in addition to the use of the u-band for classification (see respectively Fukushima et al. 2018; Thomas et al. 2018b).

When adding the *Pristine* $CaHK$ magnitudes however, as shown in the bottom panel of Figure 5.3, the purity stays at 93% and the completeness rises spectacularly to 91%.

5.2.4 Completeness and purity as a function of magnitude

Hampering a better understanding of the outermost parts of the Galactic halo are the larger photometric uncertainties at fainter magnitudes. BHB stars of 21.5 in g-band magnitude are required to enable a mapping of the halo to distances of $130 - 170$ kpc. A clean sample of such stars provides not only a detailed 3D map of the outer halo, but also serves as an excellent sample for spectroscopic follow-up, as such enabling careful kinematic studies of the outer (dark matter) halo. The distance in $u_0 - g_0$ between the BHB and BS star ridgelines is only 0.12 mag. This difference increases to ~ 0.2 mag in the $u_0 - CaHK_0$ space we use in this work. For stars with $21 < g_0 < 21.5$, the mean $CaHK$ uncertainty is 0.11 mag and the mean uncertainty on SDSS u-band photometry is 0.21 mag (as illustrated in Figure 5.4), meaning that photometric uncertainties will play an important role when attempting to separate both populations.

In Figure 5.5, we investigate the results of purity and completeness as a function of magnitude in two ways. Firstly, we investigate the magnitude dependence in the overlapping spectroscopic sample of Xue et al. (2008). Secondly, we Monte Carlo simulate mock datasets with uncertainties typical for that magnitude and subsequently apply our techniques.

For the spectroscopic overlap sample, the numbers of bona fide BHB stars drop significantly at $g_0 = 18$. Only 31 BHB stars are available between $18 < g_0 < 20$, making a critical assessment of the performance as function of magnitude difficult. Nevertheless, using this approach, this faintest magnitude bin gets 70% (61%) purity, and only 23% (61%) completeness by using SDSS only and probability cut-off values > 0.8 (> 0.5). For *Pristine* + SDSS, the purity remains similar, but the completeness is much better at 71% (84%). These values are shown as triangle symbols in Figure 5.5. It can also be seen in this figure that for stars brighter than $g_0 = 18$ the *Pristine* + SDSS method shows hardly any decline of success with magnitude.

Due to the small number of stars in the overlapping spectroscopic sample at faint magnitudes, we also Monte Carlo simulate mock datasets to estimate completeness, which produces a more robust representation of the effect of the photometric uncertainties on the results. Although this approach does of course not include systematic effects, it can be seen from

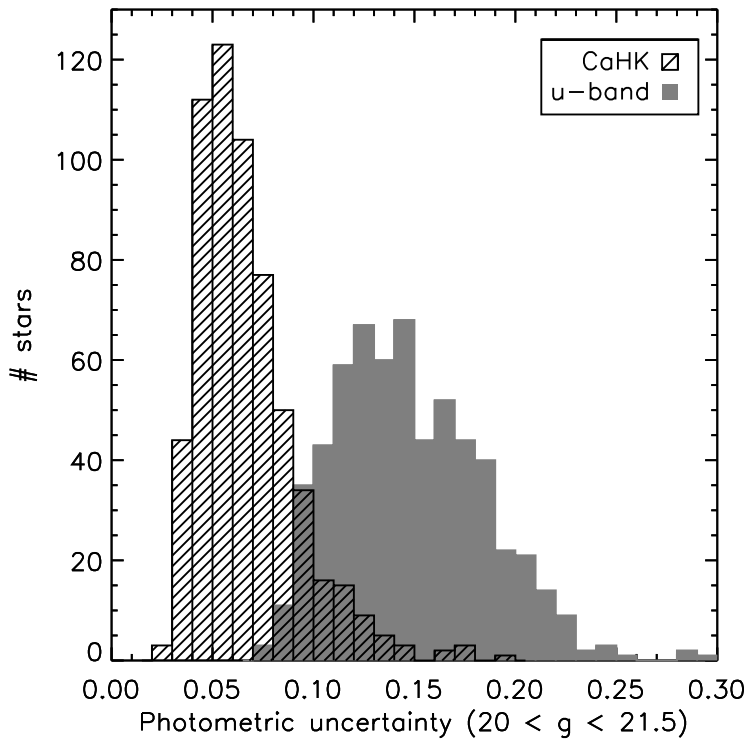

Figure 5.4 — Photometric uncertainties in u and $CaHK$ measurements respectively, for A stars with g-magnitudes between 20 and 21.5. These would probe distances of 66 – 170 kpc when classified as BHB stars (depending also on their colour).

the comparison of the modelling (lines) and data (triangles) in Figure 5.5 in the magnitude ranges where the two overlap that the model provides a reasonable description of the data. Predictions for performance of completeness are more straightforward to make than for purity, as measurements of purity also depend on the true number ratio between BHB and BS stars (it includes a measure of mis-classified BS stars). Therefore, purity predictions are shown in two panels in Figure 5.5. In the middle panel, the ratio between BHB and BS stars is modeled as 1:1, in the right panel as 1:5. The latter is more in concordance with the ratio at the faintest magnitude bin in the spectroscopic data set, but this is not necessarily a representation of the true value due to the non-trivial way in which BHB star candidates enter the SDSS selection function (see e.g., Sirko et al. 2004).

On the top x-axes of Figure 5.5 we show the distance corresponding to the magnitude for a BHB star with $(g_0 - r_0) = -0.15$. The completeness of the P > 0.8 sample very steeply declines at g_0-magnitudes beyond 19 and approaches zero at $g_0 = 20.5$ (corresponding to 100 kpc, indicated by the dotted grey line in Figure 5.5). Basically, the uncertainties become so large that the method can not distinguish BHB stars with such high probability anymore. The P > 0.5 sample, on the other hand, shows a completeness of 50% still at $g_0 = 21$, but at the cost of a lower purity, the precise value of which will depend strongly on the true number ratio of BHB and BS stars at those magnitudes.

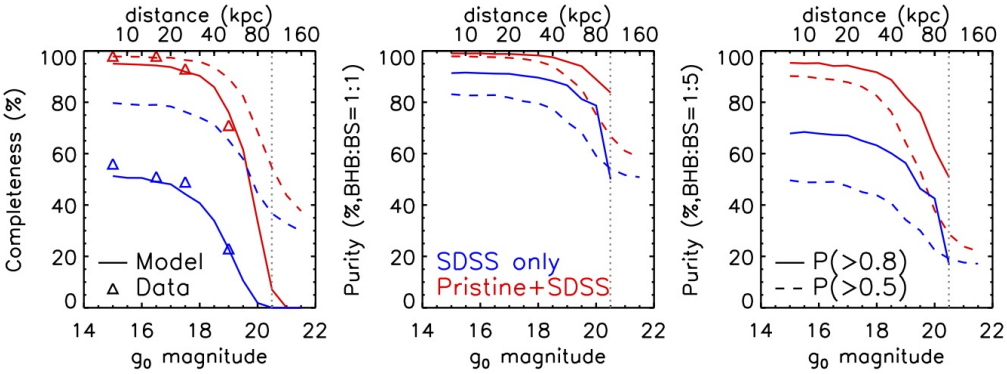

Figure 5.5 — Completeness and purity of the BHB sample evaluated as a function of magnitude for the SDSS only and *Pristine* + SDSS methods (blue and red, respectively). The triangles show the results of the overlapping spectroscopic sample of Xue et al. (2008), when binned in magnitude (the magnitude shown represents the middle of the bin), using a probability cut-off value > 0.8. The lines show a Monte Carlo model of performance based on the photometric uncertainties. Full lines use a probability cut-off value > 0.8, dashed lines > 0.5. The two rightmost panels both show results for the purity of the sample, but based on a different underlying modelled ratio of BHB to BS stars, either 1:1 (middle panel), or 1:5 (right panel). The top x-axes show a distance scale for a BHB star with $(g_0 - r_0) = -0.15$ at that apparent magnitude. A vertical dotted grey line indicates a distance of 100 kpc.

5.2.5 Future outlook

As illustrated in Figure 5.4, the crucial uncertainty in our method is not the $CaHK$ magnitude, but instead the rather shallow SDSS u-band photometry (at $g = 20$ and beyond, u-band uncertainties are significantly larger than Pristine photometry uncertainties). Excitingly, the currently on-going Canada-France Imaging Survey (CFIS, Ibata et al. 2017a,b) will provide more accurate u-band photometry in coming years. CFIS is ∼ 3 magnitudes deeper than SDSS. CFIS uncertainties in the u-band are only 0.01-0.03 mag on average (Ibata et al. 2017a) in the same magnitude range. The overlap between the *Pristine* and CFIS footprints is a few hundred square-degrees at the moment, but it is quickly increasing and will reach 100% by the end of the survey. CFIS photometry has already been used to select BHB stars very far out in the Galactic halo (Thomas et al. 2018b). With the addition of the *Pristine* photometry this will provide a golden match for photometric BHB studies providing, at the same time, unprecedented depth as well as purity and completeness.

5.3 Results: Investigating the halo with BHB and BS stars

As mentioned above, BHB stars are standard candles, which means that distances can be computed for them using their apparent magnitude. We again here follow Deason et al. (2011), who define

$$M_{g,\mathrm{BHB}} = 0.434 - 0.169(g_0 - r_0) + 2.319(g_0 - r_0)^2 + 20.449(g_0 - r_0)^3 + 94.517(g_0 - r_0)^4. \quad (5.8)$$

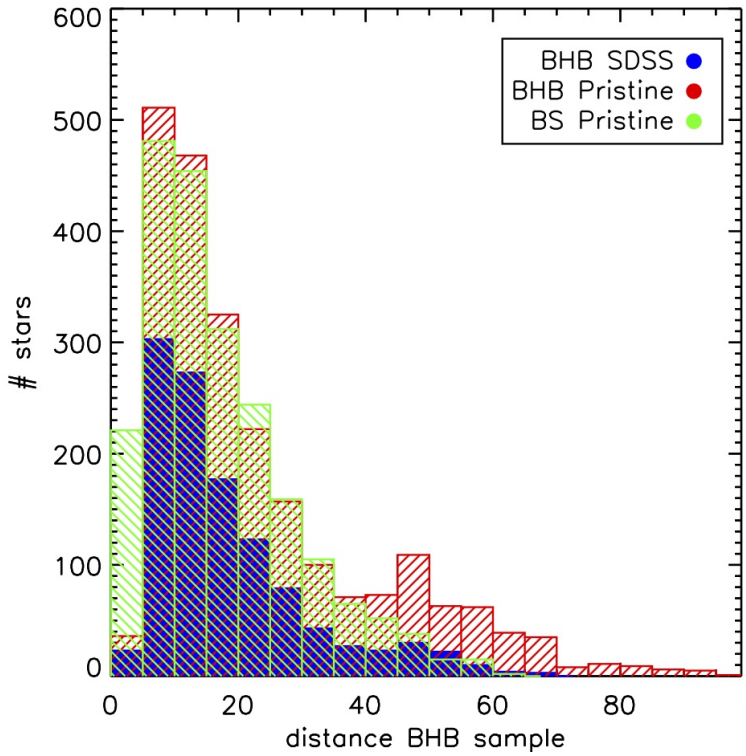

Figure 5.6 — Histogram of distances derived for the BHB population as identified in Section 5.2.1 from SDSS photometry alone (blue solid histogram) and the BHB population as well as the BS population from combining *Pristine* and SDSS photometry as described in Section 5.2.2 (red 45° and green 135° slanted histograms respectively). Targets are selected to have clean photometry. The relations to derive absolute magnitudes for these populations are given in Equations 5.8 and 5.9.

Additionally, the BS stars can be used to trace substructures (e.g., Thomas et al. 2018a). For the BS stars, we use the conversion to absolute magnitude (and hence distance) as given by Navarrete et al. (2018):

$$M_{g,\mathrm{BS}} = 2.2 + 4.557(g_0 - r_0) - 0.45[\mathrm{Fe/H}]. \tag{5.9}$$

The BS stars are less accurate standard candles and for them the distance vs. absolute magnitude relation will have a higher dispersion than for BHB stars. Additionally, the fact that they are intrinsically fainter means that they do not trace as far out and will suffer at closer distances from larger photometric uncertainties. In the absence of any robust measurement of the metallicity, we use a metallicity of [Fe/H] = −1.0 in accordance with our results in Figure 5.2. This lower metallicity will be a more accurate estimate for the BS stars further out in the halo than for those in the disk.

We note that the absolute magnitude for a typical A star will differ by about 1.5 magnitudes if it is a BHB or a BS star. As pointed out by Lancaster et al. (2019b), this means that the proper motions of the stars (if available) as measured by Gaia DR2 can provide an extra means to clean the sample. A BS star that is misclassified as a BHB star will as a result be placed much further and its tangential velocity calculated will be larger based on the same proper motion. In this work, we adopt as an extra criterion that if the tangential velocity calculated with the

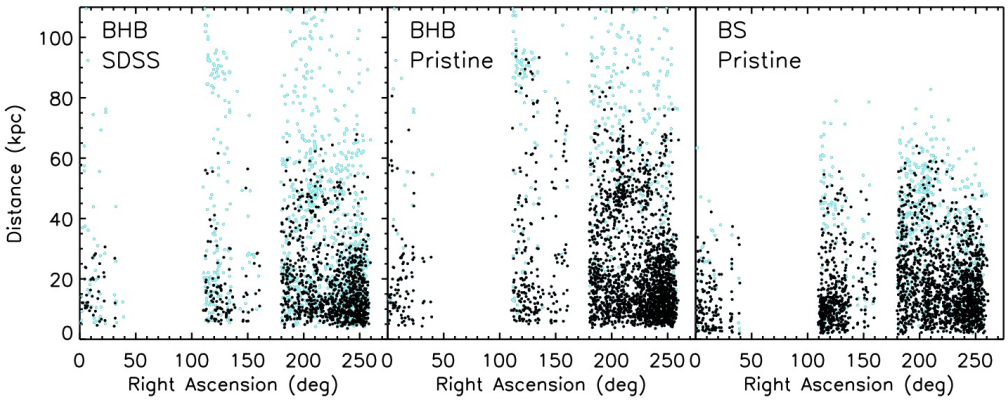

Figure 5.7 — Right ascension versus distance maps for the three different populations shown in Figure 5.6. Stars with a probability > 0.8 per class are shown in black, while stars with a probability between 0.5 and 0.8 appear as cyan points.

corresponding BHB distance for a star is above 500 km s^{-1}, a limit of 0.4 is adopted for the BHB probability. If the probability that the star is a BHB star was above this value before, it is lowered to 0.4 and the probability that the star is a BS star is raised to 0.6. This is only applied if the relative proper motion uncertainty is smaller than 0.2. The effect of this extra cleaning is small, it does not affect any stars in the overlapping spectroscopic sample and less than 1% of the stars in our total footprint (~ 2.5% of the stars with P > 0.5 of being a BHB star). We furthermore use the same colour and quality cuts as described in Section 5.2.3, but with no strict cut on $CaHK$ magnitude uncertainty (although in practice very few stars have uncertainties larger than the value of 0.1 that was used as a limit before). We verify with the second data release from Gaia (Gaia Collaboration et al. 2018) that the distances we obtain for the samples are in agreement with their independently measured parallax distances. As expected, however, most of the sample is too faint to have accurate parallax measurements. But for stars with relative parallax uncertainty smaller than 0.2, all distances are compatible with the Gaia measurements.

5.3.1 The profile of the Galactic halo

Figure 5.6 shows the histogram of heliocentric distances for three different populations: the BHB population as identified in Section 5.2.1 from only SDSS photometry and the BHB population as well as the BS population from combining *Pristine* and SDSS photometry as described in Section 5.2.2. All stars selected here have a probability of > 0.8 to belong to their class, which means that the completeness of the sample will sharply decline between ~ 50 and ~ 100 kpc (see Figure 5.5). We find an enhancement of BHB stars at all distances in the *Pristine* + SDSS sample, most likely because of the improvement in completeness. Most strikingly, there is a clear extra overdensity at 45 − 50 kpc and $200 < \alpha < 220^{o}$ in the *Pristine* + SDSS BHB sample. This feature is again seen clearly when the distances are shown as a function of right ascension in Figure 5.7. Here, stars with a probability > 0.5 of belonging to a class are shown in each panel (in cyan, 2720 stars), as well as those that have probability > 0.8 (in black, 2311 stars). As expected, the samples start to deviate at larger distances, where the photometric uncertainties are larger, and start to match the separation between the ridgelines in Figure 5.2. The larger separation between the ridgelines in the *Pristine* + SDSS sample lead to this effect being smaller at similar distances than in the SDSS only sample.

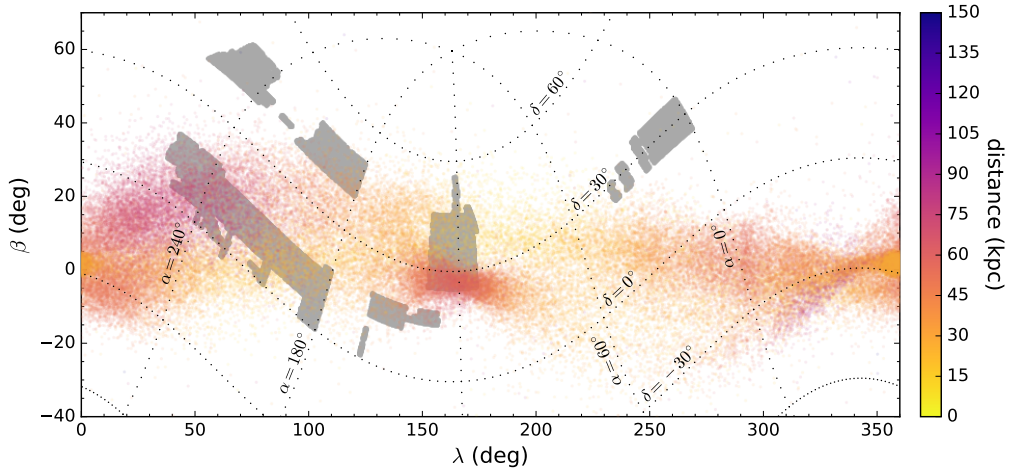

Figure 5.8 — Current footprint of the *Pristine* survey (grey filled contours) used in this work in the coordinate system of the orbital plane of the Sagittarius stream (see, e.g., Belokurov et al. 2014 for its definition). Overplotted is the Sagittarius stream modelled by Law & Majewski (2010), colour-coded by predicted distance to the various stream features.

The feature at ~ 50 kpc can be easily identified as the leading arm of the Sagittarius stream, which indeed is expected to cross the *Pristine* footprint at these distances. Another overdensity, seen at $\alpha \sim 120^{o}$ and ~ 90 kpc represents the trailing arm of the same disrupting dwarf galaxy. Figure 5.8 presents an overview of the *Pristine* footprint to date in the coordinates of the Sagittarius orbital plane (see, e.g., Belokurov et al. 2014 for its definition), with overplotted the Law & Majewski (2010) model of the Sagittarius stream and its predicted distances, matching well the observed relation with the *Pristine* footprint in Figure 5.7. Interestingly, the stream is not clearly traced in the BS star sample, but it does stand out clearly in the BHB Pristine + SDSS sample, and more clearly than the BHB stars selected from SDSS photometry alone.

Because of the overwhelming presence of the Sagittarius stream in our footprint, the *Pristine* dataset is currently not well suited to study the halo profile of the smooth(er) halo component. There has been a large body of literature already on this subject, utilising more extended and deeper datasets, including BHB stars, K-giants, or RR Lyrae (e.g., Xue et al. 2015; Xu et al. 2018; Hernitschek et al. 2018; Thomas et al. 2018b; Fukushima et al. 2018). However, given our improved statistics on completeness and purity, it is interesting to compare our findings of the halo density profile with other work. In Figure 5.9, we summarise the density profile of our *Pristine* + SDSS BHB distribution over the full footprint (filled circles) and for just the part of the footprint that is more than 30 degrees away from the Sagittarius orbital plane (a bit less than 20% of our A stars, shown as asterisks). In the full sample, clearly the influence of the Sagittarius stream at ~ 50 kpc can be detected. To address the issue of completeness, we have counted the number density of BHB stars in each distance bin in three different ways: (i) by counting the stars for which $P(\text{BHB} \mid uCaHKgr) > 0.5$; (ii) by counting only those for which $P(\text{BHB} \mid uCaHKgr) > 0.8$; (iii) by using all stars that pass the quality cuts and multiplying them with their value of $P(\text{BHB} \mid uCaHKgr)$. These three different choices are plotted with three different colours in Figure 5.9, but it is clear that the exact choice does not affect the overall results except in the outermost distance bins.

When fit with a single power-law using the least-square polynomial fit routine `poly_fit` in IDL, the density distribution favours a slope of -3.5 ± 0.1 for the full sample and a slope of -4.0 ± 0.1 for the sample off the Sagittarius plane, which is in good agreement with recent results from either RR Lyrae (-4.4 ± 0.1 Hernitschek et al. 2018), or BHB stars (measurements ranging from -3.2 to -4.0 Thomas et al. 2018b; Deason et al. 2018b; Fukushima et al. 2018), although the RR Lyrae favour a slightly steeper slope. Results from Xu et al. (2018) using a K-giant sample also favour a steeper slope of -5.0 ± 0.6 and additionally a significant flattening in the inner halo (although their preferred profile becomes nearly spherical beyond 30 kpc). Note that we do not attempt to additionally fit for the flattening of the halo. We see no strong evidence for a flattening of the slope in the inner region and/or a steeper slope at larger distances, which was reported by Deason et al. (2011) and Deason et al. (2014) who measured slopes of -2.3 up to 27 kpc, a slope of -4.6 beyond that and subsequently a strong steepening to -6 or even -10 beyond 50 kpc. When we allow for a broken power-law fit, using the fitting method outlined in Thomas et al. (2018b), the preferred break radius is at $\sim$ 73 kpc for the P > 0.5 sample, both for the full sample and the sample off the Sagittarius plane, after which the slope steepens. However, it is difficult to rule out that this is not to be attributed to the declining completeness and purity in our samples at these distances (as shown in Figure 5.5). Deeper photometry is required to investigate this further.

5.3.2 Clumpiness in the Galactic halo

Figures 5.6 and 5.7 show that there is significant deviation from smoothness in the Galactic halo between the different selected samples. Here, we aim to quantify this signal using a pair counting method to measure the amount of clustering in each sample on different scales (we refer the reader to Lancaster et al. 2019a, and Chapter 6, for a more elaborate analysis using a similar method with RR Lyrae and FGK stars, respectively).

In Figure 5.10, we show the results of a clustering analysis of the BHB samples selected with *Pristine* + SDSS against the sample selected with SDSS alone. Additionally, we include the BS star sample from the *Pristine* + SDSS selection. We select any members that have a probability > 0.5 to belong to their class. To obtain the clustering signal, 3-dimensional separations are calculated within a given distance bin. The curves shown in Figure 5.10 can be interpreted as the relative amount of clustering on a given scale and are normalised to emphasise their different shapes, rather than different numbers of stars in each class. We limit this analysis to the largest contiguous region of the footprint, for $180^\circ < \alpha < 260^\circ$, so as to avoid pair count contributions at large separations from region to region. The error bars in Figure 5.10 represent the standard deviation of 10 bootstrapping iterations of counting these pairs, each time taking a subsample of 80% of the total samples. Individual uncertainties on distances are not taken into account here, but we have tested that distance uncertainties of the order of 10% do not change the result. Typically, to measure clustering one would not only do pair counts, but also compare these to pair counts of a suitable random distribution. Given that this is a relative comparison between samples using the same underlying data quality cuts and in the same footprint, we do not have to rely on the creation of a random sample since it would be the same for all of the samples. Furthermore, creating such a (smooth) random sample is non trivial and would rely heavily on certain assumptions on the selection functions of the surveys and the underlying density gradients of the various Galactic components.

The results in Figure 5.10 are divided up in four different distance ranges. We see how in the first distance bin, up to 20 kpc, the clustering behaviour is very similar in all three samples. At $20 - 40$ kpc, there is a feature at a scale of 5 kpc for the BHB samples indicating a higher degree of clustering at small scales compared to the BS sample. At the peak of the distribution at a scale of 20 kpc, there is relatively less clustering in the BHB samples when compared to

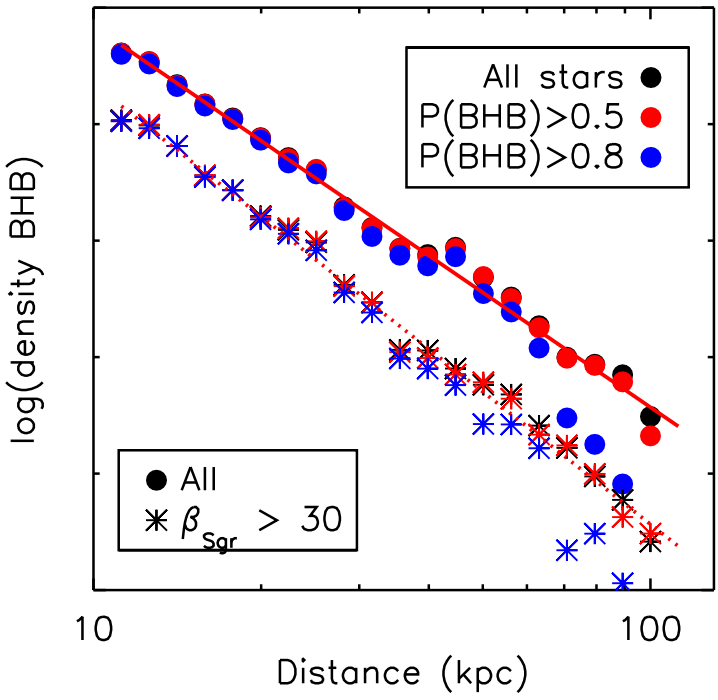

Figure 5.9 — Distance profile (in arbitrary units) of BHB stars selected in three different ways: either by taking all stars with a probability that they are a BHB star higher than 0.5 (red filled circles), or 0.8 (blue filled circles), or by taking all stars that pass the quality cuts and multiplying them with the probability they represent BHB stars (black filled circles). A fit to the red symbols is illustrated and has a negative power-law slope of 3.5 ± 0.1 (for the full sample), or 4.0 ± 0.1 for the sample at least 30 degrees off the Sagittarius orbital plane. These profile fits are shown by the full and dotted red lines respectively.

the BS stars. This is indicative that the BHB samples are better tracing the strong feature of the Sagittarius stream at these distances, which is pronounced at the edge of the distance bin, but is largely under-dense in the middle. By comparison, the BS sample is more uniformly distributed. In general, BS stars do not trace substructures as well as BHB stars, due to the larger uncertainties in their distance determinations. They also trace a younger stellar population than the BHB stars do. In the 40-60 kpc bin, all three samples show a clustering signature with the same peak, however, the BS sample shows a decrease in clustering at slightly larger scales compared to the BHB samples. Referring back to Figure 5.7, the BHB signal is again explained by a strong Sagittarius feature, whereas the BS signal comes from a different feature, characterized by a larger density of stars in the low RA half of the footprint compared to the high RA half. This either indicates some other structure that is traced by the BS stars but not the BHB stars, possibly characterized by a younger stellar population, or that there is some systematic selection effect in that region for the BS sample. Finally, in the largest distance bin ($60-100$ kpc) a hint of an effect in the clustering signal of our improved selection technique for BHB stars with *Pristine* + SDSS can be seen. Here, the sample selected on SDSS alone shows a larger clustering signal at smaller scales of roughly 40 kpc. By comparison, the *Pristine* + SDSS sample shows stronger clustering at larger separations. These differences are not very

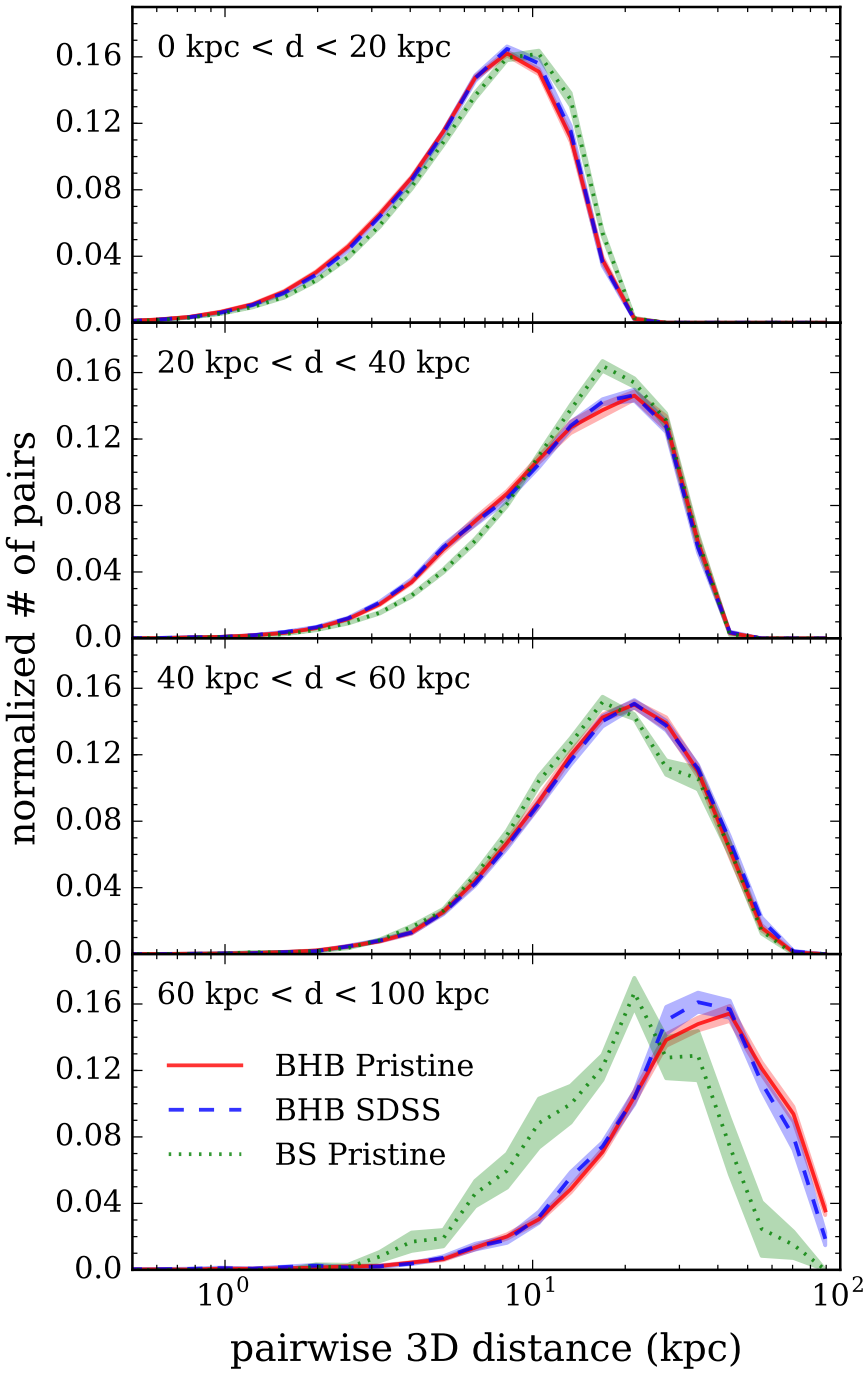

Figure 5.10 — Pairwise 3D clustering for the BHB population as identified in Section 5.2.1 from SDSS photometry alone (blue), and the BHB population and BS population from combining *Pristine* and SDSS photometry as described in Section 5.2.2 (red and green, respectively).

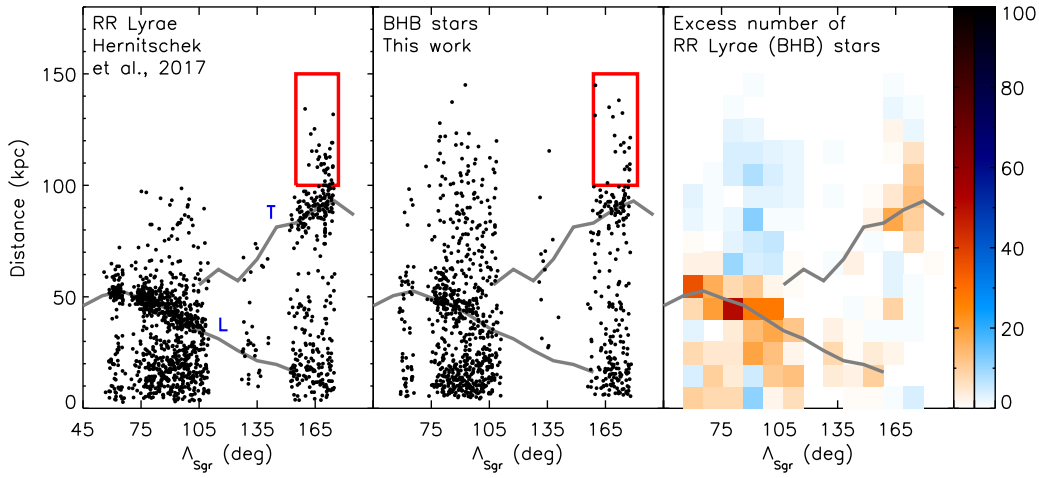

Figure 5.11 — RR Lyrae (left panel, taken from Hernitschek et al. (2017)) and BHB stars (second panel, this work, selected to have $P(\text{BHB} \mid uCaHKgr) > 0.5$) as a function of longitude along the Sagittarius orbital plane and heliocentric distance. Grey lines guide the eye along Sagittarius stream components as defined by the RR Lyrae measured by Hernitschek et al. (2017). Blue letters 'L' and 'T' indicate the leading and trailing arm respectively. Red boxes illustrate the distant "spur" feature discovered in the same RR Lyrae sample (Sesar et al. 2017a). The third panel illustrates the difference in both populations by either counting the excess of RR Lyrae stars (red colour bar), or BHB stars (blue colour bar).

significant however and need to be investigated further with larger samples and more accurate photometry at these large distances. If confirmed, the differences can be related to two things, firstly a misclassification of BS stars that should appear in the more nearby distance bins but are instead pushed into the larger distance bin in the SDSS only sample, and secondly that there is a larger number of stars at the furthest distance bin in the *Pristine* + SDSS sample (see Figure 5.6). Together, these result in a broader clustering signal which extends to larger scales for the *Pristine* + SDSS sample.

5.3.3 Tracing the Sagittarius stream further out

In Figure 5.11, we show how the Sagittarius stream is traced in the current *Pristine* survey footprint when compared with the RR Lyrae stars presented in Hernitschek et al. (2018) across the same footprint (left panel). The selection procedure for the RR Lyrae sample is described in much detail in Sesar et al. (2017b). Besides reaching distances of >120 kpc with a very high distance precision (3%), this sample is both very pure (90%) and reaches a high level of completeness (80% at 80 kpc). For the BHB sample, we use a magnitude limit for our sample to 21.5 for this analysis, allowing BHB stars in principle to be traced out to 170 kpc for the hottest stars, although probably with significant contamination. A BHB criterion with a probability of > 0.5 is adopted. Grey solid lines guide the eye along the Sagittarius main stream features as defined for RR Lyrae by Hernitschek et al. (2018). We clearly see the same structures appear in our map of BHB stars, also illustrating that the calculated distances match very well the precise distances to RR Lyrae stars.

As both populations of stars have a high completeness and purity up to 50 kpc, we can learn about the composition of these different features through the observed ratio of RR Lyrae to BHB stars. The right panels of Figure 5.11 show this comparison. It is clear that in the Sagittarius

features the signal is largely dominated by the RR Lyrae. Both RR Lyrae and BHB stars are thought to trace old stellar populations. One possible interpretation of this difference is that the outer halo in the *Pristine* survey footprint is made up from progenitor galaxies with different dominant stellar populations than the Sagittarius dwarf galaxy, for instance being smaller and more metal-poor. This would explain why the Sagittarius stream features are more dominated by RR Lyrae. Note that this could also explain the slightly steeper slope found in the halo using RR Lyrae as tracers (Hernitschek et al. 2018), when compared to BHB stars (Thomas et al. 2018b; Deason et al. 2018b; Fukushima et al. 2018). There is also a larger number of BHB stars found in the more distant regimes (most clearly visible in the region $75^\circ < \Lambda_{\rm Sgr} < 105^\circ$, between 100 and 120 kpc), but this could well be an effect of the declining performance in completeness and purity of the BHB samples.

Many authors have tried to model the Sagittarius streams, most often in Λ Cold Dark Matter (ΛCDM) cosmologies, but also some modelling has been done using alternative cosmological frameworks such as MOND (Thomas et al. 2017). No model to date has however been able to reproduce all features simultaneously (see the discussion in Fardal et al. 2019). This might point to a more complicated halo dark matter profile than is usually assumed. Additionally, in the mapping of its orbit, the internal dynamics of the dwarf galaxy and the influence of other large substructures, such as the Large Magellanic Cloud cannot always be ignored (Peñarrubia et al. 2010; Vera-Ciro & Helmi 2013) and the Sagittarius dwarf galaxy itself might perturb and alter the Milky Way potential significantly (e.g., Purcell et al. 2011; Gómez et al. 2015).

Recent observations put the apocentre distances of the leading and trailing stream at very different distances, roughly a factor two apart at $\sim$ 50 and $\sim$ 100 kpc (Belokurov et al. 2014; Sesar et al. 2017a; Hernitschek et al. 2018). Moreover, the observational data clearly hints at an extension of Sagittarius' features even beyond the furthest apocentre at $\sim$ 100 kpc. The extent and character of the "spur" feature, seen in the RR Lyrae about 30 kpc beyond the apocentre of the trailing arm itself (Sesar et al. 2017a) will provide some further crucial progress on understanding the Sagittarius stream.

While many of the modelling efforts fail to reproduce this behaviour (for instance, the model of Law & Majewski 2010, puts the furthest apocentre at $\sim$ 65 kpc), two recent modelling efforts, by Dierickx & Loeb (2017) and Fardal et al. (2019), both reproduce the observed apocentre distances of the leading and trailing stream as well as a clear "spur"-like feature. Fardal et al. (2019) interpret this feature as trailing debris from two different pericentric passages, whereby the outer material actually represents a more recent passage. In their modelling the debris reaches about 140 kpc in heliocentric distance. Dierickx & Loeb (2017) instead present a model where this feature reaches a heliocentric distance of 250 kpc, including a prediction for the velocity profile of these stars.

We are fortunate to find this intriguing "spur" feature in our footprint (marked with a red rectangle in Figure 5.11). Tentatively, we can follow this structure further out using the BHB stars instead of the RR Lyrae, but unfortunately not far enough out yet to favour the model of either Dierickx & Loeb (2017), or Fardal et al. (2019). Additionally, it is uncertain how much trust we can put in the discovery of this feature at these large distances. The completeness of the BHB sample should still be around 40%, according to our modelling shown in Figure 5.5, but the purity of the sample at these distances is unknown and could well be very poor. We will pursue a further characterisation of this feature in future work by using better quality photometry and targeted spectroscopic follow-up of candidate BHB stars at these distances.

5.4 Conclusions

In this work, we demonstrate how the narrow-band filter designed for the *Pristine* survey for its metallicity sensitivity additionally provides an excellent discriminating power to separate standard candle BHB stars from the contaminating population of intrinsically fainter BS stars (for which distance determinations are less accurate). Using this narrow-band filter in combination with SDSS u-band information the purity of the BHB-star selection stays similar while at the same time the completeness of the sample is increased from 46 to 91%.

Using this unprecedented clean and complete sample of BHB and BS stars we trace the outer halo of the Milky Way. Their distance profile follows a power-law with an almost constant negative slope of 3.5−4.0, depending on whether we select regions on or off the main Sagittarius stream. We investigate how our cleaner and more complete selection affects a quantification of the clumpiness in this part of the halo in Figure 5.10, hinting at a less tightly clumped signal at large distances than a sample that is selected solely using broad-band colours. Our mapping of the Sagittarius stream is very much in agreement with results using RR Lyrae (Hernitschek et al. 2018), as shown in Figure 5.11, illustrating however also that the BHB stars are dominating the smoother halo component while the Sagittarius stream is dominated by RR Lyrae. This likely reflects a different parent stellar population for these halo features, for instance in metallicity.

Additionally, we have laid some groundwork for subsequent follow-up of the recently discovered "spur" feature, which represents the furthest discovered (likely) feature of the Sagittarius stream. Tentatively, we trace this feature to larger distances; this, however, needs to be confirmed with better quality data. Spectroscopic follow-up could provide additional proof that these stars are indeed BHB stars as well as radial velocities of the stars in this feature, providing another avenue to constrain its origin.

Acknowledgements

We gratefully thank the CFHT staff for performing the observations in queue mode, for their reactivity in adapting the schedule, and for answering our questions during the data-reduction process. We thank Nina Hernitschek for granting us access to the catalogue of Pan-STARRS1 variability catalogue.

ES, KY, and AA gratefully acknowledge funding by the Emmy Noether program from the Deutsche Forschungsgemeinschaft (DFG). This work has been published under the framework of the IdEx Unistra and benefits from a funding from the state managed by the French National Research Agency as part of the investments for the future program. NFM, RI, and NL gratefully acknowledge support from the French National Research Agency (ANR) funded project "Pristine" (ANR-18-CE31-0017) along with funding from CNRS/INSU through the Programme National Galaxies et Cosmologie and through the CNRS grant PICS07708. ES, KY, NM, AA, JIGH, and NL benefited from the International Space Science Institute (ISSI) in Bern, CH, thanks to the funding of the Teams "The Formation and Evolution of the Galactic Halo" and "Pristine". JIGH acknowledges financial support from the Spanish Ministry project MINECO AYA2017-86389-P, and from the Spanish MINECO under the 2013 Ramón y Cajal program MINECO RYC-2013-14875.

Based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique of France, and the University of Hawaii. SDSS-IV is managed by the Astrophysical Research Consortium for the

The Pan-STARRS1 Surveys (PS1) have been made possible through contributions of the Institute for Astronomy, the University of Hawaii, the Pan-STARRS Project Office, the Max-Planck Society and its participating institutes, the Max Planck Institute for Astronomy, Heidelberg and the Max Planck Institute for Extraterrestrial Physics, Garching, The Johns Hopkins University, Durham University, the University of Edinburgh, Queen's University Belfast, the Harvard-Smithsonian Center for Astrophysics, the Las Cumbres Observatory Global Telescope Network Incorporated, the National Central University of Taiwan, the Space Telescope Science Institute, the National Aeronautics and Space Administration under Grant No. NNX08AR22G issued through the Planetary Science Division of the NASA Science Mission Directorate, the National Science Foundation under Grant No. AST-1238877, the University of Maryland, and Eotvos Lorand University (ELTE).

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss.org. SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatário Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group,Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

6 Quantifying substructure in the Milky Way halo at different metallicities

K. Youakim, E. Starkenburg, G. Matijevič, M. Krumpe, K. Venn, N. Martin, and the Pristine collaboration

— This chapter is a paper in preparation —

ABSTRACT

We perform a large scale clustering analysis of stars in the Milky Way and explore the feasibility of this method for differentiating the relative amount of substructure present at different metallicity ranges. We separate stars into metallicity bins using the photometric metallicities of the *Pristine* survey, and use the angular 2-point correlation function to quantify the amount of substructure present at different scales. We also separate the *Pristine* footprint into 6 separate regions, each sampling a different direction and galactic environment, and show that the substructure content in the Galaxy varies with observational direction. We conclude that this method and this data set are effective in recovering strong, large scale substructures, but become limited by systematics when identifying structures at very small scales. We identify a strong clustering signal at intermediate metallicities ($-2 <$ [Fe/H] < -1) in a low latitude region, which coincides with the known location and distance of the Monoceros/A13/ACS/EBS overdensity. We additionally identify a signal of enhanced clustering at small scales (< 0.5 deg, just above our detection limit) and low metallicities for the region closest to the Galactic center, that coincides with a recent discovery of a mass of diffuse material at similar metallicity detected using a stream finding algorithm in Ibata et al. (2019). Further testing and analysis are needed to robustly confirm these detections. Finally, we propose ways in which this method can be expanded upon with future data releases of the *Pristine* survey and Gaia, and discuss potential applications for Galactic archaeology studies.

6.1 Introduction

The Milky Way stellar halo has been built up over time through a complex series of accretion and mixing events. Much of the stellar content has come from the accretion and subsequent integration of other galaxies, with various masses and chemical compositions. After enough time, many of these stellar associations have been disrupted and phase mixed throughout the halo. Therefore, the stellar halo as we observe it today is highly structured, with some portion being spatially coherent in the form of star clusters, galaxies, and stellar streams, while some is moderately phase mixed and less spatially distinguishable, and some is highly phase mixed and appears as a smooth distribution of field stars. The chemical, spatial, and kinematic properties of the stars belonging to each of these components contain a wealth of information about the assembly history of the stellar halo.

One of the first and most prominent substructures discovered was the Sagittarius Dwarf galaxy and its associated stream (Ibata et al. 1994, 2001; Majewski et al. 2003). With an initial mass of a few $10^8 M_\odot$, the Sagittarius dwarf galaxy has left prominent tidal streams that wrap impressively around the entire Galaxy and extend out to large distances (Belokurov et al. 2006; Law & Majewski 2010; Sesar et al. 2017a), demonstrating very clearly the role of accretion and disruption of dwarf galaxies in the build up of the stellar halo. Since this discovery, a significant amount of effort has been invested into finding coherent, resolvable substructures in the Milky Way, including dwarf galaxies, globular clusters, and stellar streams (e.g., Belokurov et al. 2007; Grillmair 2009; Martin et al. 2014; Bernard et al. 2016; Shipp et al. 2018). Particularly in recent years, with the advent of large scale photometric surveys such as 2MASS (Skrutskie et al. 2006), SDSS (York et al. 2000), Pan-STARRS (Chambers et al. 2016) and DES, there has been a huge increase in the number of named substructures in and around the Milky Way (for compilations of known nearby dwarf galaxies and stellar streams and their properties see McConnachie 2012; Grillmair & Carlin 2016; Mateu et al. 2018). Recent work using Gaia DR2 and the STREAMFINDER algorithm (Malhan & Ibata 2018), has shown that remnant stellar streams are ubiquitous and abundant throughout the Galaxy, implying that the presence of streams may be even more prevalent in the halo than previously thought (Malhan et al. 2018; Ibata et al. 2019). This, in turn, has had a large impact on our understanding of the stellar halo and its formation, and has solidified the notion that it is full of structure and highly inhomogeneous.

In contrast to these distinct overdensities, more diffuse and larger structures that have spread more broadly over large regions of sky are harder to detect since they can not be clearly spatially distinguished from the field star population. Such structures may, however, be identifiable based on their kinematics or chemistry. For instance, recent work has shown the potential application of chemical tagging in identifying stars that, although spatially separated, have similar chemical abundance characteristics and can therefore be associated with originating from the same birth environment (e.g. Martell et al. 2016; Simpson et al. 2019). Some other studies have further shown that, using dimensionality reduction and clustering algorithms, it may be possible to link these stars back to their specific birth clusters, particularly once a large number of abundances become available from upcoming spectroscopic surveys (Kos et al. 2018; Price-Jones & Bovy 2019). For kinematic signatures, simulations have shown that even after long mixing times of > 10 Gyr, structures that are broadly distributed in density space can still be coherent in the space of integrals of motion or actions (e.g., Helmi & White 1999; Binney & Tremaine 2008; Sanderson et al. 2015). The large data set of kinematic and astrometric information provided by the recent release of Gaia DR2 has allowed for great advancements in this field. Using the full 6D-phase space information from Gaia DR2, a recent study by Helmi et al. (2018) identified a group of stars that stand out in velocity space from the bulk population of halo stars (see also Haywood et al. 2018). In particular, they identify a

prominent structure with slightly retrograde mean rotational motion, and suggest that this is the remnant of a massive satellite that merged with the Milky Way ~ 10 Gyr ago. They name this structure Gaia-Enceladus, and propose that the remnants of this accreted galaxy dominate the Milky Way inner halo that we observe today, and that this merger caused substantial dynamical heating of the resident stellar population which resulted in the formation of the thick disc. Belokurov et al. (2018) and Myeong et al. (2019) propose a similar scenario, but suggest that the inner halo is composed of stars from two distinct, large accretion events that occurred early on in the formation of the Milky Way. These are known as the Gaia Sausage, characterized by stars on highly radial orbits, and Sequoia, characterized by high energy, retrograde stars. In both of these scenarios, an inner Galactic halo that is built up of a small number of large merger events is favoured, rather than a consistent accretion of smaller galaxies.

In addition to these large accretion events, many smaller substructures have also been identified based on the angular momenta, orbits, and energies of their stars (e.g. Helmi et al. 1999; Chiba & Beers 2000; Myeong et al. 2018a; Koppelman et al. 2019). Stars can also be linked to specific substructures (e.g. Price-Whelan & Bonaca 2018; Koppelman et al. 2018), and globular clusters and stellar streams can be associated to the progenitor which brought them in based on their kinematics derived form the Gaia data (Myeong et al. 2018b,c). Detailed information about the motions of stars is thus a very powerful way to dissect the accretion history of the Galaxy, and the Gaia mission has resulted in a flurry of new discoveries in the past two years.

Several previous studies have attempted to analyze the clumpiness and substructure content of the halo in order to dissect its accretion history. For example, Bell et al. (2008) quantified the deviation from smoothness in the Sloan Digital Sky Survey (SDSS, York et al. 2000) and found a fractional RMS deviation of $\gtrsim$ 40 % between the data and a smooth, parametrized halo model, concluding that the Milky Way halo is highly structured. In another study of ~ 101 halo red giant stars from a narrow and deep pencil beam field from the Spaghetti survey, Starkenburg et al. (2009) performed a kinematic clustering analysis to quantify substructure in the Galactic halo. Even with so few stars, they were able to identify groups and associate them with known halo substructures, demonstrating the significantly non-smooth nature of the halo. More recently, this same technique has been applied to larger spectroscopic datasets from SDSS/SEGUE to quantify halo substructure using BHB stars (Xue et al. 2011) and K giants (Janesh et al. 2016). Both of these studies showed evidence for an increasing amount of substructure with Galactocentric radius. In addition, Janesh et al. (2016) also report an increasing amount of substructure with increasing metallicity, and show associations of some of their detected substructure to known structures in the halo, such as Sagittarius, the Orphan Stream, the Cetus Polar Stream, and others. Using a sample of K giants from LAMOST, Yang et al. (2019) take this even further by incorporating Gaia proper motions to compute the clustering of the halo using a 6Distance, and confirm the results of Janesh et al. (2016) that the amount of substructure increases as a function of Galactocentric radius and metallicity. Using a different approach, Lancaster et al. (2019a) lead a pioneering effort to investigate clustering as a function of distance using a 3D two-point correlation function analysis on a sample of RR Lyrae stars. Their sample has fairly complete coverage of the halo ($|b| > 30^\circ$), and a reasonable number of stars (~ 30 000 and ~ 45 000 for their two samples, respectively). With these data, they were able to show that the amount of substructure increases up to a distance of $r \sim 50$ kpc, then dies off, a result that is qualitatively consistent with Milky Way halo simulations built up entirely from accreted satellites (Bullock & Johnston 2005). McConnachie et al. (2018) pioneer a different approach of quantifying a highly structured data set on the halo of our neighbouring galaxy Andromeda. Through their use of the OPTICS clustering algorithm to characterize the hierarchical structure of Andromeda's stellar halo, they identify three new faint structures.

Taken together, these studies show that the Milky Way stellar halo, as well as Andromeda's halo, are highly clumpy and full of structure due to a long history of accretion of satellites. In this work, we push the quantitative investigation of the clustering nature of the Milky Way stellar halo further and assess how this varies as a function of metallicity, using the photometric metallicity information from the narrow-band photometric *Pristine* survey. Photometric data sets can more easily cover larger areas of sky, have a much higher completeness, and typically contain many more stars than spectroscopic data sets do. This provides a clear advantage in terms of providing a more representative interpretation of the statistical substructure content in the Galaxy.

In this chapter, we explore the feasibility of using the angular two-point correlation function in combination with the *Pristine* survey data to quantify substructure in the Milky Way at different metallicities. The goal is to recover the cumulative signals of dispersed substructures over large regions of sky, that otherwise are not detectable by traditional overddensity searching algorithms. In Section 6.2, we describe the data set and investigate the stellar content of the sample. In Section 6.3, we describe the method used for the clustering analysis, the angular two-point correlation function, and conduct tests using toy data sets to investigate the types of clustering signals that we expect to observe in the real data. We describe the production of the random comparison sample and the masking procedure in Section 6.4, and in Section 6.5 we give the results of the clustering analysis on the data. Finally, in Section 6.6 we discuss the results, the current limitations of this analysis, and the future prospects for applying this method in Galactic archaeology studies.

6.2 The data set

6.2.1 Photometric metallicities from the *Pristine* survey

The *Pristine* survey (Starkenburg et al. 2017) utilizes a specially designed narrow-band filter, mounted on the Canada France Hawaii Telescope (CFHT) on Mauna Kea in Hawaii, to target the Ca II H & K absorption lines and photometrically identify metal-poor stars. This concept is based on previous studies which have successfully used the Ca II H & K region to identify potential metal-poor star candidates without having to obtain high quality spectra for each star. This greatly increases the efficiency of these searches, as has been demonstrated using objective-prism surveys (e.g., Beers et al. 1985, 1992; Christlieb et al. 2002) and targeted photometric studies using narrow-band filters (Anthony-Twarog et al. 1991, 2000; Keller et al. 2007). Photometric metallicity measurements can be made using broadband photometry for stars close to the main sequence turnoff (MSTO), and this has been done in the past with much success (e.g., Ivezić et al. 2008). However, this photometric calibration typically loses sensitivity at low metallicities, below [Fe/H] < -2.0, although with excellent data quality this can been pushed down to the [Fe/H] ~ -2.5 regime (An et al. 2013, 2015; Ibata et al. 2017b). The *Pristine* survey, with its targeted, narrow-band filter (herafter referred to as the CaHK filter), can reliably assign photometric metallicities down into the extremely metal-poor (EMP) regime [Fe/H] < -3.0 with typical uncertainties of ~ 0.2 dex for a broad colour range, from the MSTO to the red giant branch (RGB) and the lower main sequence (Starkenburg et al. 2017; Youakim et al. 2017). The data products of this survey offer precise photometric metallicities to over 4 million FGK-type stars over $\sim 2\,500$ deg^2 in the Galactic halo ($|b| > 20$). Furthermore, follow-up spectroscopy of the metal-poor stars in this sample has revealed a purity of > 85% for stars with [Fe/H] < -2 (Youakim et al. 2017; Aguado et al. 2019). We therefore choose to separate the sample into 4 metallicity bins: [Fe/H] < -2, $-2 <$ [Fe/H] < -1.5, $-1.5 <$ [Fe/H] < -1.0, and $-0.9 <$ [Fe/H] < -0.8. The bin sizes of 0.5 dex limit the amount of contamination between adjacent bins, and also help to minimize the discrepancy between the sizes of the samples (the

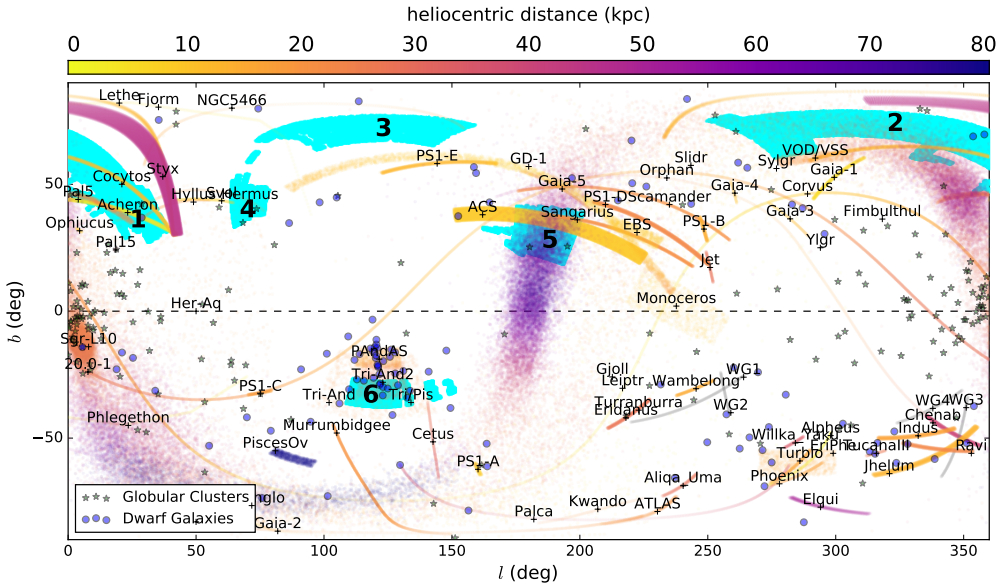

Figure 6.1 — The total *Pristine* footprint (shown in cyan), divided into 6 contiguous regions. Also shown is an up to date map of known substructures, including local group dwarf galaxies (blue circles), Milky Way globular clusters (green triangles), and stellar streams (coloured by heliocentric distance).

number of stars in each metallicity sample is given in Table 6.1). The most metal poor bin is restricted to [Fe/H] < -2 to ensure that the purity remains high, and that there are enough stars in this sample for a robust computation of the correlation function. Although photometric metallicities have somewhat larger uncertainties when [Fe/H] > -1, we can still confidently classify these stars as metal-rich, and choose the narrow range of $-0.9 <$ [Fe/H] < -0.8 to represent the metal-rich population, and also to limit this sample to a manageable size.

6.2.2 The *Pristine* footprint

Figure 6.1 shows the coverage of the *Pristine* footprint. In order to illustrate the substructures found in the different footprint regions, we also plot all known substructures using the python package galstreams (Mateu et al. 2018). The illustrated structures include the full list of local group dwarf galaxies (McConnachie 2012), the database of Milky Way globular clusters (Harris 2010, plus some new discoveries), stellar streams (Grillmair & Carlin 2016, plus some new discoveries), and the Sagittarius stream, which is plotted based on the models of Law & Majewski (2010). This work is based on the *Pristine* survey internal data release 1. Since then, the observed footprint has almost doubled, and these data are currently being homogeneously reduced and analyzed to be used in future work. We separated the *Pristine* footprint into 6 different regions to ensure contiguous data when computing the clustering signal, and to remove the irregular edges of each region (the trimmed regions can be seen in the stellar density plots of Figure 6.7 in the Appendix). This ensures that no spurious clustering signal is induced by irregular edge effects, and has the added benefit of separating the data into spatially separated regions, so that different galactic environments can be analyzed independently and compared to one another.

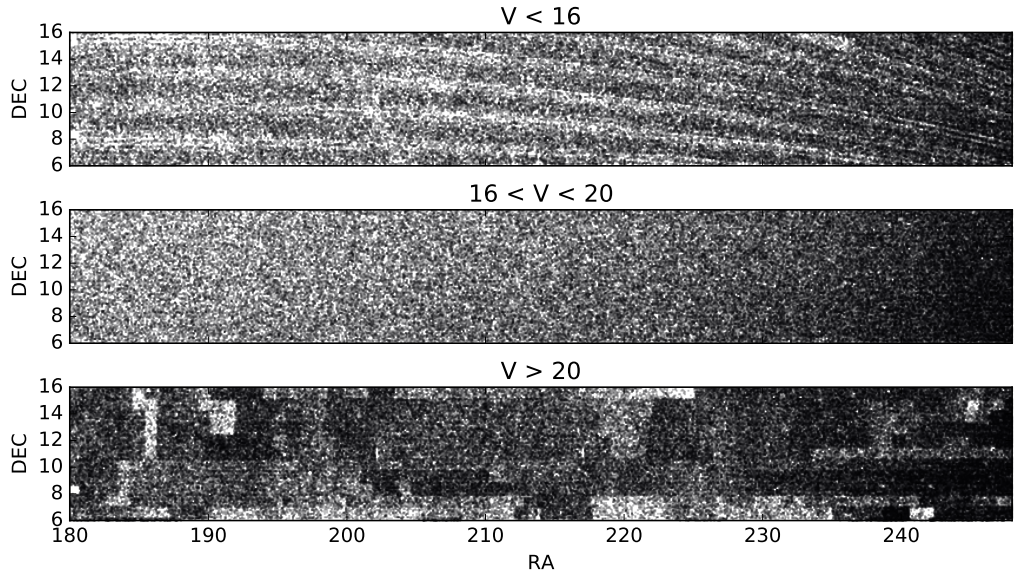

Figure 6.2 — Stellar density in different bins of magnitude for regions 1 and 2 of the Pristine footprint. The middle panel shows the magnitude cut applied to the sample, in which the stellar density appears visually smooth and is not affected by depth variations induced by the observational strategy. The top and bottom panels show the removed stars, which are brighter and fainter, respectively.

6.2.3 Quality cuts on the data

Given that we are interested in measuring the clustering signal of FGK-type stars at different metallicities, it is imperative that we remove as many artificial features in the data as possible that may introduce a systematic clustering signal. First, we take care of the differences in field depth at the magnitude limit of the survey, caused by varying weather conditions, by visually inspecting the stellar density plots for different ranges in magnitude. Figure 6.2 shows the result of this analysis for regions 1 and 2 as an example. All other regions show similar depth limits with variations of less than 0.2 mag. From this figure, it can be seen that for the bright end of the survey, there is a clear pattern of the scanning stripes of SDSS, which is likely due to variable depths of saturation from different observing sequences. At the faint end, the depth limit of the *Pristine* survey with its rectangular tiling pattern becomes apparent. For this reason, we restrict the sample to the magnitude range $16 < V < 20$, where the stellar density maps are devoid of artificial structures due to observational effects. We also implement a subset of the photometric selection criteria described in Chapter 2 (excluding the criteria for selecting only metal-poor star candidates), to remove stars with unreliable photometric metallicity determinations. These criteria remove stars flagged with bad photometry in either CaHK or SDSS magnitudes, variable objects, white dwarfs, and objects flagged as non point sources in SDSS. After making these cuts, the total number of stars in the sample is 3 642 226 (71 % of the total sample). Adding in the magnitude and metallicity cuts ($[\mathrm{Fe/H}] < -1 \cup -0.8 < [\mathrm{Fe/H}] < -0.7$), the brings the size of the sample down to 839 463 stars (16 % of the total sample).

6.2.4 What Galactic environments are being probed?

The full *Pristine* sample covers a wide magnitude and colour range, resulting in a stellar content that consists of all evolutionary types and probes a wide range of distances and

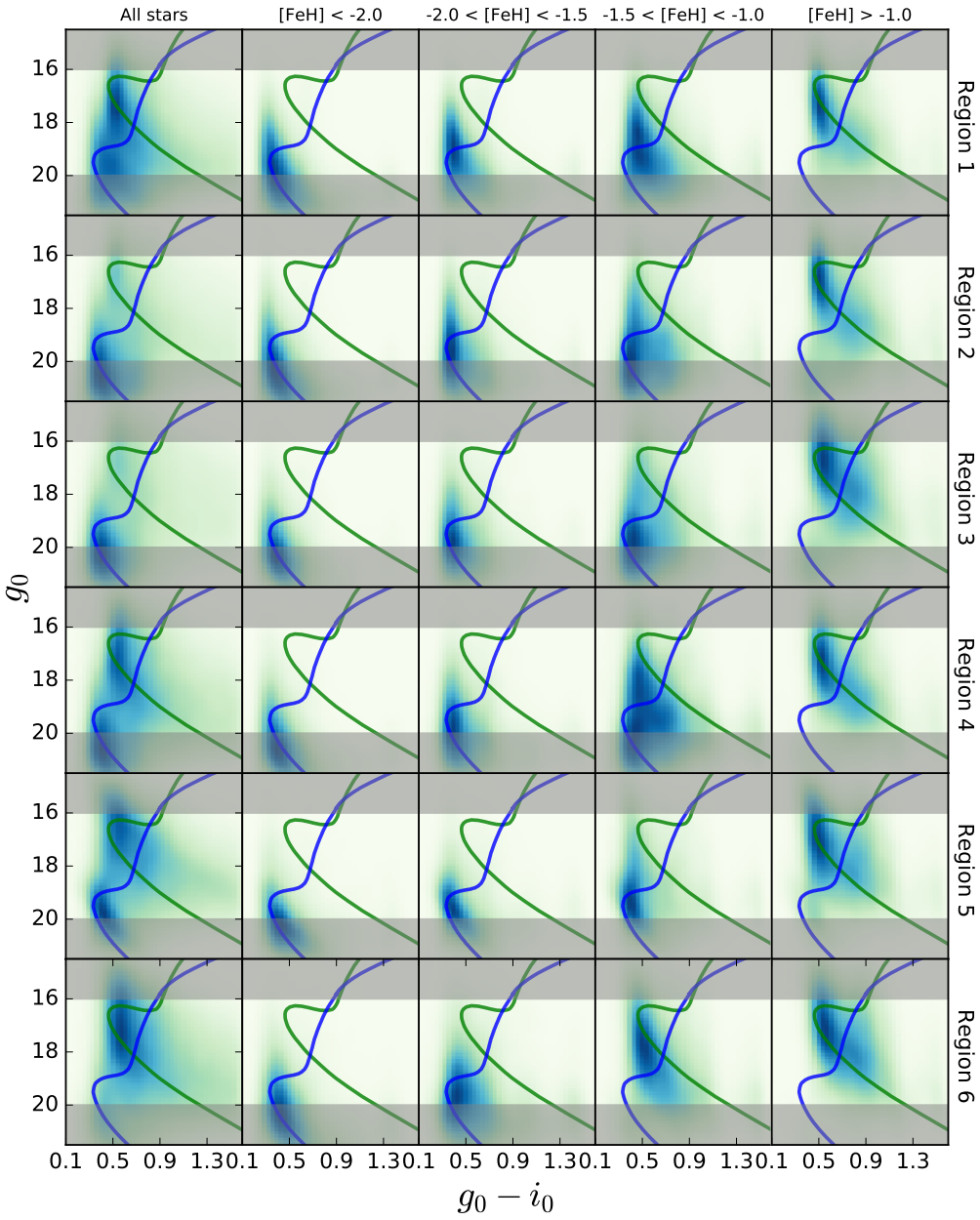

Figure 6.3 — CMDs of SDSS g_0 vs $(g_0 - i_0)$ for each region and metallicity group. The over-plotted isochrones represent a disc population (green) and a halo population (blue), corresponding to distances of ~ 3 and 11 kpc, respectively. The grey shaded regions show the magnitude cuts applied to the data.

Galactic environments. In addition, each region targets a different observational direction and contains differential contributions of galactic components, with some having relatively more disc and others relatively more halo. This is strikingly apparent from Figure 6.3, which shows the distributions in colour magnitude space with each row representing each of the 6 regions, and each column each of the metallicity samples. The magnitude cuts described in Section 6.2.3 are shaded in grey. The leftmost column shows the combined CMDs for all of the metallicity samples, with two predominant overdense regions of varying relative strengths depending on the field. Also plotted are two isochrones, generated using the PAdova and TRieste Stellar Evolution Code (PARSEC, Bressan et al. 2012). The blue isochrone corresponds to an [Fe/H] = −1.5, an age of 12 Gyr, and a distance of ∼ 11 kpc, and the green one corresponds to an [Fe/H] = −0.5, an age of 8 Gyr and a distance of ∼ 3 kpc. Based on these overlaid isochrones, the overdensities appear to be MSTO stars from two separate populations, one corresponding to a fainter, more metal-poor halo population, and the other to a brighter, more metal-rich disc population. Moving from left to right and increasing metallicity, the contribution of the sample shifts from mostly halo at [Fe/H] < −2 to mostly disc at [Fe/H] > −1. Although this may seem trivial since such metallicity cuts are routinely used to separate the halo and the disc populations, this figure shows us two important pieces of information about our sample. Firstly, that the overwhelming majority of stars - particularly in the metal-poor regime - are MSTO stars, and secondly, that they are located predominantly at a distance of ∼ 3 and 11 kpc for the metal-rich and metal-poor samples, respectively. Although some foreground and background stars are also present, distant halo giants are too sparse and the intermediate distance lower main sequence stars are too faint, meaning that the measured clustering signal can be mostly attributed to MSTO stars in the inner halo. Similarly, the metal-rich bins are made up mostly of nearby MSTO and upper MS stars, and therefore probe local substructure and features of the disc.

6.3 The two-point correlation function

In this work, we utilize the angular two-point correlation function (2-PCF) to quantify the strength of clustering for stars at different metallicities in the Galactic halo. The correlation function, $\omega(\theta)$, quantifies the excess probability of two stars having an angular separation of θ, with respect to a randomly distributed sample. Formally, it is defined as follows:

$$dP = N[1 + \omega(\theta)]d\Omega \tag{6.1}$$

where dP is the conditional probability of finding a star within the solid angle $d\Omega$ with an angular separation of θ from another star, and N is the mean number density of stars per unit area. In practice, this is typically calculated using a discrete estimator, such as the Landy-Szalay estimator (Landy & Szalay 1993):

$$\hat{\omega}(\theta) = \frac{DD(\theta) - 2DR(\theta) + RR(\theta)}{RR(\theta)} \tag{6.2}$$

where $DD(\theta)$ is the distribution of angular distances between pairs of stars in the data sample, $RR(\theta)$ the distribution for the random sample, and $DR(\theta)$ the distribution between the data and the random sample. Each term is normalized based on the number of pairs in the sample (i.e. for the DD sample the normalization factor is $1/(N_D(N_D - 1))$, where N_D is the number of stars in the data sample). To compute this, we make bins to represent a given distance scale, count the number of pairs that fall into a given angular distance bin, and then compute the correlation function according to the above equation. As can be seen from the DR and RR terms, the choice of the random background sample is crucial for a reliable clustering

estimate. Given that we are computing a two-dimensional angular correlation function over large regions of sky, the projected, large-scale density gradients coming from the halo, disc and bulge are highly variable and non-trivial. Therefore, we require a random background which appropriately models these large scale density gradients, so that we can quantify the substructure that lies underneath. The details of how this random sample is generated are given in Section 6.4.

6.3.1 Tests of the two-point correlation function

The signal that results from computing the 2-PCF is an aggregate of the combined signal over the entire sample, and therefore it can be difficult to interpret the specific features and how these relate physically to the underlying structures. In this section, we perform a series of tests on simple, mock data sets to see the size and strengths of the features which they produce in the clustering signal.

Figure 6.4 shows the results of 5 tests, investigating different configurations and strengths of overdensities and holes over the footprint. The data for each is plotted in the column with coloured points, and the grey points show the distribution of the random sample. The top row shows a simple grid configuration, with points regularly repeating with a separation of 1 deg. The resulting correlation shows the expected behaviour, that the signal produces delta functions at the specific values corresponding to the distances between points and integer multiples of these. For example, the first peak at 1 deg represents the spacing between two adjacent points in the grid, and the second peak at $\sqrt{2}$ deg represents the diagonal distance between adjacent points. The third peak at 2 deg is double the spacing between two adjacent points, the fourth at $\sqrt{5}$ deg is the diagonal distance spanned by 2 points in one direction and one in the other, and so on. This test simply validates the code and ensures that the output makes sense.

The second row shows the clustering signal produced by symmetric 2D Gaussian overdensities of different sizes, with $\sigma = 0.1, 0.5, 8$ deg for the green, blue, and magenta distributions, respectively. The overdensities consist of 240, 500, and 15 000 points, and each is combined with a random sample of 1 000 points evenly distributed over the region to simulate other stars in the data that do not belong to the cluster. Each sample is compared to a random sample of 3 000 stars to compute the correlation function. Looking at the clustering signals, we see that the signal strength begins to increase on the angular scales spanned by the overdensity, which corresponds roughly to the visual size, and is somewhat larger than the standard deviation of the Guassian distributions. There are two parameters which determine the strength of the computed clustering signal: the density of the cluster, and the relative ratio of points belonging to the cluster compared to the other points in the sample not belonging to the cluster. Corresponding to their relative point densities, the green peak is the strongest, the blue is second and the magenta the weakest. However, the differences in signal strength between the samples are smaller than they would be if point density were the only parameter of influence. Also important is the relative number of points belonging to the cluster compared to the background, which is largest for the magenta cluster, smaller for the blue cluster and the smallest for the green one. The next test, shown in the third row of the figure, demonstrates the effect of this on the clustering signal. In this panel, the three overdensities have an equal size of $\sigma = 0.5$ deg, but each have different point densities. The clusters consist of 500, 300, and 100 stars, each combined with 10 000 non member stars, resulting in membership fractions of 5%, 3%, and 1% for the blue, green and magenta samples, respectively. The right panel shows that an overdensity of this size that comprises of 5% of the total stars in the sample is easily detected, but at the 3% level it is barely detectable, and when it only contributes 1% of the stars it can not be detected at all.

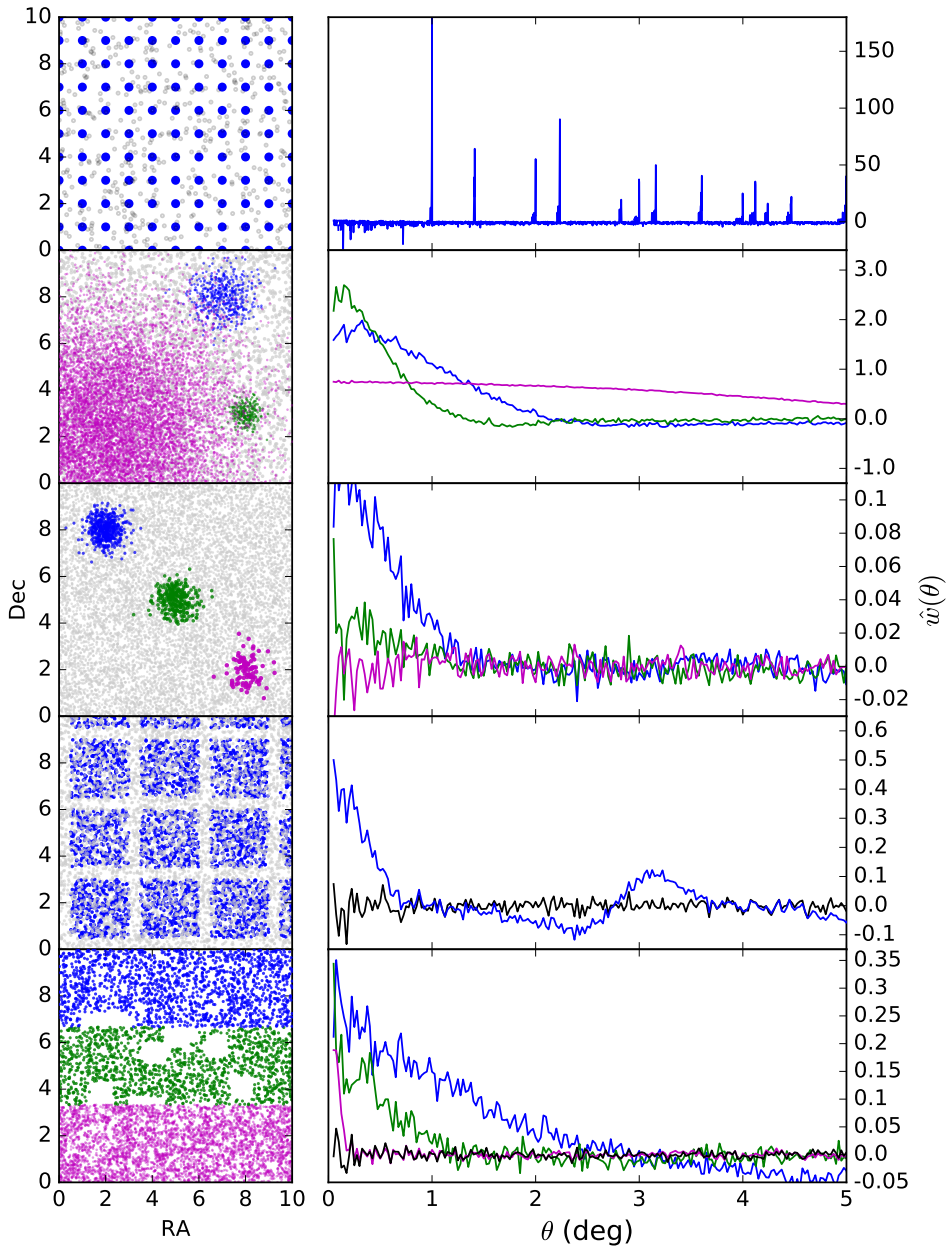

Figure 6.4 — Tests of the two-point correlation function. The coloured points in each of the panels in the left column represent the different data samples. The right column shows the corresponding clustering signal produced by computing the 2-PCF, with the colour of the lines corresponding to the sample of the same colour depicted in the left panel. The details of each of the samples used for the tests are explained in the text of Section 6.3.1.

The last two rows in the Figure show the impact of holes on the data. The second row from the bottom demonstrates the signal created by regular, grid like gaps, shown as the blue points in the left panel. The blue line in the right panel shows the resulting signal, which begins to increase on scales just larger than 0.5 deg, the width of the gaps. There is also a smaller peak at a scale of three degrees, the spacing between gaps. Additionally, the signal is negative on scales of ~ 2.5 deg, where the data sample is underdense relative to the random comparison sample, shown by the grey points. The black line shows the clustering signal when the same grid is applied to the random comparison sample. The correlation function is completely flat, consistent with what would be obtained from using two random samples, showcasing that one can correct for such features when they are well-understood. Finally, the bottom row shows the impact of different sized holes on the clustering signal. The left panel is split into three regions, each showing the scale of holes applied to the data. The blue sample has holes of 2 deg, the green sample 1 deg, and the magenta sample 0.2 deg. We see once again that there is an enhanced clustering signal on scales corresponding to the size of the holes. The blue sample also shows a negative signal at scales of 4 deg. The black sample again shows the computed clustering signal when the holes are also applied to the random comparison sample. In summary, these tests show that angular overdensities can be detected and differentiated from a smooth background using the two-point correlation function, and that the relative strength of the clustering signal depends on both the density and membership fraction of the cluster. In addition, holes in the data can produce signals that are indistinguishable from overdensities, but these can be removed completely if the holes are identically masked in the random sample. This highlights the importance to model and remove any such features as precisely and completely as possible, and to understand any potential residual structure that may be left in the data.

6.4 Creating a random sample

In the previous section, we demonstrated that clustering signals can be measured on mock data as long as they are dense enough with respect to the background stars, and an appropriate random comparison sample is used in which the holes and gaps from the data are adequately modelled. In this section, we describe the generation of the random, smooth comparison sample for computing the correlation function and the procedure used to make a mask to imprint the holes in the data onto the random sample.

6.4.1 A mock Milky Way

In order to determine the real clustering signal in the data (i.e. the signal that is above what is expected for random), we need to generate an appropriate random sample that models the underlying large scale density fluctuations. To model the complex, directionally dependant density distributions which arise from the different galactic components, we generated a smooth mock data set using the Galaxia code (Sharma et al. 2011), which implements the analytical Besançon model (Robin et al. 2003) of the Galaxy. We reproduced the depth and observational selection function of the *Pristine* survey by selecting a subset of stars from the Galaxia catalogue to match the stellar density of the *Pristine* survey. To do this, we used SDSS as a reference sample, and assumed that it is complete in the magnitude range of the *Pristine* stellar sample ($14 < V < 21.5$). We then binned the stars spatially (by each 1 x 1 deg^2 *Pristine* field), by colour (bins of 0.15 in g-i) and by magnitude (bins of 0.2 in g) and counted the number of stars in each bin. We then selected a random subset of Galaxia stars from a given bin, such that the ratio of the subset to the total was equal to the ratio of *Pristine* stars to SDSS stars

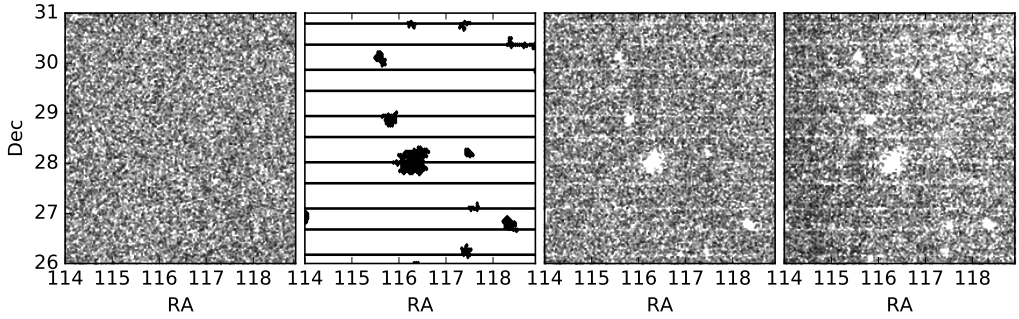

Figure 6.5 — Left panel: Random distribution of a sample region from the *Pristine* footprint. Left-middle panel: mask of the chip gaps and holes. Right-middle panel: Random sample with the mask applied. Right panel: The *Pristine* data for the same region.

in that given bin. In the case where the computed number of stars to select was higher than the total number of Galaxia stars in that bin, we simply took all the Galaxia stars, although by oversampling Galaxia by a factor of three we ensure very few bins in which this was the case. Furthermore, we note that since this version of the *Pristine* catalogue is matched with SDSS, there were no bins in which there were more *Pristine* stars than SDSS stars. Following this procedure, we ended up with a Galaxia generated sample which has the same coverage as the *Pristine* footprint, and mimics the survey's magnitude and colour distributions. This catalogue therefore reproduces the coverage, depth, and large scale density gradients of the *Pristine* survey, but does not contain any small scale substructure, making it a suitable random background sample for the 2-PCF analysis.

Additionally, we have experimented with the use of alternative background models that are fully data driven, deriving the large-scale density background from a smoothing of the data itself. These are described in the Appendix, where it is also demonstrated that the main results of this work are robust against the exact choice of background used, as long as it provides a reasonable description of the underlying density structure.

6.4.2 Including holes and chip gaps

As we demonstrated in Section 6.3.1, holes in the data can induce spurious signals in the computed correlation function, with the scale of the signal corresponding to the size of the holes. However, proper modelling of these holes in the random sample will remove this signal effectively. The *Pristine* data is cross matched with SDSS broad-band imaging in order to derive metallicities, and as such adopts any holes or peculiarities present in the SDSS footprint, in addition to those already contained in the *Pristine* data. Together, these include holes of various sizes due to saturation from nearby bright stars, photometrically unresolved overdense regions, and missing fields or gaps between adjacent fields. The CFHT/MegaCam fields are 1 x 1 deg^2, and consist of a mosaic of 36 CCDs, with 80 arcsec gaps between the rows and 13 arcsec gaps between the columns. In selecting the 6 regions, we removed any portions with high incompleteness and many missing fields, such that the data is cleared of any blatant features. The smaller holes, however, need to be modelled and added into the random sample in order to remove any spurious signal that could result from them.

We create a mask of these features in the data, taking into account the holes and chip gaps. For the chip gaps, their positions can be identified precisely based on the pointing coordinates of each field, and the well characterized specifications of the dimensions of the detector. We

therefore simply remove all stars with coordinates that fall within the width of each gap. The mask for the holes is more difficult to generate, since there are regions in which the data is sparse, resulting in naturally occurring holes (real structure that we want to keep) and artificial holes (which we want to remove) having the same size. For this reason, we are fundamentally limited in the hole size that we can model and correct for in the random data. To create this mask we first compute a list of healpixels to cover the entire footprint area, and then identify the subset of these that do not contain any data. Each empty healpixel then represents a hole, and is used as a seed to search for neighbouring empty pixels. If no empty neighbouring pixel is found, then this healpixel is removed from the list. If a neighbouring pixel that is also a hole is found, its index is added to the first to make a larger hole. We iterate over each empty pixel in this way until all pixels are either in groups with neighbouring empty pixels, or identified as single empty pixels. We then select a threshold for number of pixels, or hole size, above which we consider a hole to be real or artificial. After checking the performance of masks using NSides of 512, 1024, and 2048, corresponding to a mean spacing per pixel of 0.1145, 0.0573 and 0.0287 deg, respectively, we found that an NSide of 2048 and a minimum hole size of 10 healpixels were the best parameters for making the mask. Therefore, for holes that are near square or circular, the approximate minimum hole diameter that can be modelled with this mask is 0.0860 deg or 5.2 arcminutes. We therefore limit our analysis and interpretation of the data clustering signal to scales larger than this. Figure 6.5 shows the form of this mask and its implementation on a random sample for a selected region of the data.

6.4.3 Overlapping edges of the field

As can be seen in the two rightmost panels of Figure 6.5, there still is some overall residual difference between the random distribution and the *Pristine* data that seems to show slight density differences on the scale of the *Pristine* field size. In particular, some fields have a slight overlap with adjacent fields due to the telescope pointing. A low level of this signal persists in some regions despite our attempt to mitigate it with a more restrictive magnitude cut on the stars (as was shown in Figure 6.2) and by modelling it into our random sample by making mock observations of our Galaxia catalogues using the same field sizes and similar colour and magnitude distributions. In principle, if our survey would be complete to the magnitude level of this work (as would be expected), this should not be an issue, as the same stars would be targeted in both fields and these duplicates would be taken out of the overlapping region. The regular and repeating over- and under-densities on the scale of the field size (1 deg) in some regions of the footprint do, however, suggest that this is not working perfectly. When strongly present, these regions repeat periodically every degree (the size of a *Pristine* field) and have a width of $\sim 0.1 - 0.2$ deg. Therefore, any clustering signal generated from this feature would be induced on very small scales (< 0.2 deg) or at whole integer degree scales, and should affect all metallicity bins equally. A more homogeneous, full data reduction and cross-match of the full *Pristine* footprint is in progress and is expected to lessen many of these concerns. In the meantime, we keep this limitation in mind when interpreting our results.

6.5 Clustering results

Figure 6.6 shows the results of the 2-PCF analysis for each of the 6 regions of the *Pristine* survey footprint that were identified in Figure 6.1. The y-axis is the result of the angular correlation function computed with Equation 6.2 and represents the relative strength of the clustering, and the x-axis shows the angular scale of the clustering. The signals on scales < 0.1 deg are below the scale at which holes in the data could be properly modelled, and thus these features are

omitted from the analysis. While one would naively expect these signatures to show up at all metallicities, these are particularly prominent in the [Fe/H] < -2 bin, which shows this sharp feature at very small scales in every region. Furthermore, the periodic signal that occurs at every integer degree is also likely a systematic feature caused by repeating gaps and overdensities on the scale of the field size (as discussed in Section 6.4.3), and is also omitted from the analysis. For a detailed analysis of the signals in each region, we use the combined information from Figure 6.6 with Figures 6.1 and 6.3 to compare the spatial distributions of known substructures to the stellar content probed by a given region and metallicity bin, and assess to what degree these could be causing the observed signals.

6.5.1 Regions 1 and 2

Regions 1 and 2 show a stronger clustering signal at scales roughly smaller than 2 degrees for the metal-poor sample at [Fe/H] < -2 when compared to the other metallicity samples. For these regions in particular, this signal extends beyond the small scale for which we can not model holes in the data (as discussed in Section 6.4.2) and this peak is broader than the one present in other regions. We do note however, that upon closer inspection of the stellar density maps in region 1 for the [Fe/H] < -2 sample, we can visually see the a faint signal of overdense regions due to the pattern of observed *Pristine* fields (as discussed in Section 6.4.3). The impact of residual structure in the data due to overlap regions is expected to affect all metallicity scales, particularly since all hole masking and removal of systematic structure in the data is done on the whole sample without any consideration of metallicity. Therefore, the fact that we see a relative difference between the separate metallicity samples is suggestive that this signal is real, but differentiating with confidence whether these signals are real or residual systematic structures still present in the data requires further testing on the larger, homogeneously reduced data set.

It is intriguing nevertheless to discuss the possible nature of these signals, if they are indeed physical clustering signals from Galactic structure. As can be seen in Figure 6.1, the Hercules dwarf galaxy and the globular cluster Palomar 14 are both contained within region 1, and Boötes I (Boo I) and II dwarf galaxies are contained in region 2. Boo I is the strongest of these overdensities in the footprint and can be seen by eye in the [Fe/H] < -2 sample. However, even though it is also the nearest of these objects, it is still too distant to be sampled down to the MSTO ($\sim$ 60 kpc), and therefore consists mostly of a sparse number of giants. Within twice the tidal radius of the galaxy (26'), there are only $\sim$ 315 stars, accounting for 0.7% of the total stars in region 2 at [Fe/H] < -2. Referring back to the middle panel of Figure 6.4, for an object about twice the size of Boo I, it is undetectable below a density of $\sim$ 3% of the background. Therefore, this suggests that although Boo I and the other known clusters in regions 1 and 2 may contribute a very small amount to the observed clustering signal, they can not account for the whole signal. In addition, Hercules and Palomar 14 have higher metallicities and would not be preferentially selected in the [Fe/H] < -2 metallicity bin. To confirm this, we reran the correlation function again with these known objects excised from the footprint (making sure to excise the same sized holes from the random sample), and indeed this results in no change in the resulting clustering signal. In addition to the obvious compact objects, region 1 also contains part of the Cocytos and Archeron stellar streams. However, despite being metal-poor and nearby, these streams are substantially more diffuse than Boo I and thus their contribution to the signal is likely small.

In a recent paper, Ibata et al. (2019) report the discovery of several new streams in the inner halo using Gaia DR2 data and the STREAMFINDER algorithm. In addition to the new streams of high significance that they discover and name in their high proper motion maps, there are some kinematically coherent, low proper motion features that are too diffuse to be

resolved as individual streams, but that are still detected with $> 8\sigma$ confidence at [Fe/H] $= -2$ in the distance range $1 < d_{\odot} < 10$. Interestingly, one of these features is located at $0^{\circ} < l < 50^{\circ}$ and $20^{\circ} < b < 45^{\circ}$, and coincides with our region 1 (see Figure 5 in Ibata et al. 2019). The authors note that these features do not appear in the comparison data used as a noise model which suggests that they are real structures rather than just spurious detections, but that further investigation is required to determine the nature of these structures. Therefore, one possible interpretation of the clustering signal seen at small scales in region 1 is a diffuse conglomeration of disrupted material, possibly due to a an unresolvable group of streams, or low-metallicity debris from a larger satellite merger.

In region 2, we see a clustering signal that is above 0 at all scales for all metallicity bins. This is indicative of a discrepancy in the large scale density distributions between the data sample and the Galaxia smooth random sample. Figure 6.7 in the Appendix shows the stellar density maps of the *Pristine* data and the various random samples, with the Galaxia mock sample shown in the second panel. Although it seems the smooth density distribution in this region is reasonably well modelled, the size and the wide range of Galactic longitudes that it spans result in a more complicated density gradient due to a mixture of disc and halo populations. These discrepancies can be enhanced in each of the metallicity bins due to the changing relative contributions of disc and halo at different metallicities (the details of this are discussed further in Section 6.6.1). There is also a relative enhancement in the clustering signal for the $-1.5 <$ [Fe/H] < -1.0 metallicity sample (shown by the cyan line in panel 2 of Figure 6.6) as compared to the other metallicity bins. If a smoothed background is used (see Appendix), rather than the Galaxia model, this feature disappears. It is therefore likely that this metallicity range is even more sensitive to the discrepancies between the modelled disk and halo components in Galaxia. Additionally, it is also likely that there is some real signal here too, as this is the expected metallicity and scale of the Sagittarius stream which passes directly through this region. However, the wrap of Sagittarius is at a distance of ~ 70 kpc in this part of the sky, and would therefore only be traced by sparse distant halo giants in the sample that are greatly outnumbered the more nearby foreground. In future work, we plan to analyze this region further using a much cleaner halo giant sample only (see discussion in Section 6.6.2).

6.5.2 Regions 3, 4, and 6

Regions 3 and 4 are high latitude halo fields that are far from the disc and the main path of the Sagittarius stream. From the leftmost column of Figure 6.3, it can be seen that region 3 is predominantly populated with halo stars while region 4 has a higher contribution of disc stars. Apart from a noisy clustering signal at small scales in the metal-poor bin of region 3, the clustering analysis from these two regions shows no detectable signal for any of the metallicity samples up to the largest measured scale of 5 degrees. Referring to Figure 6.1, these two regions are entirely devoid of any known substructures apart from the globular cluster NGC 6229 with [Fe/H] $= -1.5$ and $d \sim 31$ kpc in region 4. Also in region 6, we see no clear clustering signal and no relative differences between the metallicity samples. These three regions are the smallest and have the least number of stars (see Table 6.1), and thus will really be helped by the expansion of the *Pristine* footprint in further (internal) data releases.

6.5.3 Region 5

Region 5 is at low Galactic latitude in an anti-center direction, in a region of the Galaxy known to contain a complex structure consisting of the Monoceros Ring, Anti-Center Stream (ACS), Eastern Banded Structure (EBS) and A13. These structures, as well as the current theories for their formation were previously discussed in Sections 1.3.2 and 4.5.2 of this thesis, and we

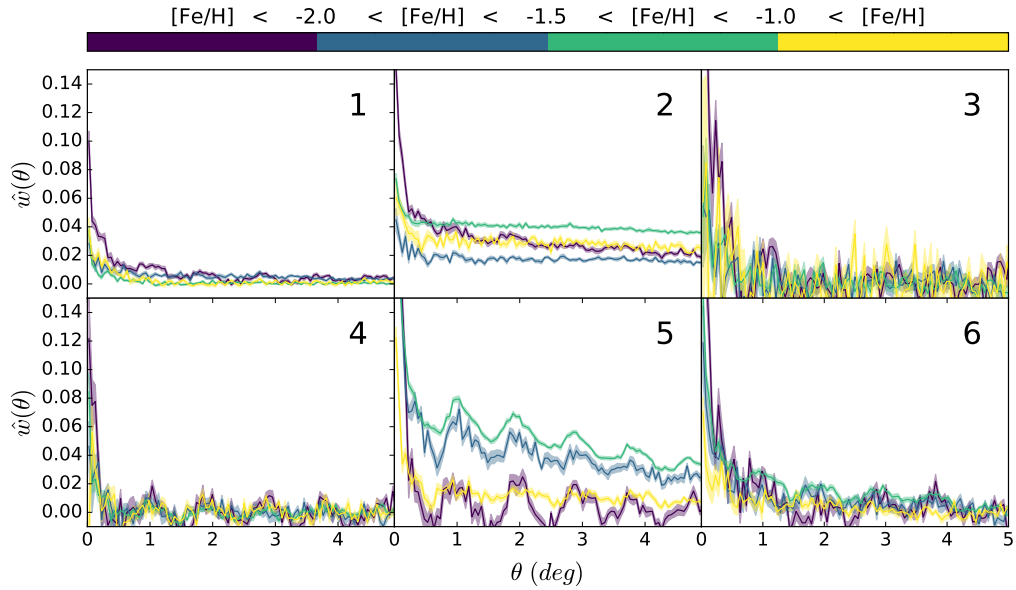

Figure 6.6 — 2-PCF clustering analysis of the 6 *Pristine* footprint regions using the Galaxia generated random sample.

refer the reader to those sections for more information. The large degree of clustering seen for the $-2 <$ [Fe/H] < -1 sample in region 5 is consistent with these features, particularly at these intermediate metallicities. Referring back to Figure 6.3, the majority of the stars in this metallicity range are concentrated near the turnoff of the blue isochrone, which is shifted to a distance $\sim$ 10 kpc, corresponding nicely to the expected distance range of these observed overdensities.

6.6 Discussion and conclusions

In this chapter, we have demonstrated the potential capabilities of the two-point correlation function as a method for quantifying substructure in the Milky Way. The advantage of such a method is that it is highly quantitative, facilitating easier comparison with cosmological simulations, and that it is particularly suitable to reveal diffuse and partially mixed structures that are characteristically different to the bound substructures, or tidal streams that other methods are well-suited to recover.

We decompose the Galaxy into metallicity bins using photometric metallicities from the *Pristine* survey. Most strikingly, we see a signal of enhanced clustering at small scales for the most metal-poor bin of [Fe/H] < -2 in regions 1 and 2, and a signal of large scale broad substructure in region 5 at intermediate metallicities of $-2 <$ [Fe/H] < -1. We tentatively associate these detections with a diffuse clump of structure identified in Ibata et al. (2019) and the large, low latitude, anti-center Monoceros/A13/ACS/EBS feature. However, a firm conclusion on the nature of these features has to await a more homogeneous data reduction and more rigorous testing of possible residual systematic features in the data. In the remainder of this section we discuss the current limitations and future directions that could be explored to improve this work.

6.6.1 Creation of the random sample

As shown in Sections 6.4.1 and 6.5, the results of our method depend heavily on the creation of a realistic random comparison sample. In this work, we chose to use the Galaxia code to generate a smooth mock *Pristine* sample with which to compute the 2-PCF. This approach has several advantages in that it produces stellar populations using physically motivated models for different Galactic components that are independent of the data, and it is relatively straightforward to imprint the footprint, field size and other properties of the survey onto the random sample. We then sampled this model to produce a mock catalogue which is an even more realistic representation of the underlying, smooth, large-scale density fluctuations in the data. However, using Galaxia does have some limitations, namely that the model uses a simplified metallicity distribution function (MDF) consisting of two gaussian components, one for each of the halo and disc populations. Although this describes the total MDF of the *Pristine* sample reasonably well, the relative contributions of the disc and halo can vary substantially between the data and the model for a given metallicity bin, and for a given observed region. This can be seen in Figure 6.7, which does not show each metallicity bin, but shows that the global stellar density is not well matched between the two samples in regions 2, 5, and 6, where there is a larger relative contribution from the disc. Due to the complexity of the Milky Way, its components can not be perfectly fit by a simplified model, and therefore some level of such discrepancies are unavoidable and would likely persist regardless of the model used.

In the Appendix, we discuss alternative methods for producing this random sample, including smoothing the data with a gaussian kernel to eliminate structure on small scales. One disadvantage to this approach is that the choice of kernel size is critical, but there is no objective way to determine the optimum size. Since the data is being used to generate its own random sample, selecting a kernel that is too small can cause substructures in the data to be imprinted on the random sample, and thus wash out the signal in the correlation function. Using too large of a kernel loses all density information of the original data set, especially when the kernel size approaches the size of the region. In addition, this method washes over many observational features such as overlapping regions or gaps between observed fields, although gaps can be masked in later. On the other hand, a large advantage to this methodology is that it can be easily applied to any data set without having to carefully model the underlying density fluctuations. Given that our main results are robust to either choice of background model, this suggests that we are probing the underlying characteristics of our data set, and that the results do not depend strongly on specific choice of the background model.

6.6.2 From 2D to 3D

One clear concern about analyzing structures in 2D are the projection effects that can cause stars separated in distance to seem falsely associated in clusters. In addition, physical associations show up at different angular scales depending on their distance, which complicates the interpretation of the signal found in a 2D analysis, as presented here. Furthermore, the various metallicity samples may probe slightly different distance ranges, causing the physical scales we are probing at various angular separations to change for the various metallicity bins.

A related, but probably more minor effect, is that stars with smaller Ca H & K absorption lines will appear brighter in the wavelength region transmitted by the *Pristine* narrow-band filter, such that for a given magnitude in g, hotter and/or more metal-poor stars will be observed to a greater depth than cooler and/or more metal-rich stars. We investigated this effect in detail in Section 4.6.1, and showed that it can cause a magnitude discrepancy of $0.2 - 0.5$ mag between metal-poor and metal-rich stars. This effect is relatively small, given the large magnitude range that we probe, but still needs to be considered as a source of uncertainty.

The obvious way to curtail these projection effects is to perform such analyses in higher dimensions, adding in distances for 3D spatial information, and even velocity and proper motion information if available. Several studies have already explored this, but these analyses are limited to spectroscopic data sets for which distances are known to high precision, and thus target specific stellar types that can be used as standard candles, such as BHB stars, RR Lyrae, and K-giants (Starkenburg et al. 2009; Xue et al. 2011; Janesh et al. 2016; Lancaster et al. 2019a; Yang et al. 2019; Starkenburg et al. 2019).

In contrast to these spectroscopic studies, the clear advantage of using photometry is the availability of much larger samples, simpler selection functions, and more complete coverage, although this comes with the trade-off of having less information per star. Naturally, if the benefit of these large samples could be combined with reliable distance information, this would be much more powerful. We have set the first steps in this direction, using Gaia DR2 parallax information to attempt to separate the sample of cooler stars into foreground dwarfs and more distant giants. We then made the same cuts on our Galaxia mock *Pristine* catalogue to check the resulting true distance distributions. Although these cuts did remove some amount of the foreground disc giants, they were not sufficient to make a clean enough giant sample with which to perform a clustering analysis. The limits of such selections were recently explored in more detail by Deason et al. (2019) using both parallax and proper motion selections from Gaia DR2. Even when they limit their study to a bright sample of stars ($14 < G < 17$), they find that they can recover the true number of RGB stars only to $\lesssim 30\%$ in $\sim 70\%$ of their selection bins in spatial, colour and magnitude space. For our purposes, this means that we still will have a high level of contamination and additionally need to induce complex selection functions in spatial, colour, kinematic, and magnitude space. We therefore leave a further exploration of a 3D correlation function to future work, when further Gaia data releases will enable a selection between giants and dwarfs with higher completeness and purity.

6.6.3 Comparisons to simulations

Galaxy simulations have become more and more successful at producing realistic Milky Way-like galaxies, particularly using zoom-in, magneto-hydrodynamic and high-resolution techniques. Recent work has taken this one step further, and produced observable mock stellar catalogues of these simulations (Lowing et al. 2015; Grand et al. 2018; Sanderson et al. 2018). These simulated data sets contain a wealth of substructure from the accretion of satellite galaxies, and may be useful for comparison with the Milky Way. A direct one to one substructure comparison may not be very meaningful, since the specific substructures in a given halo are dependent on the stochastic accretion and mixing events during the formation and evolution of the galaxy. However, using a method like the two-point correlation function to quantify substructure may be an appropriate way to make such a comparison, because this allows for a comparison of the global substructure properties of the simulations and the real Galaxy. Furthermore, using relative differences in the metallicity samples of each galaxy may further enhance the feasibility of such a comparison.

A careful comparison with simulations may also provide an effective way to investigate the projection effects associated with a 2D analysis, given that the distances are known to the stars in the simulations. Therefore, both 2D and 3D analyses can be performed, and insight can be gained on the degree to which 2D clustering signals are biased by projection effects.

Despite the advances in simulations and techniques to produce mock stellar catalogues from them, there still are several caveats to such an analysis. First, a lot of the available simulations do not reproduce basic properties of the Milky Way such as disc to halo mass fractions and realistic MDFs of the different components. Furthermore, the conversion of massive stellar particles into individual stars retains the phase space information of the parent

particle, and thus resolution of structures is limited to the resolutions and relaxation lengths of the initial simulations. It is important to understand these properties of the simulated mock catalogues in detail before attempting such an analysis and comparison. If done carefully, this may still be an enlightening exercise, but is outside of the scope of this chapter and is left for future work.

Acknowledgements

KY and ES gratefully acknowledge funding by the Emmy Noether program from the Deutsche Forschungsgemeinschaft (DFG). NFM gratefully acknowledges support from the French National Research Agency (ANR) funded project "Pristine" (ANR-18-CE31-0017) along with funding from CNRS/INSU through the Programme National Galaxies et Cosmologie and through the CNRS grant PICS07708. ES, KY, and NM benefited from the International Space Science Institute (ISSI) in Bern, CH, thanks to the funding of the Teams "The Formation and Evolution of the Galactic Halo" and "Pristine". KV thanks the Natural Sciences and Engineering Research Council for funding through the Discovery Grants program and the CREATE program in New Technologies for Canadian Observatories.

Based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique of France, and the University of Hawaii.

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss.org. SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatário Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group,Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

6.7 Appendix - Alternate random samples

In order to investigate the impact on the resulting signal, we also tested alternative methods for generating the random sample. Instead of using a galaxy model to replicate the smooth, large scale density fluctuations in the data, we smooth each *Pristine* region with a 2-dimensional gaussian kernel to generate a random sample which still preserves the large scale density variations over the size of the field, but which is smooth on small scales. As detailed in Section 6.6.1, the choice of the size of the kernel is quite critical in this approach. Here, we found that

Table 6.1 — Number of stars in each of the samples used for the clustering analysis.

	[Fe/H]	Pristine	Galaxia	1 deg	5 deg	10 deg	random
Region 1	< -2	29 846	57 717	279 771	553 121	518 856	55 444
	$[-2, -1.5]$	31 517	75 324	296 623	584 400	548 179	57 459
	$[-1.5, -1]$	52 544	207 676	493 624	971 362	904 419	86 671
	> -1	24 130	97 974	226 738	445 421	410 174	35 112
Region 2	< -2	44 451	75 899	382 285	770 933	778 544	139 625
	$[-2, -1.5]$	45 178	99 547	388 949	785 328	793 780	143 851
	$[-1.5, -1]$	63 406	168 505	548 505	1 111 239	1 134 943	217 741
	> -1	23 022	65 719	199 860	406 912	424 802	88 650
Region 3	< -2	4 903	6 525	29 151	57 624	56 406	8 766
	$[-2, -1.5]$	4 789	9 199	28 534	57 006	55 170	8 594
	$[-1.5, -1]$	6 463	12 781	38 534	76 278	74 461	11 418
	> -1	2 239	4 424	13 512	26 794	25 787	3 970
Region 4	< -2	5 585	8 611	52 590	104 172	101 413	15 695
	$[-2, -1.5]$	5 995	13 045	56 289	111 948	108 111	16 808
	$[-1.5, -1]$	10 245	32 342	96 233	191 769	185 528	28 610
	> -1	5 297	14 762	49 757	99 524	95 726	14 854
Region 5	< -2	9 839	16 168	89 684	179 392	175 200	26 785
	$[-2, -1.5]$	14 952	21 219	136 917	273 254	266 757	40 692
	$[-1.5, -1]$	42 049	61 777	387 199	772 612	753 550	114 570
	> -1	22 107	31 249	203 698	405 768	396 581	60 225
Region 6	< -2	7 065	10 954	64 974	129 701	126 963	19 610
	$[-2, -1.5]$	7 120	14 230	65 970	131 971	129 241	19 941
	$[-1.5, -1]$	21 692	40 442	201 860	403 724	395 259	60 939
	> -1	11 287	19 074	105 129	212 255	208 645	32 043

using a kernel with a standard deviation of 5 degrees adequately smooths on scales smaller than this, while still preserving the large scale density fluctuations caused by the underlying galactic components. We also compute the correlation function using 1 and 10 degree gaussian kernels to test the impact on the signal when too small or too large of a kernel is used. The results of this are shown in Figures 6.8, 6.9, and 6.10, and are discussed in the main text of the chapter.

In addition, we also generated a random sample using the simplest case of a completely random sample of stars. To generate this, we first separated the sample into metallicity bins, and generated three times the number of points in the data sample for each, taking into account projection effects as a function of increasing declination. Naturally, this approach does the poorest job of matching the underlying large scale density gradients in the data, as can be seen from the large scale clustering signals produced in some of the regions and metallicity bins in Figure 6.11. This is particularly true for the metal rich bin (yellow line) and in regions which span a larger range in either Galactic longitude or latitude, such as regions 1, 2 and 5, since in these regions there is the strongest influence of density gradients from the disc.

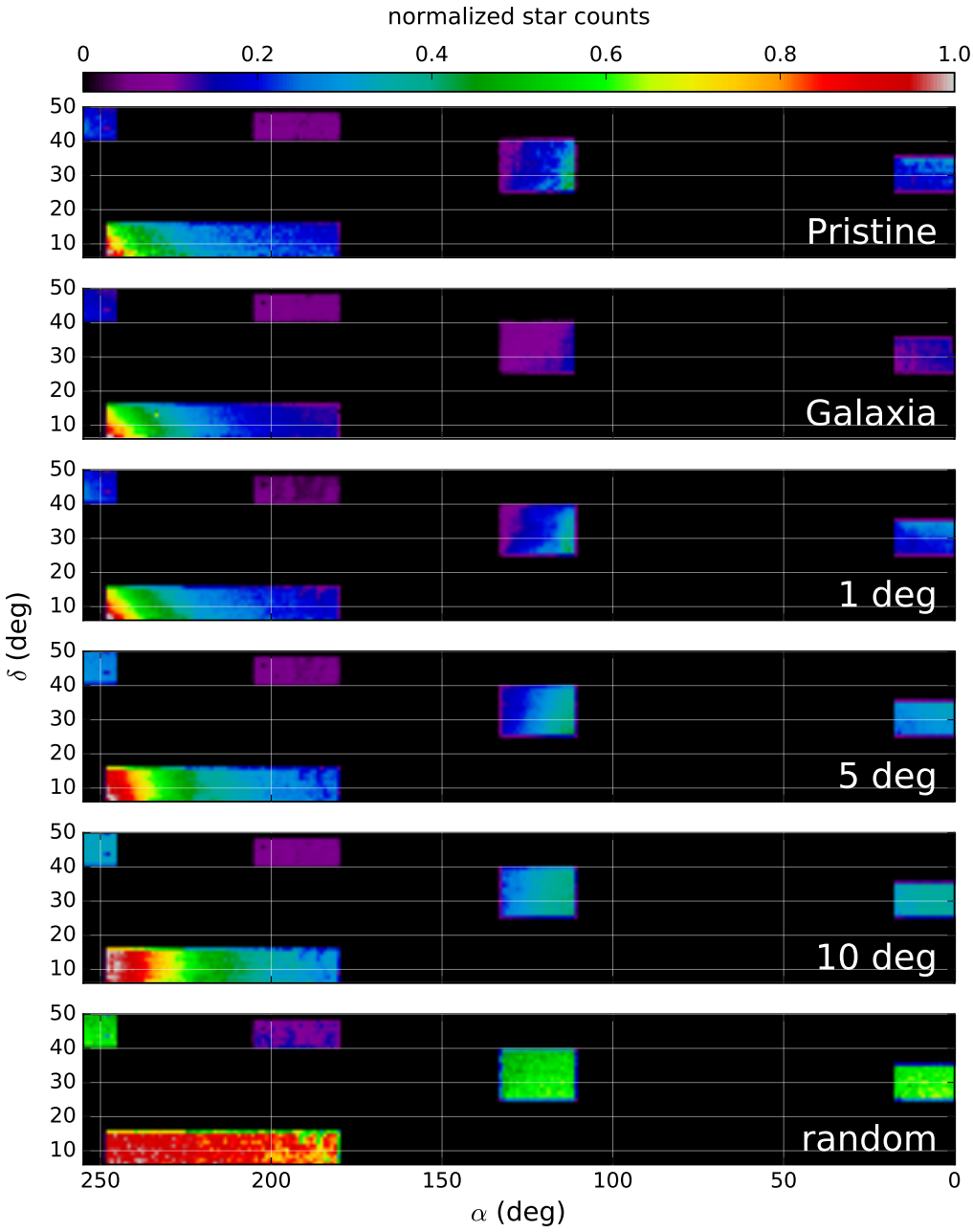

Figure 6.7 — Stellar density plots of the *Pristine* footprint (top) and the smooth background samples.

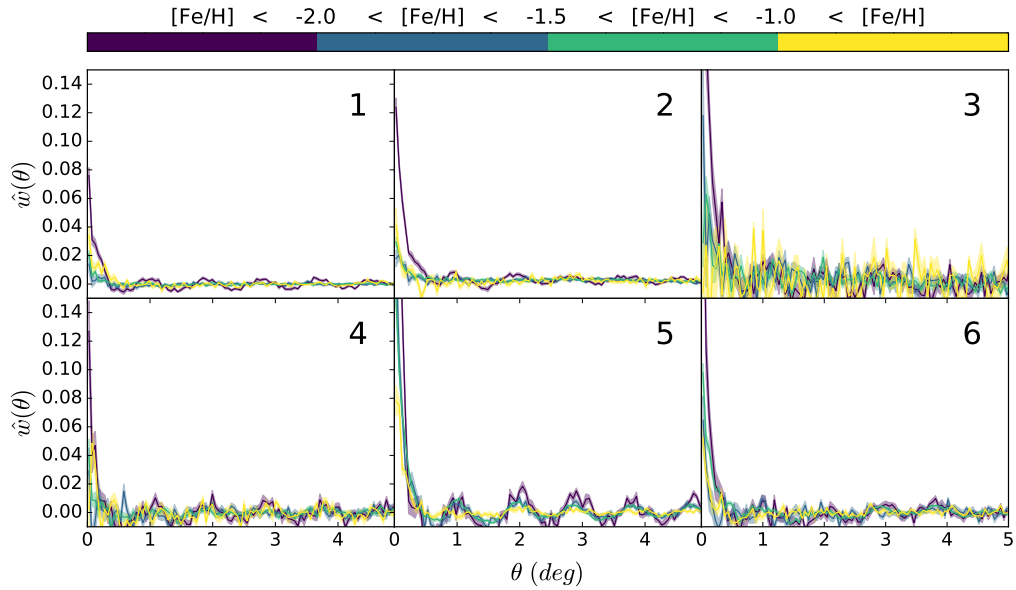

Figure 6.8 — 2-PCF of the 6 *Pristine* footprint regions using the 1 degree smoothed gaussian random sample.

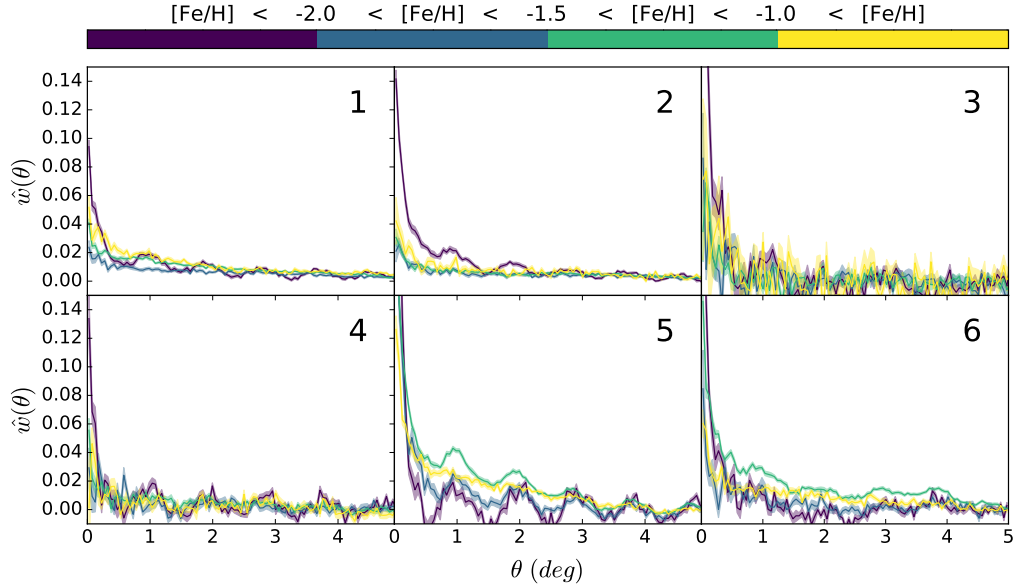

Figure 6.9 — 2-PCF of the 6 *Pristine* footprint regions using the 5 degree smoothed gaussian random sample.

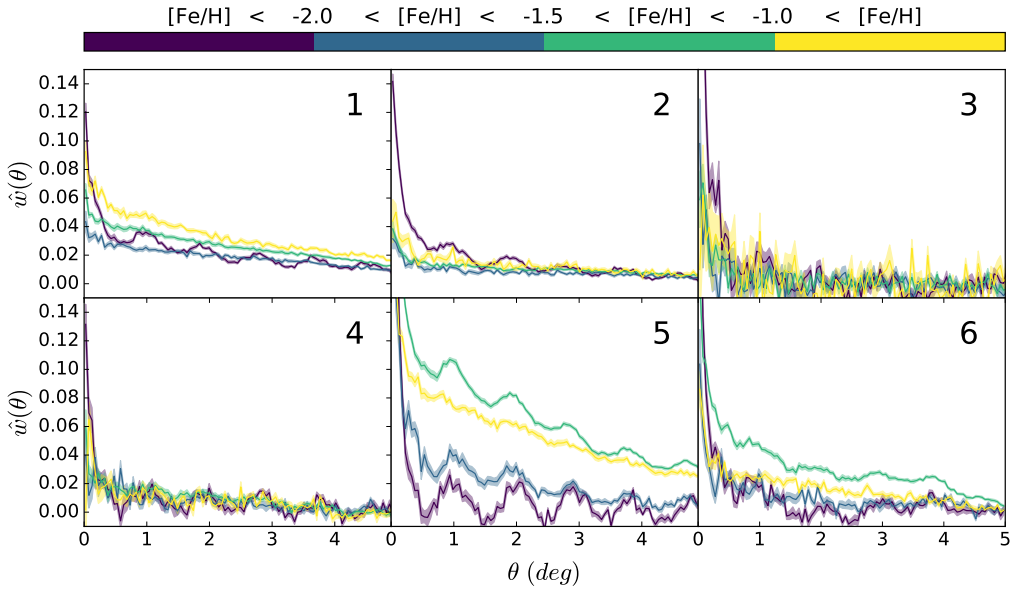

Figure 6.10 — 2-PCF of the 6 *Pristine* footprint regions using the 10 degree smoothed gaussian random sample.

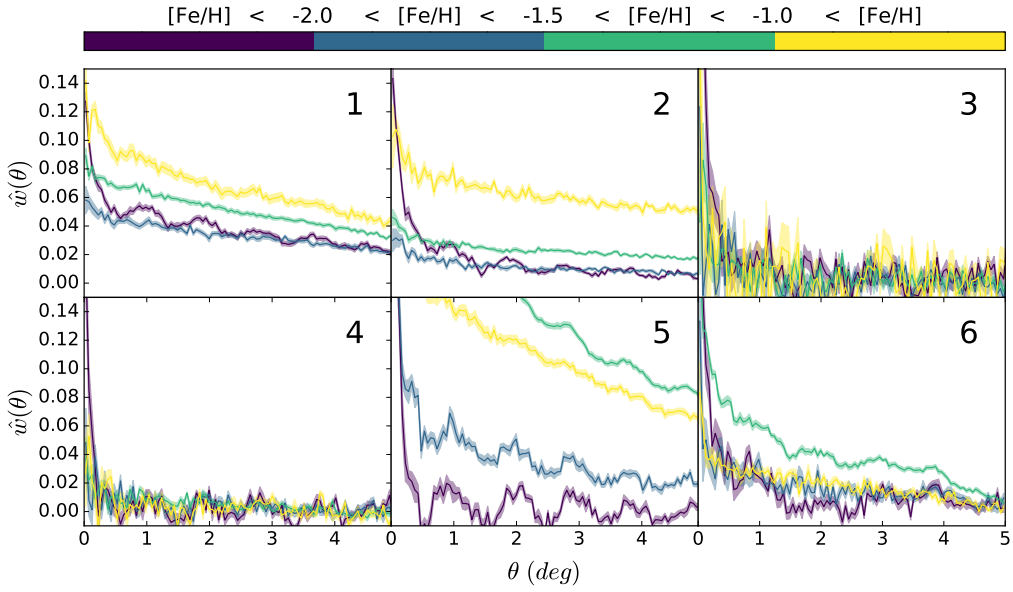

Figure 6.11 — 2-PCF of the 6 *Pristine* footprint regions using the fully randomly generated sample.

7 Conclusions & Future Outlook

7.1 Conclusions

The goal of this thesis is to contribute to the understanding of the properties of the Milky Way halo, specifically its chemical composition and substructure content. We have done this in a variety of ways, including using the *Pristine* survey as an efficient method for finding metal-poor stars and selecting blue horizontal branch (BHB) stars, providing a photometric metallicity distribution function (MDF) of the halo, and exploring statistical methods for future large scale clustering analyses. These results have not only had an immediate impact on the field of Galactic archaeology, but also provide a basis to be expanded upon in future work. In particular, some of these results have already been cited and used by other studies, and will be useful for upcoming large-scale surveys designed to target metal-poor stars for Galactic archaeology studies.

7.2 Preliminary impact of this thesis work

In Chapters 2 and 3, we characterized the success rates of *Pristine* follow-up observations, paving the way for the future discovery of large numbers of metal-poor stars by follow-up of *Pristine* candidates with upcoming multi-object spectroscopic (MOS) surveys. In Chapter 3, we have also presented the full sample of > 1 000 spectra of *Pristine* medium-resolution folllow-up stars. The catalogue contains > 700 new very metal-poor (VMP) stars, and almost 100 new extremely metal-poor (EMP) stars. It has provided a significant contribution to the known sample of metal-poor stars in the literature, and has already been used for Galactic archaeology studies to investigate the orbital dynamics of EMP stars. More specifically, Sestito et al. (2019b) showed that of the sample currently located in the disc, 31% are confined to disc like orbits, suggesting that a significant fraction of stars with [Fe/H] ≤ -2.5 formed either within, or concurrently with the Milky Way disk. During the course of this follow-up campaign we have also discovered one of the most metal-poor stars ever found, Pristine_221.8781+9.7844 (see Starkenburg et al. 2018 for a detailed analysis), which along with other discoveries of stars with similarly unusual abundance patterns has helped to constrain theories of early chemical enrichment by the first stellar generations.

We have also used the results from this work to estimate the number of expected EMP and UMP stars in the halo, namely that for a magnitude range of $14 < V < 18$ and $|b| > 30°$, we expect $\sim 1/800$ stars to be EMP, and $\sim 1/80\,000$ to be UMP. To make these calculations we have

used a UMP:EMP ratio of 1:100, which has been previously described in the literature (although with low number statistics, Schörck et al. 2009; Allende Prieto et al. 2014) and confirmed by our sample. Although it may seem trivial, a robust value of the frequency of these most metal-poor stars in the Milky Way halo is difficult to compute, due to their scarcity and the small samples that we currently have, and also due to the complex selection functions of the spectroscopic surveys with which they are found. This is mostly mitigated by the photometric *Pristine* survey, and we can therefore make relatively unbiased estimates of the frequencies for these metal-poor stars. A recent study by Magg et al. (2019) have used this result to derive a simple analytical estimate to constrain the number of surviving pristine (metal-free) stars in the Milky Way. They conclude with high confidence, that if metal-free stars are still present in the Galaxy, we should have observed at least one by now. They then suggest that this implies that either metal-free stars no longer exist today, or that metal-free stars must have formed very rarely in the Galaxy or have been polluted during their lifetimes.

In Chapter 4, we extended this concept further and provided a corrected MDF of a selected halo sample of ~ 80 000 *Pristine* stars that is largely unbiased in the metallicity range of $-3.5 < [Fe/H] < -1.5$. This sample has the advantage of a larger number of stars and less selection effects as compared to previous works that have presented MDFs of the halo using smaller spectroscopic samples. The results showed a substantially higher frequency of metal-poor stars with $-3 < [Fe/H] < -2$ as compared to the spectroscopic samples of Schörck et al. (2009) and Allende Prieto et al. (2014), but a similar number as that predicted by An et al. (2013, 2015) who used photometric metallicities derived from an SDSS sample. These results have interesting implications for our understanding of our Galaxy. For example, if this higher frequency of metal-poor stars is real, this may suggest that the paucity of metal-poor globular clusters with $[Fe/H] < -2.5$ in our Galaxy is due to a physical truncation of the MDF for globular clusters, rather than simply statistical undersampling. The MDF presented in Chapter 4 of this thesis is being applied to a more thorough investigation of this question in Gieles et al. in prep.

7.3 Future Outlook

7.3.1 The future of the *Pristine* survey: synergies with WEAVE and CFIS

In this thesis we have demonstrated some of the ways in which the *Pristine* dataset can be used for Galactic archaeology studies. The data used throughout this thesis was based on the first internal data release of the *Pristine* survey, with a coverage of ~ 2 500 deg^2, mostly in the northern Galactic halo. Currently, the survey covers a total footprint of ~ 5 000 deg^2 and continues to increase in size, with a homogeneous data reduction of the entire footprint currently underway. With this increased coverage, we will be able to expand upon the work presented in this thesis. For example, it will enable us to investigate the MDF and BHB stellar distributions for significantly larger regions of the northern hemisphere sky. Although this work still requires further development, the preliminary results shown in Chapter 6 present potential promise for using the two-point correlation function and *Pristine* data identifying and quantifying diffuse substructure in different Galactic environments. The doubling of area of the new *Pristine* survey footprint will provide additional statistical power to those regions that were too small to yield significant results in this preliminary work.

There is also a lot of potential to further expand upon the work of Chapter 5 with the enlarged footprint of the *Pristine* survey, particularly because there will be a large region of overlap with CFIS (the Canada-France Imaging Survey Ibata et al. 2017a). In this chapter we showed that the *Pristine* filter used in combination with SDSS broad-band photometry can separate BHB stars from BS stars with higher completeness and purity than previously possible with SDSS photometry alone (Deason et al. 2014). A recent paper by Thomas et al. (2018b)

has also shown that the deep u-band photometry from CFIS can also be used to make clean selections of BHB stars. The selection of the BHB sample using *Pristine* data is currently precision limited by the depth of the SDSS u-band data, but the u-band data from CFIS is ∼ 3 magnitudes deeper than SDSS, and by the projected end of CFIS this high quality data will cover the entire *Pristine* footprint. Combining these two data sets from *Pristine* and CFIS will allow for the selection of substantially fainter samples of BHB stars to probe the Galactic halo out to even larger radii.

Naturally, the doubling in size of the footprint for the *Pristine* survey will also provide a larger number of metal-poor candidates over more Galactic environments to be followed up by upcoming MOS surveys such as WEAVE. The WEAVE survey is scheduled to begin in 2020 on the William Herschel Telescope on La Palma in the Canary Islands. Located in the northern hemisphere, the main WEAVE Galactic archaeology surveys will target the outer disc and the halo. One of the primary objectives of WEAVE is to complement the data set from the Gaia space mission, and as such it will provide radial velocities for stars with $V \gtrsim 15.5$, and detailed elemental abundances for stars with $V \gtrsim 12$, in the regimes that are too faint for accurate measurements with Gaia spectra. With both low resolution (R ∼ 5 000) and high resolution (R ∼ 20 000) modes, it will provide both large spatial coverage with many observed spectra and also detailed abundances for a subset of stars. One of WEAVE's primary science goals is disentangling the assembly history of the Milky Way disc and halo. In particular the WEAVE collaboration has already signed a memorandum of understanding with the *Pristine* collaboration, and will use *Pristine* photometric candidates to search for metal-poor stars in their Galactic archaeology program. The details of this partnership have been described in the discussion of Chapter 3, but to summarize, WEAVE plans to dedicate an average of ∼ 6 fibres per deg^2 and will observe up to ∼ 30 000 *Pristine* candidate EMP stars over the projected ≥ 5 000 deg^2 overlap between the two surveys. Extrapolating from the success rates for finding EMP stars in the *Pristine* survey calculated in Chapter 3, we project that this will lead to the discovery of ∼ 5 000 − 7 000 EMP stars, and ∼ 50 − 70 UMP stars, effectively doubling the current samples of these most metal-poor stars.

7.3.2 The Milky Way in a cosmological framework

Simulations continue to improve in both resolution and the detailed physics that they can model. It is now possible using high resolution cosmological hydrodynamic zoom-in simulations and techniques to decompose massive star particles into realistic stellar populations, and to produce realistic Gaia catalogues of Milky Way mass galaxies with parameters for individual stars (e.g. Lowing et al. 2015; Grand et al. 2018; Sanderson et al. 2018). Furthermore, realistic analogues of massive accretion events have been identified in simulations, such as the Gaia sausage (Fattahi et al. 2019) and Gaia Enceledus (Bignone et al. 2019), hinting at the potential for simulations to be used to disentangle the accretion history of the Milky Way. Making comparisons with simulations allows for the investigation of the evolution of such merger events, such as the time of the merger, the size of the satellite, and the mixing time. As simulations continue to improve, we can hope not only to find more analogues of substructures in the Milky Way to explain their origins, but also to make more global comparisons of the substructure content in simulated halos as compared to that of the Milky Way, using statistical methods such as the two-point correlation function (Chapter 6 Lancaster et al. 2019a), or large scale hierarchical algorithms such as OPTICS (McConnachie et al. 2018, Thomas et al. in prep.).

7.3.3 Gaia and spectroscopic surveys in the next decade

The ESA Gaia satellite mission has already provided the largest catalogue of photometric and astrometric observations covering the full sky to date. With the second data release (DR2) in 2018, the publicly available catalogue consisted of more than 1.3 billion sources with parallaxes and proper motions in the magnitude range of $3 < G < 21$, and radial velocities to more than 7.2 million stars in the magnitude range of $4 < G < 13$. The third data release, scheduled for the second half of 2021, will provide astrometry and integrated photometry to roughly 1.8 billion sources, source classification and astrophysical parameters from BP/RP spectra for more than 300 million sources (including $T_{\rm eff}$, $\log g$, metallicity, extinction, distances, amongst others), and radial velocities to roughly 30 million sources with $G_{RVS} < 14$. The fourth data release is projected to improve the precision of parallaxes, photometry, and radial velocities by a factor of 1.7 with respect to DR2, and proper motions by a factor of 4.5. Should the mission be extended for an additional 5 years past the nominal mission time, these would increase further by a factors of 1.4 and 2.8, respectively. As this very powerful data set continues to improve, it will allow for the continued discovery of substructures and allow for these to be associated to clusters and stars for a more complete picture of the accretion history of the Galaxy. In addition, supplementing this data set with the large number of chemical abundances over large regions of the halo and disc that will be provided by WEAVE and 4MOST will add even more dimensions with which to dissect these Galactic components.

One upcoming survey, the Hobby-Eberly Telescope Dark Energy Experiment (HETDEX Hill et al. 2008) is particularly interesting for the characterisation of the halo MDF, and expanding upon the work of Chapter 4. With the Visible Integral-Field Replicable Unit Spectrographs (VIRUS), they plan to cover 420 deg^2 in the northern Galactic halo with an array of 150 integral field unit spectrographs. Although the primary objective of the survey is to measure the redshifts of Galaxies, each pointing will produce ~ 34 000 spectra, and over the course of the survey ~ 250 000 stars in the Milky Way will be serendipitously observed. This sample will be a very interesting data set with which to compute the MDF of the halo, as it combines the precision of spectroscopy with the unbiased observing strategy of a photometric survey, and will provide large samples with which to dissect different stellar populations and distance ranges.

Large scale spectroscopic surveys such as SDSS-I -IV, RAVE, GALAH and Gaia ESO, amongst others, have allowed us to make large progress in understanding the stellar populations in our Galaxy with large numbers of uniformly observed spectra. Combined, these surveys have provided on the order of ~ 10^6 spectra with high quality abundance and radial velocity information. The next generation of MOS facilities are going to increase the number of stars observed spectroscopically yet another order of magnitude, with surveys such as LAMOST, DESI, WEAVE, 4MOST, MSE, PSF and SDSS-V poised to provide spectral information for, on the order of ~ 10^7 stars. In combination with Gaia, which will supplement these observations with high quality photometry, as well as providing spectroscopic stellar parameters for on the order of ~ 10^8 objects on its own, these samples will provide an unprecedented view of our Galaxy in phantasmogoric detail. The next decade will therefore not only greatly expand the number of stars observed, but also probe more Galactic environments and much fainter magnitude ranges than previously possible. This will provide further insight into studies of Galctic substructure, the accretion history of the Milky Way, chemical and kinematic tagging, Galactic populations and abundance trends, and much more. The implications for advancing the field of Galactic archaeology are enormous.

Bibliography

Abadi, M. G., Navarro, J. F., & Steinmetz, M. 2006, MNRAS, 365, 747
Abate, C., Pols, O. R., Karakas, A. I., & Izzard, R. G. 2015, A&A, 576, A118
Abbott, T. M. C., Abdalla, F. B., Allam, S., et al. 2018, ApJS, 239, 18
Abolfathi, B., Aguado, D. S., Aguilar, G., et al. 2018, ApJS, 235, 42
Abuter, R., Amorim, A., Bauboeck, M., et al. 2019, arXiv e-prints, arXiv:1904.05721
Aguado, D. S., Allende Prieto, C., González Hernández, J. I., et al. 2016, A&A, 593, A10
Aguado, D. S., Allende Prieto, C., González Hernández, J. I., & Rebolo, R. 2018a, ApJ, 854, L34
Aguado, D. S., Allende Prieto, C., González Hernández, J. I., Rebolo, R., & Caffau, E. 2017a, A&A, 604, A9
Aguado, D. S., González Hernández, J. I., Allende Prieto, C., & Rebolo, R. 2017b, A&A, 605, A40
Aguado, D. S., González Hernández, J. I., Allende Prieto, C., & Rebolo, R. 2018b, ApJ, 852, L20
Aguado, D. S., Youakim, K., González Hernández, J. I., et al. 2019, MNRAS, 490, 2241
Allende Prieto, C., Beers, T. C., Wilhelm, R., et al. 2006, ApJ, 636, 804
Allende Prieto, C., Fernández-Alvar, E., Aguado, D. S., et al. 2015a, A&A, 579, A98
Allende Prieto, C., Fernández-Alvar, E., Schlesinger, K. J., et al. 2014, A&A, 568, A7
Allende Prieto, C., Koesterke, L., Hubeny, I., et al. 2018, A&A, 618, A25
Allende Prieto, C., Rebolo, R., García López, R. J., et al. 2000, AJ, 120, 1516
Allende Prieto, C. et al. 2015b, in American Astronomical Society Meeting Abstracts, Vol. 225, American Astronomical Society Meeting Abstracts, 422.07
Alpher, R. A., Bethe, H., & Gamow, G. 1948, Physical Review, 73, 803
Alvarez, R. & Plez, B. 1998, A&A, 330, 1109
An, D., Beers, T. C., Johnson, J. A., et al. 2013, ApJ, 763, 65
An, D., Beers, T. C., Santucci, R. M., et al. 2015, ApJ, 813, L28
Anthony-Twarog, B. J., Sarajedini, A., Twarog, B. A., & Beers, T. C. 2000, AJ, 119, 2882
Anthony-Twarog, B. J., Twarog, B. A., Laird, J. B., & Payne, D. 1991, AJ, 101, 1902
Aoki, W., Beers, T. C., Christlieb, N., et al. 2007, ApJ, 655, 492
Aoki, W., Beers, T. C., Lee, Y. S., et al. 2013, AJ, 145, 13
Arentsen, A., Starkenburg, E., Martin, N. F., et al. 2020, MNRAS, 491, L11
Arentsen, A., Starkenburg, E., Shetrone, M. D., et al. 2019, A&A, 621, A108
Arnold, R. & Gilmore, G. 1992, MNRAS, 257, 225
Asplund, M. 2005, ARA&A, 43, 481
Asplund, M., Grevesse, N., & Sauval, A. J. 2005, in Astronomical Society of the Pacific Conference Series, Vol. 336, Cosmic Abundances as Records of Stellar Evolution and Nucleosynthesis, ed. T. G. Barnes, III & F. N. Bash, 25
Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481
Auer, L. H. & Mihalas, D. 1972, ApJS, 24, 193
Barklem, P. S., Piskunov, N., & O'Mara, B. J. 2000, A&A, 363, 1091
Beers, T. C., Carollo, D., Ivezić, Ž., et al. 2012, ApJ, 746, 34
Beers, T. C. & Christlieb, N. 2005, ARA&A, 43, 531
Beers, T. C., Christlieb, N., Norris, J. E., et al. 2005, in IAU Symposium, Vol. 228, From Lithium to Uranium: Elemental Tracers of Early Cosmic Evolution, ed. V. Hill, P. Francois, & F. Primas, 175–183
Beers, T. C., Preston, G. W., & Shectman, S. A. 1985, AJ, 90, 2089
Beers, T. C., Preston, G. W., & Shectman, S. A. 1992, AJ, 103, 1987
Beers, T. C., Rossi, S., Norris, J. E., Ryan, S. G., & Shefler, T. 1999, AJ, 117, 981
Bell, E. F., Xue, X. X., Rix, H.-W., Ruhland, C., & Hogg, D. W. 2010, AJ, 140, 1850
Bell, E. F., Zucker, D. B., Belokurov, V., et al. 2008, ApJ, 680, 295
Belokurov, V. 2013, New Astron. Rev., 57, 100
Belokurov, V., Erkal, D., Evans, N. W., Koposov, S. E., & Deason, A. J. 2018, MNRAS, 478, 611

Belokurov, V., Koposov, S. E., Evans, N. W., et al. 2014, MNRAS, 437, 116
Belokurov, V., Zucker, D. B., Evans, N. W., et al. 2006, ApJ, 642, L137
Belokurov, V., Zucker, D. B., Evans, N. W., et al. 2007, ApJ, 654, 897
Bergemann, M., Sesar, B., Cohen, J. G., et al. 2018, Nature, 555, 334
Bernard, E. J., Ferguson, A. M. N., Schlafly, E. F., et al. 2016, MNRAS, 463, 1759
Bessell, M., Bloxham, G., Schmidt, B., et al. 2011, PASP, 123, 789
Bessell, M. S., Collet, R., Keller, S. C., et al. 2015, ApJ, 806, L16
Bignone, L. A., Helmi, A., & Tissera, P. B. 2019, ApJ, 883, L5
Binney, J. & Tremaine, S. 2008, Galactic Dynamics: Second Edition
Bland-Hawthorn, J. & Gerhard, O. 2016, ARA&A, 54, 529
Bonifacio, P., Caffau, E., Sestito, F., et al. 2019, MNRAS, 1324
Bonifacio, P., Caffau, E., Spite, M., et al. 2015, A&A, 579, A28
Bonifacio, P., Caffau, E., Spite, M., et al. 2018a, Research Notes of the American Astronomical Society, 2, 19
Bonifacio, P., Caffau, E., Spite, M., et al. 2018b, A&A, 612, A65
Bonifacio, P., Sbordone, L., Caffau, E., et al. 2012, A&A, 542, A87
Bonifacio, P., Spite, M., Cayrel, R., et al. 2009, A&A, 501, 519
Bressan, A., Marigo, P., Girardi, L., et al. 2012, MNRAS, 427, 127
Bromm, V. & Loeb, A. 2003, Nature, 425, 812
Buder, S., Asplund, M., Duong, L., et al. 2018, MNRAS, 478, 4513
Bullock, J. S. & Johnston, K. V. 2005, ApJ, 635, 931
Burbidge, E. M., Burbidge, G. R., Fowler, W. A., & Hoyle, F. 1957, Reviews of Modern Physics, 29, 547
Buzzoni, B., Delabre, B., Dekker, H., et al. 1984, The Messenger, 38, 9
Caffau, E., Bonifacio, P., François, P., et al. 2011a, Nature, 477, 67
Caffau, E., Bonifacio, P., François, P., et al. 2012, A&A, 542, A51
Caffau, E., Bonifacio, P., Sbordone, L., et al. 2013a, A&A, 560, A71
Caffau, E., Bonifacio, P., Starkenburg, E., et al. 2017, ArXiv e-prints
Caffau, E., Ludwig, H.-G., Steffen, M., Freytag, B., & Bonifacio, P. 2011b, Solar Physics, 268, 255
Carney, B. W., Laird, J. B., Latham, D. W., & Aguilar, L. A. 1996, AJ, 112, 668
Carollo, D., Beers, T. C., Bovy, J., et al. 2012, ApJ, 744, 195
Carollo, D., Beers, T. C., Chiba, M., et al. 2010, ApJ, 712, 692
Carollo, D., Beers, T. C., Lee, Y. S., et al. 2007, Nature, 450, 1020
Casagrande, L., Wolf, C., Mackey, A. D., et al. 2019, MNRAS, 482, 2770
Casey, A. R. & Schlaufman, K. C. 2015, ApJ, 809, 110
Castelli, F. & Kurucz, R. L. 2003, in IAU Symposium, Vol. 210, Modelling of Stellar Atmospheres, ed. N. Piskunov, W. W. Weiss, & D. F. Gray, A20
Cayrel, R., Depagne, E., Spite, M., et al. 2004, A&A, 416, 1117
Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, ArXiv e-prints
Chiaki, G., Tominaga, N., & Nozawa, T. 2017, MNRAS, 472, L115
Chiaki, G. & Wise, J. H. 2019, MNRAS, 482, 3933
Chiappini, C. 2013, Astronomische Nachrichten, 334, 595
Chiba, M. & Beers, T. C. 2000, AJ, 119, 2843
Christlieb, N., Gustafsson, B., Korn, A. J., et al. 2004, ApJ, 603, 708
Christlieb, N., Wisotzki, L., & Graßhoff, G. 2002, A&A, 391, 397
Clewley, L., Warren, S. J., Hewett, P. C., et al. 2002, MNRAS, 337, 87
Cohen, J. G., Sesar, B., Bahnolzer, S., et al. 2017, ApJ, 849, 150
Da Costa, G. S., Bessell, M. S., Mackey, A. D., et al. 2019, MNRAS, 2195
Dalton, G., Trager, S., Abrams, D. C., et al. 2018, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10702, 107021B
Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, AJ, 145, 10
de Boer, T. J. L., Belokurov, V., & Koposov, S. E. 2018, MNRAS, 473, 647
de Jong, R. S., Agertz, O., Berbel, A. A., et al. 2019, The Messenger, 175, 3
De Propris, R., Harrison, C. D., & Mares, P. J. 2010, ApJ, 719, 1582
Deason, A. J., Belokurov, V., & Evans, N. W. 2011, MNRAS, 416, 2903

Deason, A. J., Belokurov, V., Evans, N. W., et al. 2012, MNRAS, 425, 2840
Deason, A. J., Belokurov, V., & Koposov, S. E. 2018a, MNRAS, 473, 2428
Deason, A. J., Belokurov, V., & Koposov, S. E. 2018b, ApJ, 852, 118
Deason, A. J., Belokurov, V., Koposov, S. E., & Lancaster, L. 2018c, ApJ, 862, L1
Deason, A. J., Belokurov, V., Koposov, S. E., & Rockosi, C. M. 2014, ApJ, 787, 30
Deason, A. J., Belokurov, V., & Sanders, J. L. 2019, MNRAS, 2394
Deason, A. J., Mao, Y.-Y., & Wechsler, R. H. 2016, ApJ, 821, 5
Deng, L.-C., Newberg, H. J., Liu, C., et al. 2012, Research in Astronomy and Astrophysics, 12, 735
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
D'Souza, R. & Bell, E. F. 2018a, Nature Astronomy, 2, 737
D'Souza, R. & Bell, E. F. 2018b, MNRAS, 474, 5300
Efron, B., Hastie, T., Johnstone, I., & Tibshirani, R. 2004, The Annals of statistics, 32, 407
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
Fardal, M. A., van der Marel, R. P., Law, D. R., et al. 2019, MNRAS, 483, 4724
Fattahi, A., Belokurov, V., Deason, A. J., et al. 2019, MNRAS, 484, 4471
Fernández-Alvar, E., Allende Prieto, C., Beers, T. C., et al. 2016, A&A, 593, A28
Fernández-Alvar, E., Allende Prieto, C., Schlesinger, K. J., et al. 2015, A&A, 577, A81
Finkbeiner, D. P., Schlafly, E. F., Schlegel, D. J., et al. 2016, ApJ, 822, 66
François, P., Caffau, E., Bonifacio, P., et al. 2018, A&A, 620, A187
Frebel, A., Aoki, W., Christlieb, N., et al. 2005, Nature, 434, 871
Frebel, A., Christlieb, N., Norris, J. E., et al. 2006, ApJ, 652, 1585
Frebel, A., Ji, A. P., Ezzeddine, R., et al. 2019, ApJ, 871, 146
Frebel, A. & Norris, J. E. 2015, ARA&A, 53, 631
Freeman, K. & Bland-Hawthorn, J. 2002, ARA&A, 40, 487
Fukushima, T., Chiba, M., Homma, D., et al. 2018, PASJ, 70, 69
Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1
Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1
Gallart, C., Bernard, E. J., Brook, C. B., et al. 2019, Nature Astronomy, 407
Geier, S., Raddi, R., Gentile Fusillo, N. P., & Marsh, T. R. 2019, A&A, 621, A38
Gentile Fusillo, N. P., Gänsicke, B. T., & Greiss, S. 2015, MNRAS, 448, 2260
Gómez, F. A., Besla, G., Carpintero, D. D., et al. 2015, ApJ, 802, 128
Grand, R. J. J., Helly, J., Fattahi, A., et al. 2018, MNRAS, 481, 1726
Gratton, R. G., Carretta, E., Claudi, R., Lucatello, S., & Barbieri, M. 2003, A&A, 404, 187
Grillmair, C. J. 2006, ApJ, 651, L29
Grillmair, C. J. 2009, ApJ, 693, 1118
Grillmair, C. J. & Carlin, J. L. 2016, in Astrophysics and Space Science Library, Vol. 420, Tidal Streams in the Local Group and Beyond, ed. H. J. Newberg & J. L. Carlin, 87
Grillmair, C. J., Carlin, J. L., & Majewski, S. R. 2008, ApJ, 689, L117
Gu, J., Du, C., Jia, Y., et al. 2015, MNRAS, 452, 3092
Gustafsson, B., Edvardsson, B., Eriksson, K., et al. 2008, A&A, 486, 951
Hansen, T., Hansen, C. J., Christlieb, N., et al. 2014, ApJ, 787, 162
Hansen, T. T., Andersen, J., Nordström, B., et al. 2016a, A&A, 586, A160
Hansen, T. T., Andersen, J., Nordström, B., et al. 2016b, A&A, 588, A3
Harmsen, B., Monachesi, A., Bell, E. F., et al. 2017, MNRAS, 466, 1491
Harris, W. E. 2010, ArXiv e-prints
Hartwick, F. D. A. 1976, ApJ, 209, 418
Haywood, M., Di Matteo, P., Lehnert, M. D., et al. 2018, ApJ, 863, 113
Helmi, A. 2004, ApJ, 610, L97
Helmi, A. 2008, A&ARv, 15, 145
Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, Nature, 563, 85
Helmi, A. & White, S. D. M. 1999, MNRAS, 307, 495
Helmi, A., White, S. D. M., de Zeeuw, P. T., & Zhao, H. 1999, Nature, 402, 53

Henden, A. & Munari, U. 2014, Contributions of the Astronomical Observatory Skalnate Pleso, 43, 518
Henden, A. A., Levine, S., Terrell, D., & Welch, D. L. 2015, in American Astronomical Society Meeting Abstracts, Vol. 225, American Astronomical Society Meeting Abstracts, 336.16
Henden, A. A., Welch, D. L., Terrell, D., & Levine, S. E. 2009, in American Astronomical Society Meeting Abstracts, Vol. 214, American Astronomical Society Meeting Abstracts #214, 669
Hernitschek, N., Cohen, J. G., Rix, H.-W., et al. 2018, ApJ, 859, 31
Hernitschek, N., Schlafly, E. F., Sesar, B., et al. 2016, ApJ, 817, 73
Hernitschek, N., Sesar, B., Rix, H.-W., et al. 2017, ApJ, 850, 96
Hill, G. J., Gebhardt, K., Komatsu, E., et al. 2008, The Hobby-Eberly Telescope Dark Energy Experiment (HETDEX): Description and Early Pilot Survey Results
Howes, L. M., Asplund, M., Keller, S. C., et al. 2016, MNRAS, 460, 884
Howes, L. M., Casey, A. R., Asplund, M., et al. 2015, Nature, 527, 484
Huang, Y., Chen, B.-Q., Yuan, H.-B., et al. 2019, arXiv e-prints
Huxor, A. P. & Grebel, E. K. 2015, MNRAS, 453, 2653
Ibata, R., Lewis, G. F., Irwin, M., Totten, E., & Quinn, T. 2001, ApJ, 551, 294
Ibata, R. A., Gilmore, G., & Irwin, M. J. 1994, Nature, 370, 194
Ibata, R. A., Lewis, G. F., McConnachie, A. W., et al. 2014, ApJ, 780, 128
Ibata, R. A., Malhan, K., & Martin, N. F. 2019, ApJ, 872, 152
Ibata, R. A., McConnachie, A., Cuillandre, J.-C., et al. 2017a, ApJ, 848, 128
Ibata, R. A., McConnachie, A., Cuillandre, J.-C., et al. 2017b, ApJ, 848, 129
Irwin, M. & Lewis, J. 2001, New Astron. Rev., 45, 105
Ishigaki, M. N., Chiba, M., & Aoki, W. 2012, ApJ, 753, 64
Ishigaki, M. N., Tominaga, N., Kobayashi, C., & Nomoto, K. 2014, ApJ, 792, L32
Ivezić, Ž., Sesar, B., Jurić, M., et al. 2008, ApJ, 684, 287
Iwamoto, N., Umeda, H., Tominaga, N., Nomoto, K., & Maeda, K. 2005, Science, 309, 451
Janesh, W., Morrison, H. L., Ma, Z., et al. 2016, ApJ, 816, 80
Johnston, K. V., Law, D. R., & Majewski, S. R. 2005, ApJ, 619, 800
Jorden, P. R. 1990, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 1235, Instrumentation in Astronomy VII, ed. D. L. Crawford, 790–798
Keller, S. C., Bessell, M. S., Frebel, A., et al. 2014, Nature, 506, 463
Keller, S. C., Schmidt, B. P., Bessell, M. S., et al. 2007, Pub. of the Astr. Soc. of Australia, 24, 1
Kinman, T. D., Suntzeff, N. B., & Kraft, R. P. 1994, AJ, 108, 1722
Koch, A., McWilliam, A., Preston, G. W., & Thompson, I. B. 2016, A&A, 587, A124
Koesterke, L., Allende Prieto, C., & Lambert, D. L. 2008, ApJ, 680, 764
Kollmeier, J. A., Zasowski, G., Rix, H.-W., et al. 2017, arXiv e-prints, arXiv:1711.03234
Koposov, S. E., Belokurov, V., & Evans, N. W. 2013, ApJ, 766, 79
Koposov, S. E., Belokurov, V., Zucker, D. B., et al. 2015, MNRAS, 446, 3110
Koppelman, H., Helmi, A., & Veljanoski, J. 2018, ApJ, 860, L11
Koppelman, H. H., Helmi, A., Massari, D., Price-Whelan, A. M., & Starkenburg, T. K. 2019, arXiv e-prints, arXiv:1909.08924
Kos, J., Bland-Hawthorn, J., Freeman, K., et al. 2018, MNRAS, 473, 4612
Lamb, M., Venn, K., Andersen, D., et al. 2017, MNRAS, 465, 3536
Lancaster, L., Belokurov, V., & Evans, N. W. 2019a, MNRAS, 484, 2556
Lancaster, L., Koposov, S. E., Belokurov, V., Evans, N. W., & Deason, A. J. 2019b, MNRAS, 486, 378
Landy, S. D. & Szalay, A. S. 1993, ApJ, 412, 64
Laporte, C. F. P., Johnston, K. V., Gómez, F. A., Garavito-Camargo, N., & Besla, G. 2018, MNRAS, 481, 286
Laporte, C. F. P., Johnston, K. V., & Tzanidakis, A. 2019, MNRAS, 483, 1427
Law, D. R., Johnston, K. V., & Majewski, S. R. 2005, ApJ, 619, 807
Law, D. R. & Majewski, S. R. 2010, ApJ, 714, 229
Law, D. R. & Majewski, S. R. 2016, in Astrophysics and Space Science Library, Vol. 420, Tidal Streams in the Local Group and Beyond, ed. H. J. Newberg & J. L. Carlin, 31
Lenz, D. D., Newberg, J., Rosner, R., Richards, G. T., & Stoughton, C. 1998, ApJS, 119, 121
Levi, M., Bebek, C., Beers, T., et al. 2013, arXiv e-prints, arXiv:1308.0847
Li, H., Aoki, W., Matsuno, T., et al. 2018, ApJ, 852, L31

Li, H., Tan, K., & Zhao, G. 2018, The Astrophysical Journal Supplement Series, 238, 16
Li, H.-N., Aoki, W., Honda, S., et al. 2015, Research in Astronomy and Astrophysics, 15, 1264
Li, J., FELLOW, L., Liu, C., et al. 2019, ApJ, 874, 138
Li, J., Newberg, H. J., Carlin, J. L., et al. 2012, ApJ, 757, 151
Li, T. S., Sheffield, A. A., Johnston, K. V., et al. 2017, ApJ, 844, 74
Liu, C., Deng, L.-C., Carlin, J. L., et al. 2014, ApJ, 790, 110
Lokhorst, D., Starkenburg, E., McConnachie, A. W., et al. 2016, ApJ, 819, 124
Longeard, N., Martin, N., Starkenburg, E., et al. 2018, MNRAS, 480, 2609
Longeard, N., Martin, N., Starkenburg, E., et al. 2019, arXiv e-prints, arXiv:1902.02780
Lowing, B., Wang, W., Cooper, A., et al. 2015, MNRAS, 446, 2274
Magg, M., Klessen, R. S., Glover, S. C. O., & Li, H. 2019, MNRAS, 487, 486
Majewski, S. R., Ostheimer, J. C., Rocha-Pinto, H. J., et al. 2004, ApJ, 615, 738
Majewski, S. R., Skrutskie, M. F., Weinberg, M. D., & Ostheimer, J. C. 2003, ApJ, 599, 1082
Malhan, K. & Ibata, R. A. 2018, MNRAS, 477, 4063
Malhan, K., Ibata, R. A., & Martin, N. F. 2018, MNRAS, 481, 3442
Mardini, M. K., Li, H., Placco, V. M., et al. 2019a, The Astrophysical Journal, 875, 89
Mardini, M. K., Placco, V. M., Taani, A., Li, H., & Zhao, G. 2019b, The Astrophysical Journal, 882, 27
Martell, S. L., Shetrone, M. D., Lucatello, S., et al. 2016, ApJ, 825, 146
Martin, N. F., Ibata, R. A., Rich, R. M., et al. 2014, ApJ, 787, 19
Mashonkina, L., Gehren, T., Shi, J. R., Korn, A. J., & Grupp, F. 2011, A&A, 528, A87
Mashonkina, L., Korn, A. J., & Przybilla, N. 2007, A&A, 461, 261
Mateu, C., Read, J. I., & Kawata, D. 2018, MNRAS, 474, 4112
Matteucci, F. & Recchi, S. 2001, ApJ, 558, 351
McConnachie, A. 2019, in American Astronomical Society Meeting Abstracts, Vol. 233, American Astronomical Society Meeting Abstracts #233, 128.03
McConnachie, A., Babusiaux, C., Balogh, M., et al. 2016a, arXiv e-prints, arXiv:1606.00043
McConnachie, A. W. 2012, AJ, 144, 4
McConnachie, A. W., Babusiaux, C., Balogh, M., et al. 2016b, ArXiv e-prints
McConnachie, A. W., Ibata, R., Martin, N., et al. 2018, ApJ, 868, 55
McWilliam, A. 1997, ARA&A, 35, 503
Mészáros, S., Allende Prieto, C., Edvardsson, B., et al. 2012, AJ, 144, 120
Meynet, G., Ekström, S., & Maeder, A. 2006, A&A, 447, 623
Meynet, G., Hirschi, R., Ekstrom, S., et al. 2010, A&A, 521, A30
Munari, U., Sordo, R., Castelli, F., & Zwitter, T. 2005, A&A, 442, 1127
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., Sand ers, J. L., & Koposov, S. E. 2018b, MNRAS, 478, 5449
Myeong, G. C., Evans, N. W., Belokurov, V., Sand ers, J. L., & Koposov, S. E. 2018c, ApJ, 863, L28
Myeong, G. C., Vasiliev, E., Iorio, G., Evans, N. W., & Belokurov, V. 2019, arXiv e-prints, arXiv:1904.03185
Navarrete, C., Belokurov, V., Catelan, M., et al. 2018, arXiv e-prints
Newberg, H. J., Yanny, B., Rockosi, C., et al. 2002, ApJ, 569, 245
Nomoto, K., Kobayashi, C., & Tominaga, N. 2013, ARA&A, 51, 457
Nordlander, T., Amarsi, A. M., Lind, K., et al. 2017, A&A, 597, A6
Nordlander, T., Bessell, M. S., Da Costa, G. S., et al. 2019, MNRAS, 488, L109
Norris, J. E., Christlieb, N., Bessell, M. S., et al. 2012, ApJ, 753, 150
Norris, J. E., Christlieb, N., Korn, A. J., et al. 2007, ApJ, 670, 774
Norris, J. E. & Hawkins, M. R. S. 1991, ApJ, 380, 104
Norris, J. E. & Yong, D. 2019, ApJ, 879, 37
Oman, K., Starkenburg, E., & Navarro, J. 2017, Galaxies, 5, 33
Pagel, B. E. J. & Patchett, B. E. 1975, MNRAS, 172, 13
Peñarrubia, J., Belokurov, V., Evans, N. W., et al. 2010, MNRAS, 408, L26
Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825
Pier, J. R. 1983, ApJS, 53, 791
Pila-Díez, B., Kuijken, K., de Jong, J. T. A., Hoekstra, H., & van der Burg, R. F. J. 2014, A&A, 564, A18
Placco, V. M., Frebel, A., Beers, T. C., et al. 2014, ApJ, 781, 40

Placco, V. M., Frebel, A., Lee, Y. S., et al. 2015, ApJ, 809, 136
Plez, B. 2008, Physica Scripta Volume T, 133, 014003
Price-Jones, N. & Bovy, J. 2019, arXiv e-prints
Price-Whelan, A. M. & Bonaca, A. 2018, ApJ, 863, L20
Price-Whelan, A. M., Johnston, K. V., Sheffield, A. A., Laporte, C. F. P., & Sesar, B. 2015, MNRAS, 452, 676
Purcell, C. W., Bullock, J. S., Tollerud, E. J., Rocha, M., & Chakrabarti, S. 2011, Nature, 477, 301
Robertson, B., Bullock, J. S., Font, A. S., Johnston, K. V., & Hernquist, L. 2005, ApJ, 632, 872
Robin, A. C., Reylé, C., Derrière, S., & Picaud, S. 2003, A&A, 409, 523
Rocha-Pinto, H. J., Majewski, S. R., Skrutskie, M. F., Crane, J. D., & Patterson, R. J. 2004, ApJ, 615, 732
Rodgers, A. W., Harding, P., & Sadler, E. 1981, ApJ, 244, 912
Ryan, S. G. & Norris, J. E. 1991, AJ, 101, 1865
Salaris, M. & Cassisi, S. 2005, Evolution of Stars and Stellar Populations (John Wiley & Sons, Ltd.), 400
Salvadori, S., Schneider, R., & Ferrara, A. 2007, MNRAS, 381, 647
Sandage, A. R. & Eggen, O. J. 1959, MNRAS, 119, 278
Sanderson, R. E., Helmi, A., & Hogg, D. W. 2015, ApJ, 801, 98
Sanderson, R. E., Wetzel, A., Loebman, S., et al. 2018, arXiv e-prints, arXiv:1806.10564
Schlaufman, K. C. & Casey, A. R. 2014, ApJ, 797, 13
Schönrich, R., Asplund, M., & Casagrande, L. 2011, MNRAS, 415, 3807
Schönrich, R., Asplund, M., & Casagrande, L. 2014, ApJ, 786, 7
Schörck, T., Christlieb, N., Cohen, J. G., et al. 2009, A&A, 507, 817
Schwarzschild, M., Searle, L., & Howard, R. 1955, ApJ, 122, 353
Searle, L. & Sargent, W. L. W. 1972, ApJ, 173, 25
Sérsic, J. L. 1963, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 6, 41
Sesar, B., Hernitschek, N., Dierickx, M. I. P., Fardal, M. A., & Rix, H.-W. 2017a, ApJ, 844, L4
Sesar, B., Hernitschek, N., Mitrović, S., et al. 2017b, AJ, 153, 204
Sestito, F., Longeard, N., Martin, N. F., et al. 2019a, MNRAS, 46
Sestito, F., Martin, N. F., Starkenburg, E., et al. 2019b, arXiv e-prints, arXiv:1911.08491
Shapley, H. 1918, PASP, 30, 42
Sharma, S., Bland-Hawthorn, J., Johnston, K. V., & Binney, J. 2011, ApJ, 730, 3
Sharma, S., Johnston, K. V., Majewski, S. R., et al. 2010, ApJ, 722, 750
Sheffield, A. A., Price-Whelan, A. M., Tzanidakis, A., et al. 2018, ApJ, 854, 47
Shi, W. B., Chen, Y. Q., Carrell, K., & Zhao, G. 2012, ApJ, 751, 130
Shipp, N., Drlica-Wagner, A., Balbinot, E., et al. 2018, ApJ, 862, 114
Simpson, J. D., Martell, S. L., Da Costa, G., et al. 2019, MNRAS, 482, 5302
Sirko, E., Goodman, J., Knapp, G. R., et al. 2004, AJ, 127, 899
Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163
Skúladóttir, Á., Tolstoy, E., Salvadori, S., et al. 2015, A&A, 574, A129
Slater, C. T., Nidever, D. L., Munn, J. A., Bell, E. F., & Majewski, S. R. 2016, ApJ, 832, 206
Sommer-Larsen, J., Christensen, P. R., & Carter, D. 1989, MNRAS, 238, 225
Spite, M., Caffau, E., Bonifacio, P., et al. 2013, A&A, 552, A107
Starkenburg, E., Aguado, D. S., Bonifacio, P., et al. 2018, MNRAS, 481, 3838
Starkenburg, E., Helmi, A., Morrison, H. L., et al. 2009, ApJ, 698, 567
Starkenburg, E., Hill, V., Tolstoy, E., et al. 2013, A&A, 549, A88
Starkenburg, E., Martin, N., Youakim, K., et al. 2017, MNRAS, 471, 2587
Starkenburg, E., Oman, K. A., Navarro, J. F., et al. 2017b, MNRAS, 465, 2212
Starkenburg, E., Shetrone, M. D., McConnachie, A. W., & Venn, K. A. 2014, MNRAS, 441, 1217
Starkenburg, E., Youakim, K., Martin, N., et al. 2019, MNRAS, 2542
Takada, M., Ellis, R. S., Chiba, M., et al. 2014, PASJ, 66, R1
Takahashi, K., Umeda, H., & Yoshida, T. 2014, ApJ, 794, 40
Thomas, G. F., Famaey, B., Ibata, R., Lüghausen, F., & Kroupa, P. 2017, A&A, 603, A65
Thomas, G. F., Laporte, C. F. P., McConnachie, A. W., et al. 2018a, MNRAS
Thomas, G. F., McConnachie, A. W., Ibata, R. A., et al. 2018b, MNRAS, 481, 5223
Tibshirani, R. 1996, Journal of the Royal Statistical Society. Series B (Methodological), 267
Tody, D. 1986, in Proc. SPIE, Vol. 627, Instrumentation in astronomy VI, ed. D. L. Crawford, 733

Tody, D. 1993, in Astronomical Society of the Pacific Conference Series, Vol. 52, Astronomical Data Analysis Software and Systems II, ed. R. J. Hanisch, R. J. V. Brissenden, & J. Barnes, 173
Tolstoy, E., Hill, V., & Tosi, M. 2009, ARA&A, 47, 371
Tominaga, N., Iwamoto, N., & Nomoto, K. 2014, ApJ, 785, 98
Torrealba, G., Belokurov, V., Koposov, S. E., et al. 2019, MNRAS, 488, 2743
Tumlinson, J. 2010, ApJ, 708, 1398
Umeda, H. & Nomoto, K. 2003, Nature, 422, 871
Umeda, H. & Nomoto, K. 2005, ApJ, 619, 427
Venn, K., Kielty, C., Sestito, F., et al. 2019, arXiv e-prints, arXiv:1910.06340
Vera-Ciro, C. & Helmi, A. 2013, ApJ, 773, L4
Vickers, J. J., Grebel, E. K., & Huxor, A. P. 2012, AJ, 143, 86
Vrug, A., ter Braak, C., Dicks, C., et al. 2009, International Journal of Nonlinear Sciences & Numerical Simulation, 10, 273:290
Wallerstein, G. 1962, ApJS, 6, 407
Watkins, L. L., Evans, N. W., Belokurov, V., et al. 2009, MNRAS, 398, 1757
Watkins, L. L., van der Marel, R. P., Bellini, A., Libralato, M., & Anderson, J. 2019, arXiv e-prints, arXiv:1908.00789
White, S. D. M. & Springel, V. 2000, in The First Stars, ed. A. Weiss, T. G. Abel, & V. Hill, 327
Wilhelm, R., Beers, T. C., & Gray, R. O. 1999, AJ, 117, 2308
Wolf, C., Onken, C. A., Luvaul, L. C., et al. 2018, Pub. of the Astr. Soc. of Australia, 35, e010
Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868
Xu, Y., Liu, C., Xue, X.-X., et al. 2018, MNRAS, 473, 1244
Xue, X.-X., Rix, H.-W., Ma, Z., et al. 2015, ApJ, 809, 144
Xue, X.-X., Rix, H.-W., Yanny, B., et al. 2011, ApJ, 738, 79
Xue, X. X., Rix, H. W., Zhao, G., et al. 2008, ApJ, 684, 1143
Yang, C., Xue, X.-X., Li, J., et al. 2019, ApJ, 880, 65
Yanny, B., Newberg, H. J., Kent, S., et al. 2000, ApJ, 540, 825
Yanny, B., Rockosi, C., Newberg, H. J., et al. 2009, AJ, 137, 4377
Yong, D., Norris, J. E., Bessell, M. S., et al. 2013a, ApJ, 762, 26
Yong, D., Norris, J. E., Bessell, M. S., et al. 2013b, ApJ, 762, 27
Yoon, J., Beers, T. C., Placco, V. M., et al. 2016, ApJ, 833, 20
Yoon, J., Beers, T. C., Tian, D., & Whitten, D. D. 2019, ApJ, 878, 97
York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, AJ, 120, 1579
Youakim, K., Starkenburg, E., Aguado, D. S., et al. 2017, MNRAS, 472, 2963
Zuo, W., Du, C., Jing, Y., et al. 2017, ApJ, 841, 59

Acknowledgements

In order to accomplish great things you first have to assemble a strong support network to help you get there. I was very fortunate to have had a wonderful group of friends, family, colleagues and mentors to help me succeed in completing this thesis. This section is to recognize all of you who have made this possible, and who have contributed to making these last 4 years as special as they were.

First of all, I want to give a shout out to my editing team – Else, Brendan, Friedrich, Lotti, Johannes and Jenn. If there are any typos in here I blame you!...Just kidding, without you guys this thesis would be no where near as silky and smooth as it is. So thank you!

To my roommates, Greta and Johannes, it's been an absolute blast living with you guys for the past 4 years. Thank you for the good times, and thank you for your patience during the busy times when I couldn't contribute as much as I should have around the apartment. You guys definitely picked up the slack, and without you I don't know if I would have survived long enough to finish!

To my family, mom, dad and Alex, thanks for supporting me unconditionally, and for being so enthusiastic and excited about me doing this PhD. It made it a lot easier to be so far away from home knowing that you guys were behind me 100%. Grandpa, Claudette, Grandma, Uncle Doudi and Aunt Lori, thanks for keeping in touch and making home seem not quite so far away as it otherwise would have. And it was always fun having you guys come and visit to go on road trips and explore so many cool places (and of course, I can't not mention Greg and Kath here too!).

Anke, my PhD sister from another mister, your abilities in astronomy never cease to amaze me. I have no doubt that you will achieve greatness...just don't forget about us normal people when you reach the top!

David, you are the one man, you are the one! You are a star my friend. It's been a great pleasure getting to know you over these past years, and collaborating with you was hands down one of the funnest parts of the PhD.

Kim, this acknowledgements section would not be complete without you in it. I remember when I was a faceless undergrad who wanted to get into astronomy, but just needed someone to give me a chance. That person was you, and without that, none of what came after would have been possible. Keep on inspiring young students. You are the coolest astronomer I know!

To all of the great friends I made at the AIP – Cam, Guillaume, Dalal, Brazili-Anna, Ismael, Petra, Eloy, Javi, Friedrich – thanks to you guys most days hardly felt like work. And Jenn, we had some amazing times! (for a more detailed list of the shenanigans we got up to, see the acknowledgments in Jenn's thesis). You might even say we had a "nose smashing" good time. Thanks for being an awesome friend and mentor.

To my friends back home that kept me connected and up to date with life back in Canada: Trent, Pete, Kelsey, Brando, Nate, Brendan, Nick, Molly, Liam, Bryan, Jared, Brad, Julena, thanks for making an effort to stay in touch, even with the nasty time change.

Gal, it's been a wild ride! What else can I say but that I hope it continues for many years to come. It was pure good fortune that we overlapped pretty much perfectly at the AIP, and even though looking back now it seems to have gone by so fast, at least we can say we definitely made the most of it. So here's to many more stimulating discussions and interesting discoveries in the future. You are one of a kind my friend!

To my supervisor Else, I cannot in this short paragraph describe what a pleasure it has been to be your student. You always made sure to give my well being and happiness top priority. There were a lot of difficult challenges that I had to face during the course of this PhD, and you had my back the entire time. I am immensely grateful for your supervision, mentorship, and the fact that you are just such a downright great person. From the bottom of my heart, thank you.

Thank you to everyone else who I did not mention, but who have been a part of these last 4 years and contributed to this wonderful phase in my life. I will never forget the time that I have spent here in Berlin and Potsdam, and all of the fond memories and experiences that I was lucky enough to share with you all.