



**HAL**  
open science

## The Pristine survey XIII: uncovering the very metal-poor tail of the thin disc

Emma Fernández-Alvar, Georges Kordopatis, Vanessa Hill, Else Starkenburg, Akshara Viswanathan, Nicolas Martin, Guillaume Thomas, Julio Navarro, Khyati Malhan, Federico Sestito, et al.

► **To cite this version:**

Emma Fernández-Alvar, Georges Kordopatis, Vanessa Hill, Else Starkenburg, Akshara Viswanathan, et al.. The Pristine survey XIII: uncovering the very metal-poor tail of the thin disc. Monthly Notices of the Royal Astronomical Society, 2021, 508 (1), pp.1509-1525. 10.1093/mnras/stab2617. hal-03374625

**HAL Id: hal-03374625**




**<https://hal.science/hal-03374625v1>**

Submitted on 21 Apr 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The Pristine survey XIII: uncovering the very metal-poor tail of the thin disc

Emma Fernández-Alvar<sup>1</sup><sup>1\*</sup>, Georges Kordopatis,<sup>1</sup> Vanessa Hill,<sup>1</sup> Else Starkenburg,<sup>2</sup> Akshara Viswanathan,<sup>2</sup> Nicolas F. Martin<sup>3</sup><sup>3</sup>, Guillaume F. Thomas,<sup>4,5</sup> Julio F. Navarro,<sup>6</sup> Khyati Malhan,<sup>7</sup> Federico Sestito<sup>6</sup><sup>6</sup>, Jonay I. González Hernández<sup>4,5</sup> and Raymond G. Carlberg<sup>8</sup>

<sup>1</sup>Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Bd de l'Observatoire, CS 34229, F-06304 Nice cedex 4, France

<sup>2</sup>Kapteyn Astronomical Institute, University of Groningen, PO Box 800, NL-9700 AV Groningen, the Netherlands

<sup>3</sup>Université de Strasbourg, CNRS, Observatoire astronomique de Strasbourg, UMR 7550, F-67000 Strasbourg, France

<sup>4</sup>Instituto de Astrofísica de Canarias, E-38205 La Laguna, Tenerife, Spain

<sup>5</sup>Universidad de La Laguna, Dpto. Astrofísica, E-38206 La Laguna, Tenerife, Spain

<sup>6</sup>Department of Physics and Astronomy, University of Victoria, Victoria, BC V8P 5C2, Canada

<sup>7</sup>The Oskar Klein Centre for Cosmoparticle Physics, Department of Physics, Stockholm University, AlbaNova, SE-10691 Stockholm, Sweden

<sup>8</sup>Department of Astronomy & Astrophysics, University of Toronto, Toronto, ON M5S 3H4, Canada

Accepted 2021 September 6. Received 2021 August 7; in original form 2021 June 8

## ABSTRACT

We evaluate the rotational velocity of stars observed by the Pristine survey towards the Galactic anticentre spanning a wide range of metallicities from the extremely metal-poor regime ( $[\text{Fe}/\text{H}] < -3$ ) to nearly solar metallicity. In the Galactic anticentre direction, the rotational velocity ( $V_\phi$ ) is similar to the tangential velocity in the galactic longitude direction ( $V_\ell$ ). This allows us to estimate  $V_\phi$  from *Gaia* early data release 3 (Gaia EDR3) proper motions for stars without radial velocity measurements. This substantially increases the sample of stars in the outer disc with estimated rotational velocities. Our stellar sample towards the anticentre is dominated by a kinematical thin disc with a mean rotation of  $\sim -220 \text{ km s}^{-1}$ . However, our analysis reveals the presence of more stellar substructures. The most intriguing is a well-populated extension of the kinematical thin disc down to  $[\text{Fe}/\text{H}] \sim -2$ . A scarcer fast-rotating population reaching the extremely metal-poor regime down to  $[\text{Fe}/\text{H}] \sim -3.5$  is also detected but without statistical significance to unambiguously state whether this is the extremely metal-poor extension of the thin disc or the high-rotating tail of hotter structures (like the thick disc or the halo). In addition, a more slowly rotating kinematical thick disc component is also required to explain the observed  $V_\ell$  distribution at  $[\text{Fe}/\text{H}] > -1.5$ . Furthermore, we detect signatures of a ‘heated disc’, the so-called *Splash*, at metallicities higher than  $\sim -1.5$ . Finally, at  $[\text{Fe}/\text{H}] < -1.5$  our anticentre sample is dominated by a kinematical halo with a net prograde motion.

**Key words:** Galaxy: abundances – Galaxy: disc – Galaxy: kinematics and dynamics.

## 1 INTRODUCTION

Our understanding of the processes that formed the Galaxy benefits from information encoded in the properties of stars, mainly their chemical composition and their orbital motion. In particular, the analysis of the first stellar populations formed in the Universe contributes valuable hints to the earliest stages of galactic formation (Freeman & Bland-Hawthorn 2002). These stars are characterized by their very low metal content and have been the object of thorough searches over the last decades – see the reviews by Beers & Christlieb (2005) and Frebel & Norris (2015).

To unveil which processes led to the current configuration of the Galaxy, we need the orbital characterization of its stars. The velocity of a star with respect to the Sun is measured in three components: the radial velocity ( $v_{\text{rad}}$ ) and the proper motions ( $\mu_\alpha$  and  $\mu_\delta$ ), which correspond to the velocity projections along the line of sight, the right ascension, and the declination directions,

respectively. For convenience, sometimes the proper motions are expressed in the galactic reference system ( $\mu_\ell$  and  $\mu_b$ ). The vast improvement in terms of accuracy, precision, and coverage of astrometric data (proper motions and parallaxes) provided by the *Gaia* mission (Gaia Collaboration 2016, 2018a, 2020) has revealed the dynamical configuration of the stellar Galactic populations with an unprecedented level of detail. As some examples, we have now an expanded insight into the degree at which the Galaxy is out of equilibrium (Gaia Collaboration 2018c), evidence that the Galactic halo in the inner regions is dominated by an accreted system and a heated protodisc (Haywood et al. 2018; Helmi et al. 2018; Myeong et al. 2018; Di Matteo et al. 2019; Gallart et al. 2019; Myeong et al. 2019; Belokurov et al. 2020a).

The *Gaia* mission is also planning to provide spectroscopic metallicities – from the CaII near-infrared triplet region, as explained in Recio-Blanco et al. (2016). However, no chemical abundance estimates are published yet, radial velocities are available only for the brighter part of the catalogue, and even from ground-based complement, and samples of low-metallicity stars with full kinematics are still of modest size. The Pristine survey (Starkenburg

\* E-mail: emma.fernandez-alvar@oca.eu, emmafalvar@gmail.com

et al. 2017) was conceived as a narrow-band photometric survey centred in the CaII doublet H&K at 3933 and 3968 Å, a spectral feature very sensitive to the metallicity<sup>1</sup> content of the star when combined with broad-band photometry. Its main goal is to provide a comprehensive view of the stellar metallicity distribution of the metal-poor Galaxy, with a particular focus on discovering the pristine metallicity stars born from the products of the first episodes of star formation. The survey, carried out with the 3.6-m optical/infrared Canadian–French–Hawaiian Telescope (CFHT) on the Mauna Kea Hawaiian Observatory, already covers over 5000 square degrees on the sky, observing millions of targets over most of the northern Galactic hemisphere. From their photometry, a metallicity estimate is derived for every object observed. Thus, the combination of the Pristine and the *Gaia* surveys opens huge possibilities for the characterization of the first stellar populations and their link with the subsequent star formation in the Galaxy.

Indeed, kinematical analyses of metal-poor stars ( $[\text{Fe}/\text{H}] < -1$ ) making use of *Gaia* astrometry have already provided new results regarding the characterization of the first generations of stars. For instance, Sestito et al. (2019) discovered extremely and ultra metal-poor stars ( $[\text{Fe}/\text{H}] < -3$  and  $[\text{Fe}/\text{H}] < -4$ ) rotating fast in prograde orbits close to the Galactic plane, never detected before. Di Matteo et al. (2019) showed signature of rotation in stars down to  $[\text{Fe}/\text{H}] < -2$ . Carollo et al. (2019) obtained evidence that the thick disc comprises two distinct and overlapping stellar populations with different kinematic properties and chemical compositions. More studies after them have confirmed these results.

Unfortunately, radial velocities can be measured only from stellar spectra,<sup>2</sup> and only  $\sim 0.5$  per cent of stars in the Pristine survey have a spectroscopic follow-up (Caffau et al. 2017; Youakim et al. 2017; Aguado et al. 2019a; Bonifacio et al. 2019; Caffau et al. 2020; KIELTY et al. 2020; Venn et al. 2020). Consequently, the fraction of Pristine stars with full velocity characterization (proper motions and radial velocity measurements) is extremely low, and it is worth searching for specific directions on the sky where geometry cancels the contribution of the radial velocity in some of the velocity components with respect to the Galactic centre.

A number of studies have already made use of this kind of approach in the past. For instance, before the advent of the *Gaia* survey, there were few accurate proper motions measurements, but a number of publications used privileged directions where the stellar rotational velocity,  $V_\phi$ , is identical to  $v_{\text{rad}}$  in order to evaluate the rotation of the Galaxy (e.g. Morrison, Flynn & Freeman 1990; Wyse 2016; Kordopatis et al. 2013b, 2017).

With the advent of *Gaia*, the problem has now become the opposite: we have high-quality proper motions but the lack of spectra prevents us from having a  $v_{\text{rad}}$  estimation for the majority of stars in the survey. At exactly the anticentre direction  $V_\phi$  is perpendicular to the line of sight and latitude directions and, consequently, the projection of  $V_\phi$  along these directions is null. In other words, to derive  $V_\phi$  from the observed  $v_{\text{rad}}$ ,  $\mu_b$  and  $\mu_\ell$  velocity components, only the latter is required. This fact allows us to study the  $V_\phi$  distributions from proper motions only, without radial velocities. This is still a good approximation up to a few degrees away from the anticentre as we will demonstrate in this paper.

<sup>1</sup>From now on we will refer to metallicity as the abundance of iron with respect to hydrogen, and we will express it as the difference with respect to the Sun abundance in logarithmic scale:  $[\text{Fe}/\text{H}] = \log_{10}(\text{N}_{\text{Fe}}/\text{N}_{\text{H}})_{\text{star}} - \log_{10}(\text{N}_{\text{Fe}}/\text{N}_{\text{H}})_{\odot}$

<sup>2</sup>However, see Lindegren & Dravins (2021).

Some previous works have already inferred the rotational velocity of stars from proper motions only towards the anticentre. For example, Thomas et al. (2019) explored the disc flaring with this approximation, using blue stragglers. An & Beers (2020a) and An & Beers (2021) showed how stars rotating slower than the disc, i.e. the kinematical halo, are structured in several components based on their location on the  $V_\phi$  versus  $[\text{Fe}/\text{H}]$  space. They detected signatures of the *Splash* thick disc (Belokurov et al. 2020a), the metal-weak thick disc (Norris, Bessell & Pickles 1985a), the *Gaia*–Enceladus–Sausage – GES (Belokurov et al. 2018; Helmi et al. 2018) accreted system, and an additional component dominated by stars on retrograde orbits. In addition, as part of the science verification of the *Gaia* early third data release (*Gaia* EDR3), *Gaia* Collaboration (2021a) revealed the complex structure of the disc towards the anticentre with, among others, the detection of velocity oscillations throughout the plane and the identification of stellar overdensities associated to Monoceros, TriAnd, and the anticentre stream. Further structures have been detected recently in Laporte, Koposov & Belokurov (2021) also using this approximation.

Stars in the outer disc of the Milky Way are less affected by the Galactic potential and the dynamical relaxing time-scales are longer, thus favouring the detection of kinematical signatures of the processes that the Galaxy underwent (e.g. Binney & Tremaine 2008). The Pristine survey provides a map of stars in a wide range of metallicities, with uniquely robust estimates at the lowest metallicities when compared to other photometric metallicity inferences (Starkenburger et al. 2017; Youakim et al. 2017; Aguado et al. 2019b), giving key information about star formation from the earliest stages of galactic formation. For all these reasons, the rotational velocity distribution of the stars presented in this work provides a unique insight on this regard.

In this work, we aim to combine the Pristine metallicities with *Gaia* astrometry to explore the rotational velocity distribution of stars in the Galaxy towards the anticentre direction with a special emphasis on metal-poor stars. Indeed, Pristine, with its robust photometric metallicity estimates in the very (VMP,  $[\text{Fe}/\text{H}] < -2$ ) and extremely (EMP,  $[\text{Fe}/\text{H}] < -3$ ) metal-poor regime, provides a unique opportunity to probe the kinematics of these populations in the anticentre direction. We explain our data selection, the methodology to determine the rotational velocity from proper motions only, and the adopted stellar metallicities in Section 2. Section 3 explores the rotational velocity distributions as a function of  $[\text{Fe}/\text{H}]$ . We discuss the implications of our results in Section 4. Finally, our conclusions are summarized in Section 5.

## 2 METHODOLOGY AND DATA SELECTION

We cross-correlate Pristine observations, obtained between 2015 and 2020, with *Gaia* EDR3 (*Gaia* Collaboration 2021b) and the Sloan Digital Sky Survey photometry (Fukugita et al. 1996; Gunn et al. 1998; Abazajian et al. 2009; Doi et al. 2010) in order to estimate  $[\text{Fe}/\text{H}]$  and the rotational velocity component. In order to justify the selection criteria applied, we will first explain the bases of our methodology.

### 2.1 Azimuthal velocities derived from the proper motions towards the anticentre

First, we present the geometrical relations between proper motions and the rotational velocity component. Proper motions (in  $\text{mas yr}^{-1}$ ) are transformed from the equatorial to the galactic coordinate system

$(\ell, b)$  through the equations in Poleski (2013) that perform the rotations needed to change reference system:

$$\mu_\ell = \frac{1}{\cos(b)(c_1\mu_\alpha + c_2\mu_\delta)} \quad (1)$$

$$\mu_b = \frac{1}{\cos(b)(c_2\mu_\alpha + c_1\mu_\delta)}, \quad (2)$$

with  $c_1$  and  $c_2$  defined as

$$c_1 = \sin(\delta_{\text{GP}})\cos(\delta) - \cos(\delta_{\text{GP}})\sin(\delta)\cos(\alpha - \alpha_{\text{GP}}) \quad (3)$$

$$c_2 = \cos(\delta_{\text{GP}})\sin(\alpha - \alpha_{\text{GP}}) \quad (4)$$

and  $(\alpha_{\text{GP}}, \delta_{\text{GP}}) = (192.860, 27.128)$  degrees are the equatorial coordinates of the North Galactic Pole (Perryman 1997).

Then, we transform from angular to linear velocities:

$$V_\ell = 4.74d_{\text{sun}}\mu_\ell \quad (5)$$

$$V_b = 4.74d_{\text{sun}}\mu_b, \quad (6)$$

where  $d_{\text{sun}}$  is the distance in kpc inferred from the inverse of the *Gaia* parallaxes.<sup>3</sup>  $V_\ell$  and  $V_b$  are the linear projections of the proper motions over the galactic coordinates  $(\ell, b)$ .

We adopt the right-handed galactocentric Cartesian and cylindrical coordinate systems, i.e. the  $x$ -axis positive towards the Sun, the  $z$ -axis positive towards the North Galactic Pole, and the rotation of the Local Standard of Rest (LSR) negative. In this system, then, a negative value of  $V_\phi$  corresponds to a prograde rotation. We also adopt the galactocentric distance of the Sun  $R_{\text{sun}} = 8.3$  kpc (Schönrich 2012).

The Cartesian velocity components  $U$ ,  $V$ , and  $W$  with respect to the LSR are linked to  $V_\ell$  and  $V_b$  through the following equations:

$$U = v_{\text{rad}} \cos(\ell) \cos(b) - V_\ell \sin(\ell) - V_b \cos(\ell) \sin(b) \quad (7)$$

$$V = v_{\text{rad}} \sin(\ell) \cos(b) + V_\ell \cos(\ell) - V_b \sin(\ell) \sin(b) \quad (8)$$

$$W = v_{\text{rad}} \sin(b) + V_b \cos(b) \quad (9)$$

and we correct them for the solar motion with respect to the LSR  $(U_\odot, V_\odot, W_\odot) = (11.1, 12.24, 7.25)$  km s<sup>-1</sup> (Schönrich, Binney & Dehnen 2010) and the rotational velocity component of the LSR  $V_{\text{LSR}} = -238$  km s<sup>-1</sup> (Schönrich 2012).

$U$ ,  $V$ , and  $W$  can be transformed to the cylindrical coordinate system defined by the coordinates  $R, \phi$  and  $z$  by:

$$V_R = U \cos(\phi) + V \sin(\phi) \quad (10)$$

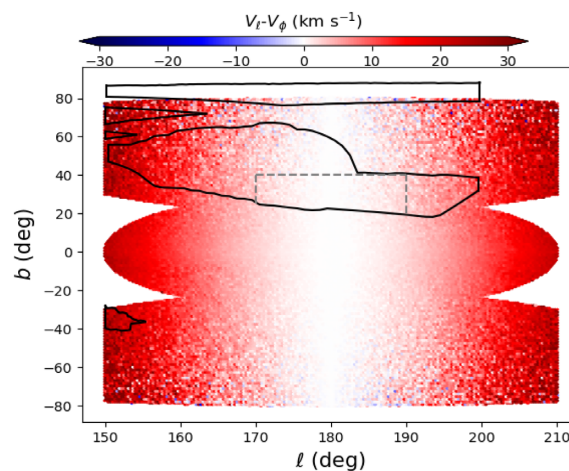
$$V_\phi = -U \sin(\phi) + V \cos(\phi) \quad (11)$$

$$V_z = W, \quad (12)$$

where  $V_R$ ,  $V_\phi$ , and  $V_z$  are the radial, rotational, and vertical velocity components. Then,  $V_\phi$  can be expressed as a function of  $(v_{\text{rad}}, V_\ell, V_b)$  by:

$$\begin{aligned} V_\phi = & v_{\text{rad}}[\sin(\ell) \cos(b) \cos(\phi) - \cos(\ell) \cos(b) \sin(\phi)] \\ & + V_\ell[\sin(\ell) \sin(\phi) + \cos(\ell) \cos(\phi)] \\ & + V_b[\cos(\ell) \sin(b) \sin(\phi) - \sin(\ell) \sin(b) \cos(\phi)]. \end{aligned} \quad (13)$$

<sup>3</sup>In this paper, we consider only stars with a low relative uncertainty in parallaxes <20 per cent. As a consequence, the subsequent uncertainties in distances are also low and there is no need to apply priors to obtain realistic distances – see Bailer-Jones (2015).



**Figure 1.** Average  $V_\phi - V_\ell$  as a function of galactic coordinates  $(\ell, b)$  measured in the *Gaia* Mock catalogue. The plot is a superposition of two projected cones centred at  $(\ell, b) = (180, 0)^\circ$ ,  $(\ell, b) = (0, -50)^\circ$ , and  $(\ell, b) = (0, 50)^\circ$  with a radius of  $30^\circ$ . The grey dashed line delimits the selected region for our analysis.

Thus, for a star located in the direction of the anticentre where  $\ell = 180^\circ$ ,  $b = 0^\circ$ , and  $\phi = 0^\circ$ , the rotational velocity component of the star,  $V_\phi$ , is identical to the linear projection of the proper motion measured over the direction of the galactic longitude,  $V_\ell$ .

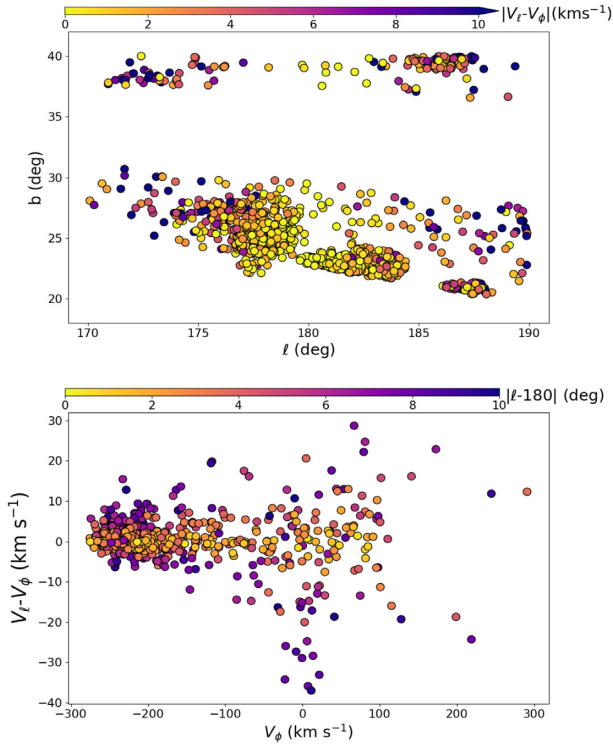
In the case of stars observed away from the anticentre, there is a dependence of  $V_\phi$  on  $v_{\text{rad}}$  and  $V_b$ , which is not zero. However, assimilating  $V_\phi$  to  $V_\ell$  remains a good approximation in the surroundings of the anticentre. Considering the sky area observed by Pristine towards the anticentre (marked with black contours in Fig. 1), we evaluated which are the regions where we can use  $V_\ell$  as a good approximation of  $V_\phi$  for our Pristine stars. We evaluated the *Gaia* eDR3 mock catalogue provided by the *Gaia* Archive<sup>4</sup> in the directions  $(\ell, b) = (180, 0)^\circ$ ,  $(\ell, b) = (0, -50)^\circ$ , and  $(\ell, b) = (0, 50)^\circ$  with a radius of  $30^\circ$ . We calculated  $V_\phi$  following equations (7)–(12), and we measured the differences between  $V_\phi$  and  $V_\ell$  (obtained based on equations 1–6), as shown with the colour scale in Fig. 1. This analysis shows that  $V_\ell$  follows  $V_\phi$  with an accuracy better than  $10$  km s<sup>-1</sup> in  $(170^\circ < \ell < 190^\circ)$ . Fig. 1 shows that at latitudes higher than  $b = 40^\circ$  and lower than  $b = -40^\circ$ , the deviation increases more rapidly when departing from  $\ell = 180^\circ$  than at  $-40 < b < 40^\circ$ .

Looking at the Pristine footprint, we decided to constrain our analysis to the region  $(170^\circ < \ell < 190^\circ, 20^\circ < b < 40^\circ)$  to select a monolithic region in  $(\ell, b)$  with low  $|V_\ell - V_\phi|$  deviations to keep a simpler relation with the distance to the plane and the centre,  $z$  and  $R$ .

As an additional verification, we computed the  $V_\phi$  for the Pristine stars in the anticentre that were also observed spectroscopically (the spectroscopic sample used for the metallicity calibration that we describe in Section 2.2) and had a measured  $v_{\text{rad}}$  from previous works (see Yanny et al. 2009 and Aguado et al. 2019b for more details) and compared their  $V_\phi$  with the resulting  $V_\ell$  from proper motions. We evaluated the stars located in the selected region  $(170^\circ < \ell < 190^\circ, 20^\circ < b < 40^\circ)$  and applied the same astrometry and photometry quality criteria as for the main analysis that we explain in Section 2.2. The comparison of  $V_\ell$  with  $V_\phi$ , shown in Fig. 2 reveals as expected that

<sup>4</sup>We made use of the `gaiaedr3.gaia_source.simulation` catalogue available through <https://gea.esac.esa.int/archive/>.



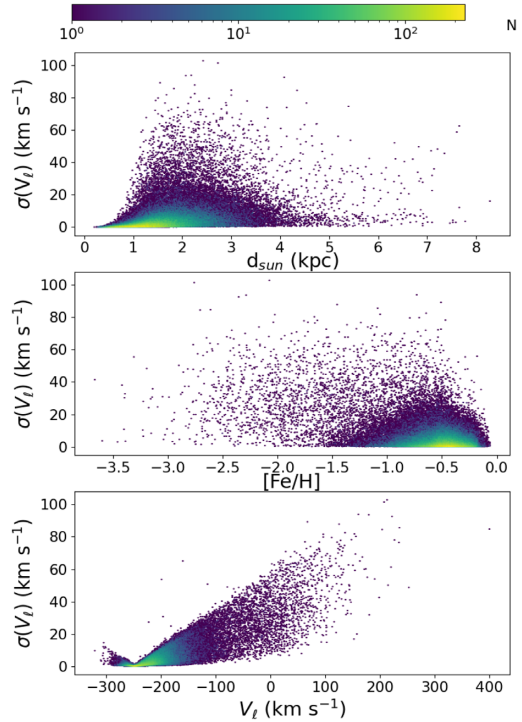


**Figure 2.** Top panel:  $\ell$  and  $b$  distribution of our training sample located in the *ac* colour coded with the difference of  $V_\ell$  compared with  $V_\phi$  measured from Gaia EDR3 proper motions and the spectroscopic  $v_{\text{rad}}$ . Bottom panel: The same  $V_\ell - V_\phi$  differences as shown in the top panel but now as a function of the  $V_\phi$ , colour coded by  $|\ell - 180^\circ|$ .

the scatter in  $V_\ell - V_\phi$  increases with  $|\ell - 180|$  (i.e.  $\ell$  departs from the anticentre). For the bulk of stars the differences are within the expected uncertainty from the simulations, i.e. lower than  $10 \text{ km s}^{-1}$ , with differences as large as  $40 \text{ km s}^{-1}$  in some cases. The stars showing largest differences are those with  $V_\phi$  close to 0 or positive, i.e. those stars with no rotation or moving on retrograde orbits. These stars are expected to have a high radial velocity component. For such stars located away from the anticentre, their  $V_\ell$  is large because of the contribution of the projection of the radial velocity component, explaining the discrepancy with their low or opposite sign  $V_\phi$ .

We thus show that it is possible to characterize the rotational motion of stars without radial velocity measurements in the anticentre direction within the area ( $170^\circ < \ell < 190^\circ$ ,  $-40^\circ < b < 40^\circ$ ). This gives us the opportunity to combine the huge Gaia EDR3 data base of proper motions with the Pristine targets, which, in the anticentre, corresponds to a sample of 414 292 stars. We correct parallaxes from the zero-point obtained from quasars,  $-0.017 \text{ mas}$  (Lindgren et al. 2020a), and remove those stars with a relative uncertainty in parallaxes higher than 20 per cent and parallaxes equal to or lower than 0.

We measured the typical uncertainties in  $V_\ell$  via 1000 Monte Carlo realizations drawing from normal distributions of the uncertainties in parallax and proper motions. The mean of the derived uncertainties in  $V_\ell$  is  $4 \text{ km s}^{-1}$ . The parameter that dominates the  $V_\ell$  uncertainty is the error in the parallax. Fig. 3 shows the dependence of the uncertainties with respect to the line-of-sight distance,  $[\text{Fe}/\text{H}]$  and  $V_\ell$  itself. Fig. 3 also shows that for some stars, in particular, those with larger  $V_\ell$ , the uncertainties are significantly much larger than  $4 \text{ km s}^{-1}$ . We decided to include them and verify in the subsequent analysis that such large uncertainties do not affect our results.



**Figure 3.**  $V_\ell$  uncertainties as a function of the line-of-sight distance,  $d_{\text{sun}}$ ,  $[\text{Fe}/\text{H}]$  and  $V_\ell$ . The uncertainties are computed as the standard deviation of 1000  $V_\ell$  Monte Carlo realizations calculated by drawing from normal distributions of the parallax and proper motions uncertainties.

## 2.2 The $[\text{Fe}/\text{H}]$ determination

We follow the same methodology as described in Starkenburg et al. (2017) – hereafter S17 – to convert from the Pristine narrow-band Ca II-HK photometric observations to metallicity estimates. The Ca II-HK doublet at 3933 and 3968 Å is a very prominent spectral feature very sensitive to stellar metallicity variations, in particular, in metal-poor stars with  $[\text{Fe}/\text{H}] < -1$ . The Ca II-HK doublet is also temperature sensitive; thus, a metallicity estimation can be performed by combining the CaHK magnitude with broad-band colours in a colour–colour space that separates out the temperature and metallicity effects (see right-hand panel on figure 3 in S17). The position of a star in this space can then be transformed into a metallicity estimate. In brief, we calibrate metallicity estimates using a stellar sample (that we will call training sample), composed of stars with spectroscopic metallicities, complemented by colours computed from synthetic spectra at the most metal-poor end. In the Pristine plus broad-band colour–colour plane, we define ‘pixels’ for which we derive a mean metallicity based on all the training sample stars that are located in this colour–colour pixel. Thus, every location in the photometric space has a metallicity value associated to it, and a Pristine target, based on its Ca II-HK and SDSS photometry alone, is then assigned a metallicity based on this map. For a more detailed explanation, we refer the reader to S17.

The spectroscopic sample used on this calibration (training sample) has evolved since S17, and now include metallicities inferred from the SEGUE Spectroscopic Parameters Pipeline and released as part of the SEGUE DR12 (Yanny et al. 2009; Alam et al. 2015) of all Pristine targets with a SEGUE counterpart, and also metallicity estimates of very metal-poor stars ( $[\text{Fe}/\text{H}] \lesssim -2$ ) from the Pristine collaboration follow-up spectroscopy (Aguado et al. 2019b),

complemented with a relation derived from synthetic spectra at the most metal-poor end (as described in S17).

The resulting calibration provides reliable photometric metallicity estimates for metal-poor FGK stars (S17 Aguado et al. 2019b). An upper metallicity cut at  $[\text{Fe}/\text{H}] = -0.2$  is applied to the results, because of the high errors and biases found for those. Indeed, for metal-rich stars ( $[\text{Fe}/\text{H}] > -1$ ), the Ca II-HK lines are heavily saturated and depend heavily on the stars' luminosity (or gravity). In the following, we show how we improve the situation in the intermediate metallicity range ( $[\text{Fe}/\text{H}] \sim -1$ ) by taking the stellar luminosity into account in the calibration.

### 2.2.1 Refining the Pristine metallicity calibration with dwarf/giant classification

At relatively high metallicities,  $[\text{Fe}/\text{H}] > -1$ , a position in the Pristine colour–colour space is solely defined not only by the metallicity and effective temperature [traced by the broad-band ( $g - i$ )<sub>0</sub> colour] but also by the stellar surface gravity, making the calibration uncertain, unless the gravity is also known accurately, and taken into account in the calibration (thus also requiring a training sample with accurate gravities).

In this work, we improve the adopted  $[\text{Fe}/\text{H}]$ , making use of the *Gaia* absolute  $M_G$  magnitude, and  $BP - RP$  colour in order to classify our stars as dwarfs or giants. We estimate the colour excess of each target due to dust reddening based on the Schlegel maps (Schlegel, Finkbeiner & Davis 1998). These maps provide the colour excess corresponding to the  $B - V$  Johnson & Cousins magnitude system based on their location on the sky. We transform them to the *Gaia* magnitude system through the following equations:

$$G_0 = G - 0.86117 * 3.1 * E(B - V) \quad (14)$$

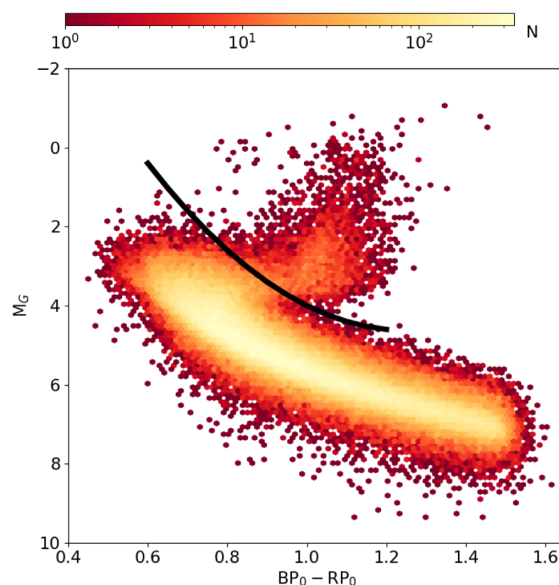
$$BP_0 = BP - 1.06126 * 3.1 * E(B - V) \quad (15)$$

$$RP_0 = RP - 0.64753 * 3.1 * E(B - V), \quad (16)$$

where  $G_0$ ,  $BP_0$ , and  $RP_0$  are the de-reddened  $G$ ,  $BP$ , and  $RP$  *Gaia* apparent magnitudes, and the coefficients multiplying  $E(B - V)$  are obtained through the estimator in <http://stev.oapd.inaf.it/cgi-bin/cm d.3.4> based on the relations of Cardelli, Clayton & Mathis (1989) and O'Donnell (1994). We use the SDSS de-reddened magnitudes and the Pristine de-reddening is described in S17.

We perform two different metallicity calibrations, each following the same principle as in S17: the first uses a training sample containing only dwarf stars, while the second is based on a training sample containing only giants, classified based on their  $M_G$  versus  $(BP - RP)_0$  colour–magnitude diagram (CMD), from *Gaia* EDR3. Then, we assign to each star the  $[\text{Fe}/\text{H}]$  derived from the corresponding metallicity calibration, depending on their locus in the same *Gaia* CMD shown in Fig. 4. The internal metallicity uncertainties computed for our anticentre sample are of 0.1 in average, with a dispersion of 0.04 for dwarfs, and 0.1 with a sigma of 0.02 for giants.

In Fig. 5 (panel c), we show the overall difference between the calibration with and without taking into account the luminosities of the training sample stars, while in Fig. 6, we show the location of the training sample and our anticentre sample in a CMD and in the Pristine colour–colour space in order to verify the impact of our dwarf/giant metallicity estimation. As expected, in the colour range where the stellar gravities are similar on the main sequence and giant branch, there is little difference between the two calibrations (e.g. base of the RGB, subgiants, and main sequence around G stars). In

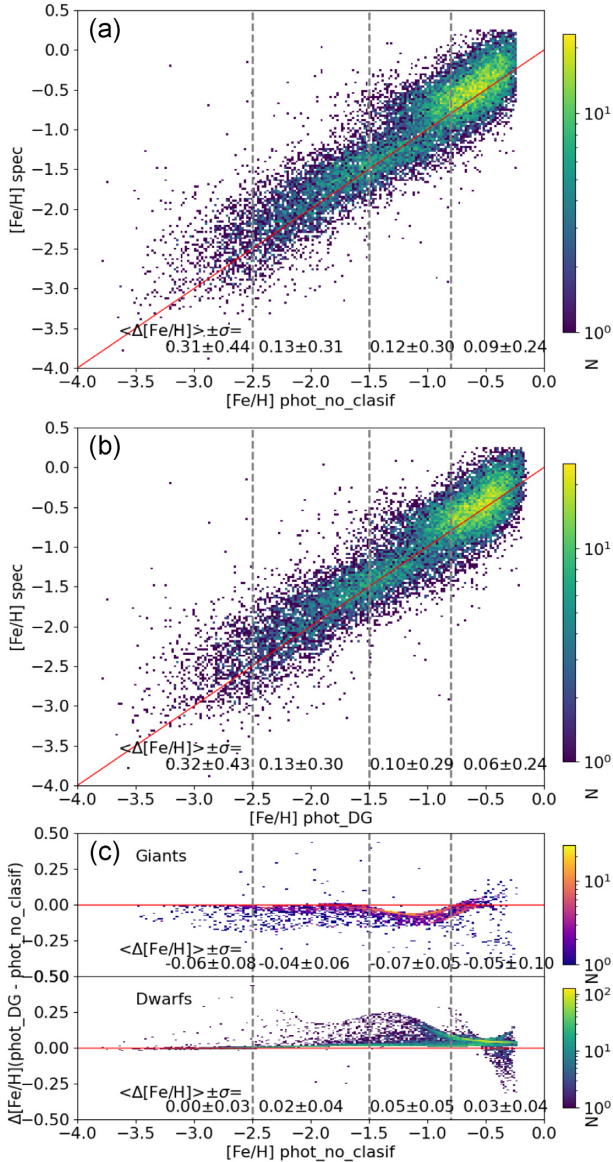


**Figure 4.** Colour–magnitude diagram of our Pristine sample towards the anticentre based on *Gaia* EDR3 photometry. The black line indicates the boundary chosen to classify these stars in dwarf and giants.

fact, in most of the parameter space, the difference between the global and dwarf/giants metallicity calibrations is smaller or equal to 0.2, and, as expected, the differences are most important in the metal-rich regime. There are, however, special areas where the differences are larger. In particular, for the reddest colours, the metallicity estimates for the global and dwarf/giants metallicity calibrations can reach up to 0.4 and even 1.0. While it is indeed at the reddest colours that the stellar gravities differ the most between a giant and a dwarf star, and hence where one expects the largest impact on the metallicity estimate, we note that the training sample exhibits a severe lack of metal-poor main sequence stars (K dwarfs). As it is clearly visible in both the CMD and, most prominently, in the Pristine colour–colour space, our anticentre sample extends to the regions where metal-poor K dwarfs are expected to lie, yet, with hardly any training sample, casting doubts on the metallicities estimated for these stars. We therefore excluded from our sample stars with  $(g - i)_0$  higher than 1.3. We stress that this cut further discriminates against metal-rich stars.

The comparison of these photometric metallicities with spectroscopic  $[\text{Fe}/\text{H}]$  is displayed in Fig. 5 (panels a and b). The overall performance of the calibration with and without taking into account the luminosity of stars is rather similar, with a very slight decrease of the overall bias and dispersion of the retrieved photometric metallicities with our new calibration separating stars according to their luminosities. The overall behaviour is preserved, with a bias against the most metal-rich stars because of the cut imposed by the calibration at  $[\text{Fe}/\text{H}] > -0.2$ , and a slight underestimation of photometric metallicities at the lowest metallicities ( $[\text{Fe}/\text{H}] < -2$ ), where the mean offset and dispersion of the differences are the highest,  $0.32 \pm 0.43$ . One should, however, bear in mind that spectroscopic measurements for these very metal-poor stars are also less precise and certainly concur to inflate the dispersion in that regime.

There are, however, subtle differences between the calibrations with and without taking the stellar luminosity into account, which will play a role in the present investigation. In particular, around spectroscopic metallicities  $[\text{Fe}/\text{H}]_{\text{phot}}$  around  $-0.7$ , a regime where the spectroscopic metallicities are the most reliable, a number of stars deviate from the one-to-one relation and show photometric



**Figure 5.** 2D histograms showing the comparison between the [Fe/H] inferred spectroscopically (by the SDSS stellar parameters pipeline) and the ones inferred photometrically. Top panel (a): the adopted photometric [Fe/H] is the one based on the global calibration. Medium panel (b): the adopted photometric [Fe/H] takes into account the dwarf/giant classification. The two bottom panels (c) show the differences between both photometric calibrations, separately for giants and dwarfs. Mean differences and dispersions ( $\Delta[\text{Fe}/\text{H}] \pm \sigma$ ) are given for various metallicity intervals (defined by the vertical dashed lines).

metallicities in the range  $[\text{Fe}/\text{H}]_{\text{phot}} = -0.8$  to  $-1.5$  in the original calibration. This population of deviating stars is reduced with our new calibration. Panel (c) of Fig. 5 shows that this is the metallicity domain where the two calibrations differ most, both for giants where the new calibration gives slightly lower metallicities and for dwarfs where the new calibration gives higher metallicities. The dwarf stars in the training sample thus contribute to reduce the bias and dispersion observed between spectroscopic and photometric metallicities with our new calibration in the range  $[\text{Fe}/\text{H}]_{\text{phot}} = -0.8$  to  $-1.5$ . Since our anticentre sample comprises mostly dwarf stars (the ratio of dwarfs to giants is higher than in the training sample,

see Fig. 4), we expect this new calibration to be even better in that problematic metallicity range. There, however, remains a potential contamination of metal-rich stars when selecting low-metallicity targets based on photometric estimations that should be taken into consideration when interpreting our results. We will address the impact of such contamination in sub-Section 3.2.

### 2.2.2 Additional quality selection cuts

There are several caveats regarding the metallicity calibration based on this approach, which could potentially affect [Fe/H] estimates for some of our targets. For example, we detect a group of white dwarfs among our anticentre sample although this stellar type is not included in our training sample. White dwarfs lack Ca absorption lines and therefore they end up with an extremely low-metallicity estimate if we misinterpret them as main sequence or giant stars. We filter them out by removing targets with  $M_G < 10$ . In Youakim et al. (2020), it was shown that young stars are another case that may be prone to erroneous [Fe/H] estimate. Following Youakim et al. (2020), we reject targets with  $(g - i)_0 < 0.6$ ,  $(u - g)_0 > 1.15$  and  $(u - g)_0 > 1.5(g - i)_0 + 0.6$ . We finally reject stars in an unphysical location in the colour–colour space, as defined by the synthetic spectra computed with no metals (i.e. above the black curve in the left-hand panel of fig. 3 of S17).

As explained before, Pristine metallicities partially rely on broadband SDSS photometry obtained several years before the actual Pristine survey (which started in 2016). Therefore, any variability in the stellar flux may result in a biased derived metallicity. For that reason, we clean our sample of possible variable stars.

We use the standard deviation  $\sigma$  flux of the  $G$ -band fluxes (which have been obtained from 2014 to 2017) as an indicator of their variability. It is provided by the Gaia EDR3 archive<sup>5</sup> as the `phot_g_mean_flux_over_error` parameter, normalized to the square root of the number of observations contributing to the  $G$  mean flux,  $\sqrt{\text{phot\_g\_n\_obs}}$ :

$$\sigma \text{flux} = \frac{\sqrt{\text{phot\_g\_n\_obs}}}{\text{phot\_g\_mean\_flux\_over\_error}}. \quad (17)$$

We remove objects verifying the following empirically defined relation:

$$\sigma \text{flux} > \frac{(\sigma \text{flux} - 16)^2}{800} * \max(\sigma \text{flux} - 15, 0) + 0.015. \quad (18)$$

In order to avoid binary systems that could lead to wrong astrometric estimates, we use the Renormalized Unit Weight Error (RUWE) parameter provided by the *Gaia* archive, keeping only targets with a RUWE parameter lower than 1.4, as suggested by the Gaia Collaboration (Lindgren et al. 2020b).

Our final sample after applying all the selection cuts comprises 152 079 stars.

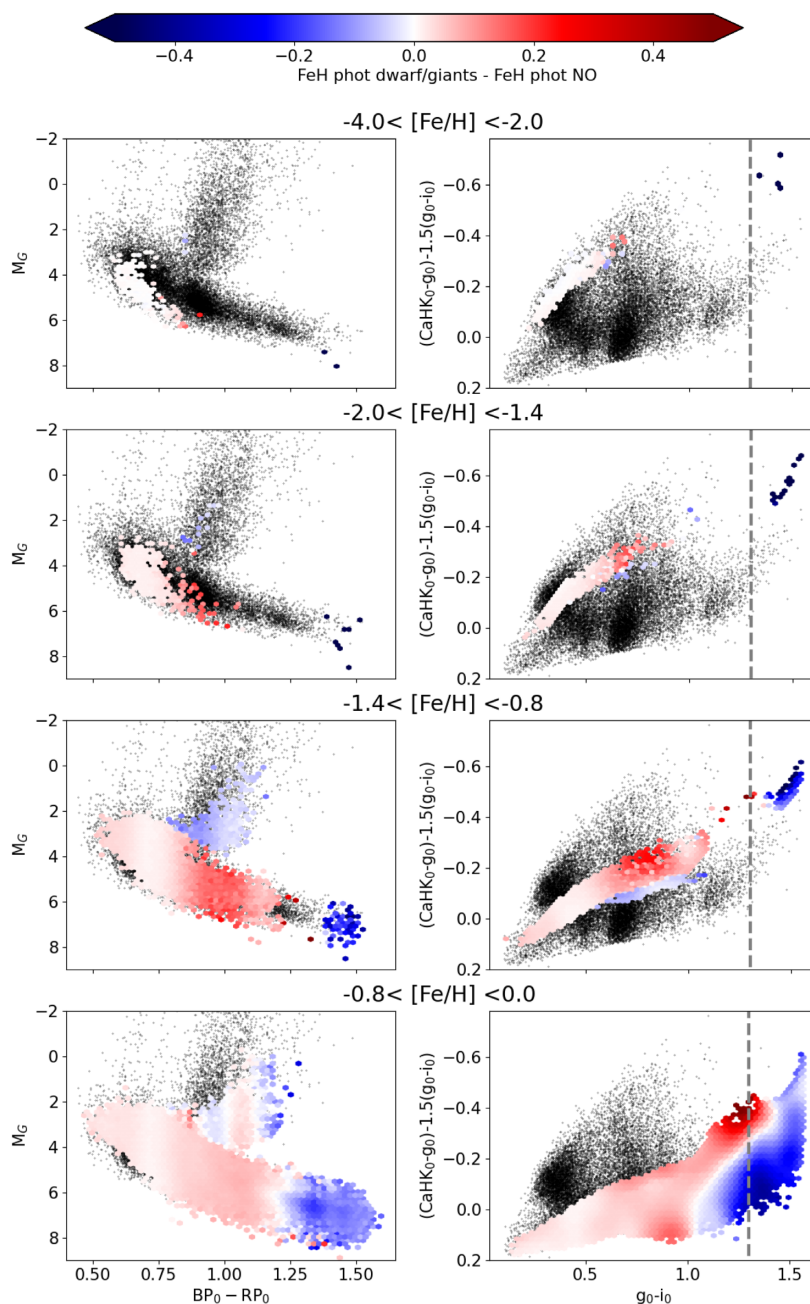
## 3 RESULTS

Our working sample covers galactocentric distances  $R$  from 8 to  $\sim 15$  kpc, with a peak at  $R \sim 9$  kpc and  $z \sim 0.5$  kpc, increasing in distance from the plane,  $z$ , as  $R$  increases, as shown in Fig. 7.

The  $V_\phi$  distribution is a good discriminant for Galactic stellar populations, whereas metallicity alone overlaps significantly between populations (e.g. the Galactic thin and thick discs, or the Galactic

<sup>5</sup><https://gea.esac.esa.int/archive/>





**Figure 6.** Left-hand panels: colour–magnitude diagram based on Gaia EDR3 photometry. The background black dots correspond to the training sample. The mean differences (calculated in hexagonal bins on the CMD) of the  $[\text{Fe}/\text{H}]$  estimates resulting from the calibration with the dwarf/giant classification and the global one for our anticentre Pristine sample (split in metallicity bins) are overlotted. Right-hand panels: Same as left-hand panels but in the colour–colour space (combining Pristine and SDSS photometry) used for the  $[\text{Fe}/\text{H}]$  calibration.

thick disc and halo). Thus, different Galactic populations should appear as distinct groups in the  $V_\phi$  (or, in our case,  $V_\ell$ ) versus  $[\text{Fe}/\text{H}]$  plane. With our  $V_\ell$  derived from the Gaia EDR3 proper motions (equations 1–6) and the photometric  $[\text{Fe}/\text{H}]$  in hands, we then proceed to examine the  $V_\ell$  distribution of Milky Way stars towards the anticentre as a function of metallicity.

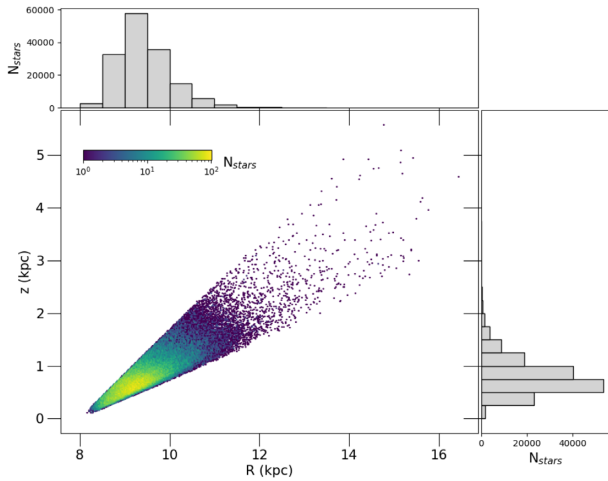
We will evaluate the  $V_\ell$  distributions in the following metallicity bins:  $[\text{Fe}/\text{H}] > -0.8$  (to include the region where we would expect to find the bulk of the thin disc),  $-1.5 < [\text{Fe}/\text{H}] < -0.8$  (the metallicity range where, based on the comparison with the spectroscopic  $[\text{Fe}/\text{H}]$  estimates, we see a large systematic deviation in the error function),  $-2 < [\text{Fe}/\text{H}] < -1.5$  (where the *metal-weak* thick disc has been

detected in previous work – Norris, Bessell & Pickles 1985b; Ruchti et al. 2011; Kordopatis et al. 2013a, among others), and  $-4 < [\text{Fe}/\text{H}] < -2$  (the most metal-poor stars in our sample). The typical mean  $V_\ell$  uncertainty, as derived from *Gaia* proper motions following the method presented in Section 2.1, is of 3, 7, 22, and 28  $\text{km s}^{-1}$ , respectively, in each metallicity range.

### 3.1 Examining the Galactic anticentre on the $V_\ell$ versus $[\text{Fe}/\text{H}]$ plane

Fig. 8 shows  $V_\ell$  as a function of  $[\text{Fe}/\text{H}]$  for our sample. We recall that negative  $V_\ell$  corresponds to prograde motion and positive  $V_\ell$  to





**Figure 7.** Distribution of the height above the plane,  $z$ , versus galactocentric radius,  $R$ , for our Pristine anticentre stellar sample, colour coded by the number of stars. Histograms at the top and at the right correspond to the  $R$  and  $z$  distributions, respectively.

retrograde orbits. The contour lines, showing the 33, 66, 98, 99, and 99.9 per cent of the cumulative distribution of the stars in this plane, highlight several substructures that we discuss below.

We see that the anticentre is dominated by metal-rich stars, with  $[\text{Fe}/\text{H}]$  higher than  $-1$ . The  $[\text{Fe}/\text{H}]$  and  $V_\ell$  distributions on the top and right-hand panels of Fig. 8 quantitatively verify that our sample is dominated by a chemical and kinematical thin disc, with a metallicity distribution function (MDF) and a  $V_\ell$  distribution that peak at  $-0.5$  and  $-225 \text{ km s}^{-1}$ , respectively. We will, however, refrain from analyzing too deeply this population, since we know that our sample is biased against the more metal-rich stars that would be part of this thin disc.

It is worth noting the intermediate-metallicity population ( $\sim -0.7$ ) extending from  $V_\ell \sim -200 \text{ km s}^{-1}$  to  $V_\ell \sim +100 \text{ km s}^{-1}$ , evidencing the presence of the thick disc and of the so-called *Splash* component (Belokurov et al. 2020a). The *Splash*, detected from Gaia DR2 data – see also Gaia Collaboration (2018b), Haywood et al. (2018), Fernández-Alvar et al. (2019), and Di Matteo et al. (2019) – is thought to be a stellar population comprising disc stars heated to halo-like kinematics likely because of the merger with the satellite galaxy Gaia–Enceladus–Sausage around 10 Gyr ago. The continuous overdensity of stars at  $[\text{Fe}/\text{H}] > -1$  extending towards increasing  $V_\ell$  resembles such a disc-heated population.

One of the most puzzling stellar substructures unveiled by this analysis is the continuous sequence of stars moving with thin disc-like  $V_\ell$ , not limited to the regime  $[\text{Fe}/\text{H}] > -1$  but extending down to extremely low metallicity values,  $[\text{Fe}/\text{H}] < -3$ . This rotating thin disc-like sequence is clearly detectable down to  $[\text{Fe}/\text{H}] \sim -3$ , as shown by the blue histogram on the top panel of Fig. 8.

The  $V_\ell$  distribution of the stars at  $[\text{Fe}/\text{H}] < -1.5$  is less conspicuous around thin/thick disc-like velocities, i.e.  $V_\ell \approx -150 \text{ km s}^{-1}$ , in favour of a less rotating and hotter halo-like distribution. At those metallicities, stellar velocities are centred around  $V_\ell \sim 0 \text{ km s}^{-1}$ , with a significant dispersion, of the order of  $70\text{--}90 \text{ km s}^{-1}$ . Interestingly, the contour line that contains 99 per cent of the sample in Fig. 8 shows a slightly overdense area centred at a  $V_\ell \sim 0 \text{ km s}^{-1}$  but with a narrower dispersion than the bulk of the halo stars, and it extends from the metallicities where the *Splash* is detected down to  $[\text{Fe}/\text{H}] \sim -2$ . This hints at a possible link between the halo (at low metallicities) and the *Splash* (at higher metallicities), which will be

further discussed in Section 4.2 – see Kordopatis et al. (2020) with a similar conclusion.

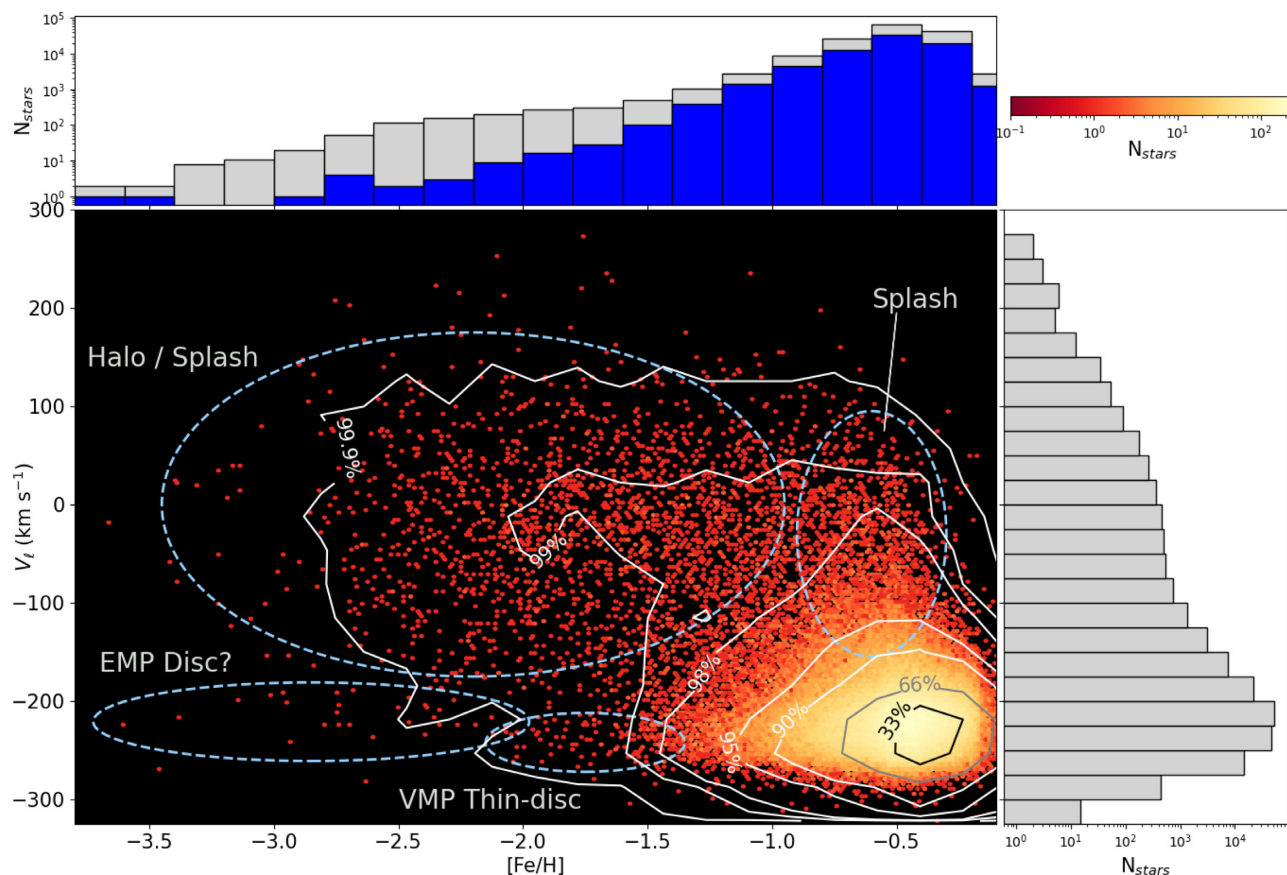
At low latitudes towards the anticentre, the vertical velocity component,  $V_z$ , is approximately  $V_b \cos(b)$  (see equations 9 and 12). We take advantage of this fact to evaluate how the different detected populations are characterized based on their vertical velocities. Fig. 9 shows the  $V_\ell$  versus  $[\text{Fe}/\text{H}]$  plane, colour-coded by  $V_b$ , split in bins of  $z$ . Before inspection, we recall that, because of our selection function, there is a correlation between  $z$  and  $R$ , i.e. stars at high  $z$  are also located at large  $R$  (see Fig. 7). Close to the plane, at  $z < 0.5 \text{ kpc}$ , the dominating population is the disc, with metallicities higher than  $-1$ , prograde orbits, and low vertical velocities. In lower proportions, there are signatures of the halo and the *Splash*. The number of retrograde stars, however, is very low and none with large  $+V_\ell$ . It is worth noting the presence of fast-rotating very metal-poor stars down to  $[\text{Fe}/\text{H}] < -3$  showing very low  $V_b \cos(b)$ , suggesting that these stars remain close to the plane. At  $0.5 < z < 1 \text{ kpc}$ , where most of our sample is located, the fraction of non-rotating stars increases and so does the fraction of retrograde stars. The thin disc tail towards very low-metallicity values is also clearly detectable, with vertical velocities close to  $V_b \cos(b) \sim 0$ . As  $z$  increases, this very low-metallicity thin disc tends to disappear, and those stars with  $[\text{Fe}/\text{H}] < -2$  moving in prograde orbits have also high vertical velocities, suggesting that they are the prograde tail of the velocity distribution of the halo.

The presence of a disc and a halo population is clear in stars with metallicities from  $-2$  to  $-0.2$ , up to  $z \sim 2$ . At  $z > 2 \text{ kpc}$  (and  $R > 11 \text{ kpc}$ ), stars show the more classical  $V_\ell$  distribution of a kinematical halo centred at  $0 \text{ km s}^{-1}$  with a large dispersion at low metallicities and disc stars located at high metallicities, with no clear sign of very low-metallicity disc or high-metallicity *Splash*. The anticentre is close to the line of nodes of the Galactic warp (López-Corredoira et al. 2002; Yusifov 2004; Momany et al. 2006; Reylé et al. 2009; Chen et al. 2019; Skowron et al. 2019) and, thus, the presence of disc stars at high  $z$  due to this feature is not expected. Besides, the amplitude of the warp was always found lower than  $2 \text{ kpc}$  at  $R < 14 \text{ kpc}$  – see Cheng et al. (2020) and references therein. Furthermore, it is not expected to have a gradient of more than a few  $\text{km s}^{-1}$  ( $5\text{--}6$  between  $8$  and  $14 \text{ kpc}$ ) in vertical velocities and no variation in the rotational velocity – see fig. 3 in Poggio et al. (2020). Consequently, the effect of the warp on our spatial and kinematical distribution is negligible.

Fig. 10 shows the  $V_\ell$  distribution, now split in low- and high-metallicity stars, with  $[\text{Fe}/\text{H}] < -1.5$  and  $[\text{Fe}/\text{H}] > -1.5$ , respectively, and in high and low vertical velocities, with  $|V_b \cos(b)| < 50 \text{ km s}^{-1}$  and  $|V_b \cos(b)| > 50 \text{ km s}^{-1}$ , respectively. We choose  $50 \text{ km s}^{-1}$  as the boundary to distinguish between disc and halo stars because this quantity is approximately  $2\sigma$  the dispersion of the thin disc vertical motion, i.e.  $\sigma = 25 \text{ km s}^{-1}$  (Gaia Collaboration 2021a). A vertical red dashed line marks the  $V_\ell$  value lower than which stars move with thin disc-like rotation. This figure clearly shows that stars with  $[\text{Fe}/\text{H}] < -1.5$  rotating like the thin disc also have low vertical velocities, and there are almost no such kind of stars with high vertical motions.

### 3.2 Contamination of metal-rich stars at low metallicities

Even though the comparison of photometric and spectroscopic estimates is overall very good, we expect, nevertheless, our low-metallicity sample to be contaminated by some higher metallicity stars that have been badly calibrated. Because metal-rich stars outnumber metal-poor stars, even a low percentage of failures can



**Figure 8.**  $V_\ell$  as a function of  $[\text{Fe}/\text{H}]$ , colour coded by density. The density is measured in bins of  $10 \text{ km s}^{-1}$  and  $0.1 \text{ dex}$ . Contour lines show the 33, 66, 98, 99, and 99.9 per cent of the cumulative distribution. The most evident stellar substructures are pointed out with annotated blue ellipses. These are: a presumed extremely metal-poor disc (EMP disc), the very metal-poor thin disc (VMP thin disc), the *Splash*, and the halo (with the likely contribution of metal-poor *Splash* stars). Stars on prograde motion are those with  $V_\ell < 0$ , while stars on retrograde motion have  $V_\ell > 0$ . Top panel: Metallicity distribution function of all the anticentre sample in light grey and of the fast rotators ( $V_\ell < -220 \text{ km s}^{-1}$ ) in blue. Right-hand panel:  $V_\ell$  distribution of all the anticentre sample. Both distributions are plotted in logarithmic scale. Note: the colour code lower limit is set to 0.1 due to colour contrast purposes.

create a significant contamination at the lowest metallicity bins. It is therefore crucial to estimate this contamination for the correct interpretation of the properties of very metal-poor stars.

In particular, it is important to verify whether the sequence of fast-rotating metal-poor stars is not, in fact, metal-rich contaminants with an underestimated photometric  $[\text{Fe}/\text{H}]$ . We recall that  $V_\ell$  has a very small uncertainty, and therefore this disc-like sequence cannot be the result of a low accuracy in  $V_\ell$  estimation. To investigate the impact of the photometric  $[\text{Fe}/\text{H}]$  uncertainties on this sequence, we model a thin disc-like MDF and convolve it with the error function of Pristine’s photometric metallicities.

We decide to model the thin disc MDF as a single Gaussian, with a mean of  $-0.15$  and a sigma of  $0.22$ . We obtain these values from Hayden et al. (2014) based on the  $(R, z)$  location of the stars within our volume. We derive the error function of the Pristine  $[\text{Fe}/\text{H}]$  metallicities from the differences between the photometric and spectroscopic  $[\text{Fe}/\text{H}]$  estimates for stars in the range  $-0.2 < [\text{Fe}/\text{H}]_{\text{spec}} < 0$ . Then, we convolve the MDF by the error function, as shown in Fig. 11.

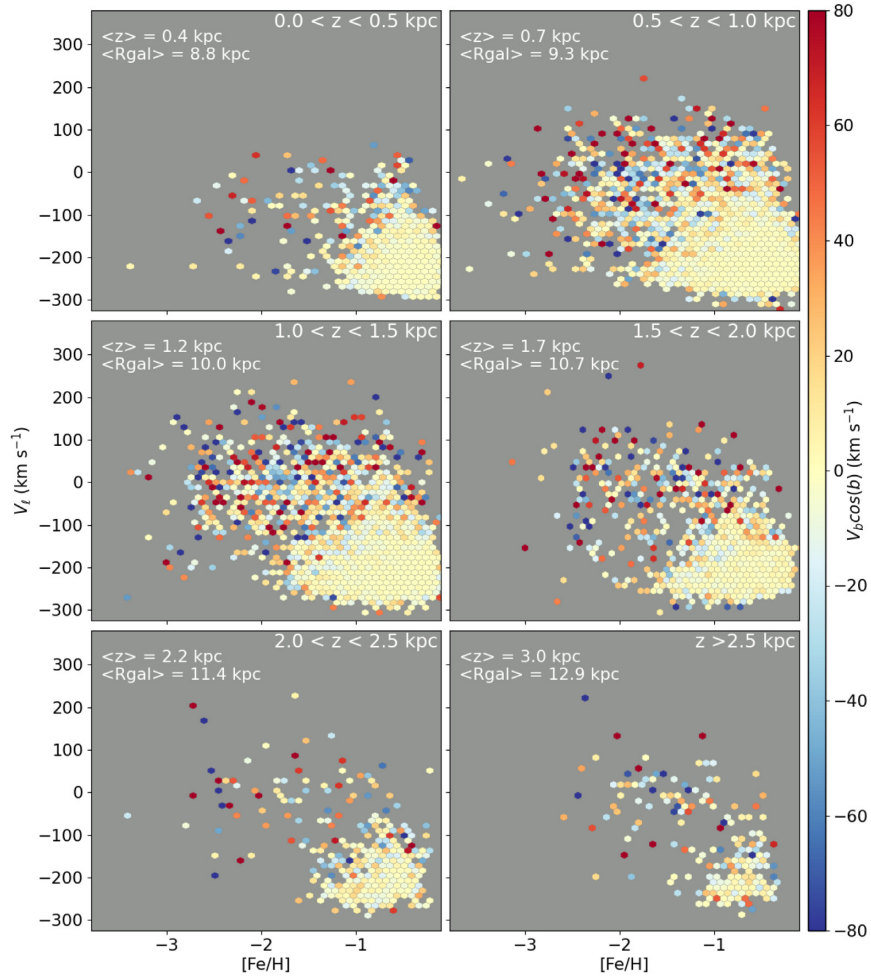
We estimate the number of thin disc contaminants present in the metallicity range where no thin disc stars are expected, i.e. lower than  $\sim -0.7$  (Bensby, Feltzing & Oey 2014) by scaling our toy model to the total number of stars moving with thin disc-like kinematics present in our sample. We obtain the total number of kinematical thin

disc stars by fitting our whole anticentre sample with a mixture of Gaussians using a Gaussian Mixture Model.<sup>6</sup> This model relies on a clustering technique to calculate the probability of a star to belong to a particular data cluster for a pre-determined number of clusters and a convergence threshold. We obtain the number of Gaussians that best account for the data evaluating the fits that result from considering Gaussian Mixture models comprising one up to six clusters without any priors. The range of considered Gaussians is an educated guess based on Fig. 8 (we expect that all of the dominant populations of the Galaxy are present in our sample). The best fit is chosen by minimizing the Bayesian Information Criteria (BIC). Then, we infer the weight of the Gaussian centred on the value closer to  $-238 \text{ km s}^{-1}$  ( $V_\phi$  of the LSR). We find that 78 per cent of our anticentre stars move with thin disc like  $V_\ell$ .

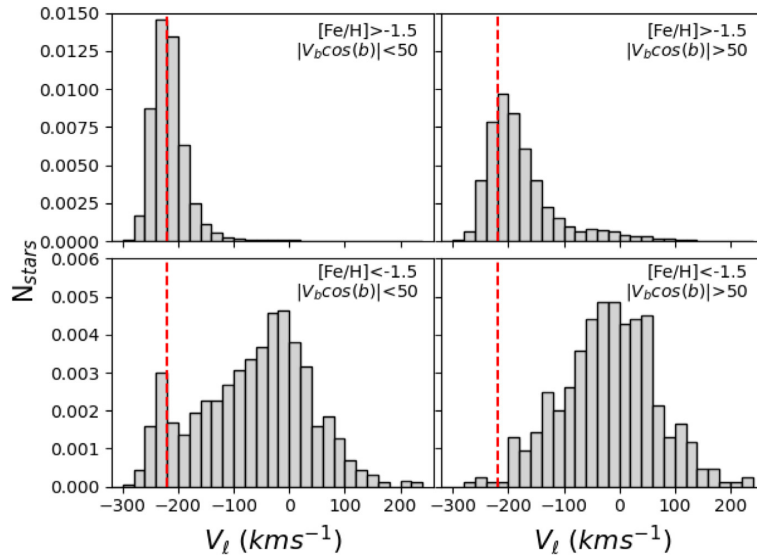
Based on our conservative toy model,<sup>7</sup> we calculate the number of contaminants in each of the metallicity bins. At metallicities  $-2 < [\text{Fe}/\text{H}] < -1.5$ , and  $-1.5 < [\text{Fe}/\text{H}] < -0.8$ , we find that the contamination of metal-rich stars due to the  $[\text{Fe}/\text{H}]$  error function

<sup>6</sup>We make use of the PYTHON module `sklearn.mixture` (Pedregosa et al. 2011).

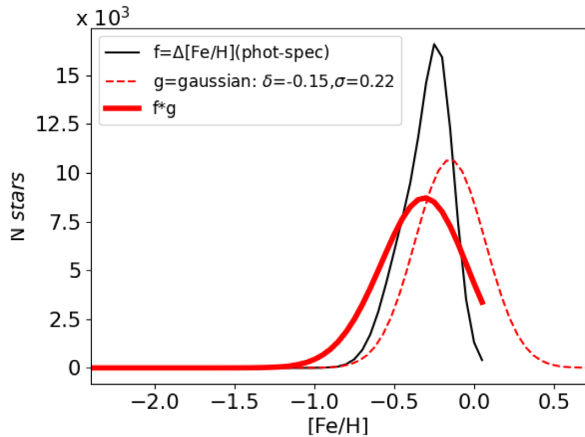
<sup>7</sup>A more complex MDF, with an asymmetric low-metallicity tail, has also been investigated, resulting to fewer contaminants at low metallicities. The simpler MDF we therefore adopt is justified as being a conservative estimate of the contamination.



**Figure 9.**  $V_l$  as a function of  $[\text{Fe}/\text{H}]$ , colour coded by  $V_b \cos(b)$ , in bins of  $z$ . The colour code is computed as the mean  $V_b \cos(b)$  in hexagonal bins, and the mean height above the plane  $\langle z \rangle$  and galactocentric radius  $\langle R_{\text{gal}} \rangle$  is given in each panel.



**Figure 10.**  $V_l$  distribution for stars with  $[\text{Fe}/\text{H}] > -1.5$  (top panels) and with  $[\text{Fe}/\text{H}] < -1.5$  (bottom panels), and with low vertical velocities  $|V_b \cos(b)| < 50$   $\text{km s}^{-1}$  (left-hand panels) and high vertical velocities  $|V_b \cos(b)| > 50$   $\text{km s}^{-1}$  (right-hand panels). Velocities lower than  $V_l = -220$   $\text{km s}^{-1}$  (the red dashed line) are those characteristic of thin disc stars.



**Figure 11.** This figure illustrates the metallicities that Pristine is expected to recover (red continuous line) if we feed the survey with solely thin disc targets distributed according to a Gaussian MDF centred at  $-0.15$  with a sigma of  $0.22$  (red dashed line). The transformation is obtained by convolving this MDF with an error function (in black), derived from targets in the range  $[\text{Fe}/\text{H}]_{\text{spec}} = [-0.2, 0]$  of Fig. 5.

is of 4 and 4045, respectively. The number of stars in our sample with velocities compatible with the thin disc at these two metallicity ranges is 176 and 9316, respectively. It is expected, then, that less than 1 per cent of stars are metal-rich contaminants in the range  $-2 < [\text{Fe}/\text{H}] < -1.5$ , and around 45 percent between  $-1.5 < [\text{Fe}/\text{H}] < -0.8$ . Therefore, we conclude that, at low metallicities, the population of stars rotating with thin disc  $V_\ell$  cannot be explained solely by the contamination from higher metallicity stars. In the next section, we will evaluate the statistical significance of this scarce population.

### 3.3 Statistical significance of the observed stellar substructures

We proceed to quantitatively evaluate how the  $V_\ell$  distribution changes as a function of the metallicity and to interpret such distributions in terms of underlying stellar populations. For this purpose, we use Gaussian Mixture Models and evaluate models combining from one up to six Gaussians, in order to derive the number of components that best fit the  $V_\ell$  distributions in each metallicity range.

Previous works have shown that for the thin and thick discs the shape of the  $V_\phi$  distribution is not Gaussian (e.g. Sharma & Bland-Hawthorn 2013, and references therein). However, at first order it is a good approximation and simple to model. We will discuss our results in this context, keeping in mind the non-Gaussianity of the true distributions.

The results are shown in Figs 12–14. We plot the Gaussian mixture fit for which we obtained the minimum BIC, along with the corresponding individual Gaussians in dotted lines and the position of their means in dashed vertical lines. In addition, we show with vertical arrows the typical  $V_\phi$  values for the thin and thick discs and the halo (red, green, and blue, respectively):  $-238 \text{ km s}^{-1}$  for the thin disc (our adopted LSR value),  $-208 \text{ km s}^{-1}$ , assuming a  $30 \text{ km s}^{-1}$  lag for the thick disc with respect to the thin disc (Recio-Blanco et al. 2014), and  $0 \text{ km s}^{-1}$  for the halo (Fermani & Schönrich 2013). We also plot with thinner longer arrows the expected mean  $V_\phi$  values for the thin and thick discs in each metallicity range, taking into account the correlation of  $V_\phi$  with metallicity measured by Kordopatis et al. (2017) – see also Spagna et al. (2010), Kordopatis et al. (2011), Lee et al. (2011), Adibekyan et al. (2013), Allende Prieto, Kawata &

Cropper (2016), and Re Fiorentin, Lattanzi & Spagna (2019), i.e.  $-12 \pm 3 \text{ km s}^{-1} \text{ dex}^{-1}$  for the thin disc and  $+47 \pm 10 \text{ km s}^{-1} \text{ dex}^{-1}$  for the thick disc. The underlying histograms depict the observed  $V_\ell$  distributions in bins of  $10 \text{ km s}^{-1}$  for each corresponding metallicity range.

#### 3.3.1 Stellar substructures at high metallicities, $[\text{Fe}/\text{H}] > -1.5$ : the thin disc, the thick disc, and the Splash

The panels in Fig. 12 show the best Gaussian mixture fit for the two most metal-rich ranges:  $-0.8 < [\text{Fe}/\text{H}] < -0.2$  and  $-1.5 < [\text{Fe}/\text{H}] < -0.8$ . The models that best fit the data are a mix of 5 and 4 Gaussians, respectively. At the top right corner of each panel, we plot the BIC differences between the other models and the best fit. BIC differences larger than 2 provide positive evidence that the model with the lower BIC is indeed the preferred fit, while differences larger than 10 imply strong and robust evidence (Raftery 1995). The resulting BIC differences in our Gaussian Mixture models are overall higher than 10, confirming the robustness of the results.

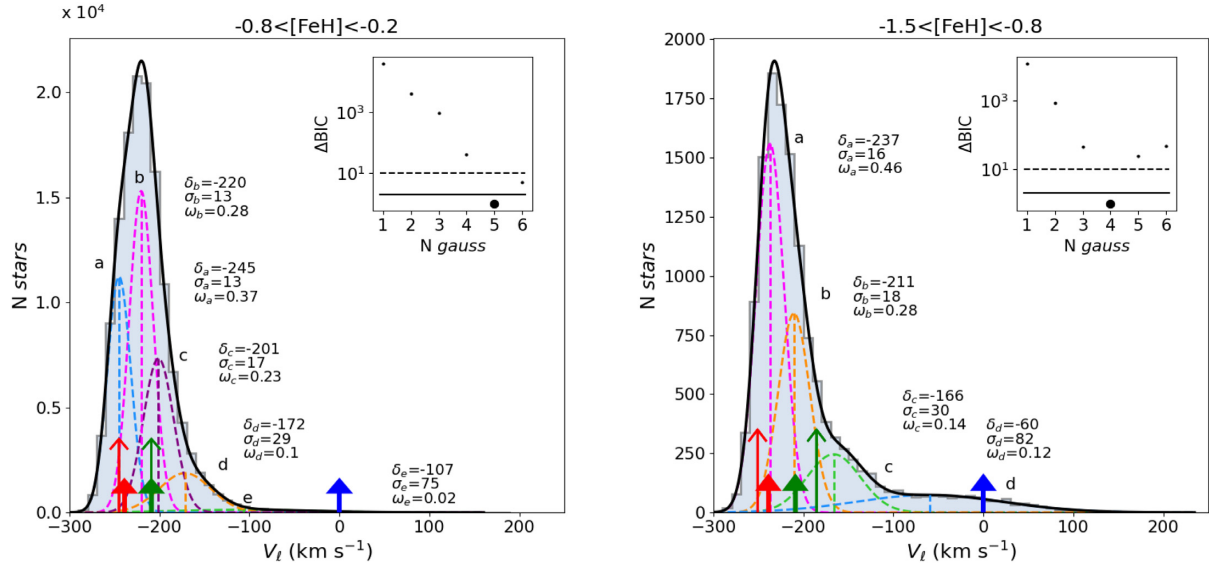
In order to interpret the resulting Gaussian mixtures, we compare the centres of the individual Gaussians with the  $V_\ell$  values at which we would expect to find the mean thin disc and thick disc distributions (red and green arrows, respectively; see beginning of Section 3.3). Looking at the most metal-rich bin,  $-0.8 < [\text{Fe}/\text{H}] < -0.2$ , we verify that Gaussian *a* is centred at the mean thin disc  $V_\phi$  value considering the correlation with  $[\text{Fe}/\text{H}]$ . The mean of Gaussians *b* and *c* is close to the expected  $V_\phi$  for the thick disc. The other two Gaussians, *d* and *e*, show a significant lag with respect to the thick disc, with mean  $V_\ell$  values of  $-172$  and  $-107 \text{ km s}^{-1}$ , respectively. Taking into account that the thin disc and thick disc  $V_\phi$  distributions are not Gaussian, Gaussian *b* likely comprises the lower rotation tail of the thin disc and a fraction of thick disc stars, and Gaussian *d* accounts for the lower rotation thick disc and some of the *Splash*.

The same applies for the metallicity range  $-1.5 < [\text{Fe}/\text{H}] < -0.8$ , with the fast-rotating Gaussian *a* indicating the presence of the thin disc, Gaussian *b* the thick disc, and Gaussian *c* likely containing contribution from both the thick disc and the *Splash*. Both the thick disc and the *Splash* increase in proportion as the metallicity decreases.

The mean value of each Gaussian shifts to lower rotation values (higher  $V_\ell$ ) with decreasing metallicity. For the thick disc, this is in line with what found in Kordopatis et al. (2017). However, the thin disc also decreases its rotation as metallicity decreases, contrary to what is measured at higher metallicities,  $-0.5 < [\text{Fe}/\text{H}] < 0.5$  – see Kordopatis et al. (2017) and Re Fiorentin et al. (2019).

In order to assess the robustness of this result, we re-evaluate  $V_\ell$  from 100 Monte Carlo realizations of the parallax and proper motions, based on their associated uncertainties. This analysis confirms our results, i.e. that at high metallicities,  $[\text{Fe}/\text{H}] > -1.5$ , the thin disc is the dominating population, but the  $V_\ell$  distribution requires at least two additional components. The best Gaussian mixture models that fit the data reveal, on the one hand, that the thick disc is clearly present in the anticentre at such metallicity values and, on the other hand, that there is significant presence of a stellar component of much smaller weight and a very wide  $V_\ell$  dispersion that describes well the characteristics of the *Splash*, as shown by Belokurov et al. (2020b). They found slightly lower mean rotation and dispersion values ( $V_\phi = 25 \text{ km s}^{-1}$  and  $\sigma(V_\phi) = 54 \pm 11 \text{ km s}^{-1}$ , but their analysis was performed at higher distances from the plane ( $2 < |z| < 3 \text{ kpc}$ ), where only the stars heated to higher energies can reach.





**Figure 12.** Best Gaussian Mixture Model fits of the  $V_\ell$  values for stars with  $-0.8 < [\text{Fe}/\text{H}] < -0.2$  (left-hand panel) and  $-1.5 < [\text{Fe}/\text{H}] < -0.8$  (right-hand panel) are plotted with continuous black lines. The corresponding individual Gaussians are in coloured dashed lines, together with their means, widths, and relative weights annotated next to them. At the top right corner, we show the differences between the BIC of each model and the one with the best fit, as well as the reference values of statistical significance for BIC differences of 10 (dashed black line) and 2 (continuous black line).

### 3.3.2 Stellar substructures at low metallicities, $[\text{Fe}/\text{H}] < -1.5$ : the extremely metal-poor disc and the halo

Fitting the data at  $-2 < [\text{Fe}/\text{H}] < -1.5$  returns as the optimal model a mix of two Gaussians. Fig. 13 shows that their means are centred on  $V_\ell$  values close to: (a) the LSR, and (b) the halo. As done for the more metal-rich stars, we evaluate the robustness of these results by performing 1000 Monte Carlo simulations.

The right top panel of Fig. 13 shows the distribution of the number of components of the model with the minimum BIC for each of the 1000 Monte Carlo simulations. For most of the realizations, the best model is a mix of two Gaussians, although there are some cases where it is a mix of three and even a mix of four. For the realizations for which the number of components of the best-fitting model is the same, we calculated the mean centre and  $\sigma$  of each of the Gaussian components and display the Gaussians inferred from these means in the panels below. As in the previous plots, we also plot in red, green, and blue arrows the expected rotational velocities of the thin and thick discs and the halo.

At the top-right corner of each plot, we evaluate the significance of the best-fitting model relative to the other ones by displaying the mean BIC differences between the first and formers, as well as their standard deviations as error bars.

We see from the BIC differences that when the best fit is composed by two Gaussians, then it is robustly the preferred model. For realizations for which the best fit is composed of three or four Gaussians, the BIC difference with the model of two Gaussians is close to 2, which implies that three or four Gaussian models are not significantly better than two-Gaussian model even if they show the minimum BIC. Also notable is the fact that the mean of Gaussian *b* in the case of three Gaussians is very poorly defined, showing a huge dispersion. This gives strong support to the fact that, taking into account the uncertainties in our measurements, the two-Gaussian model is the best fit for the data.

Interestingly, the two Gaussians that compose this model are: (1) one dominating the distribution, centred at low  $V_\ell$  close to 0 and showing a large dispersion, that we associate to the halo; and (2) one

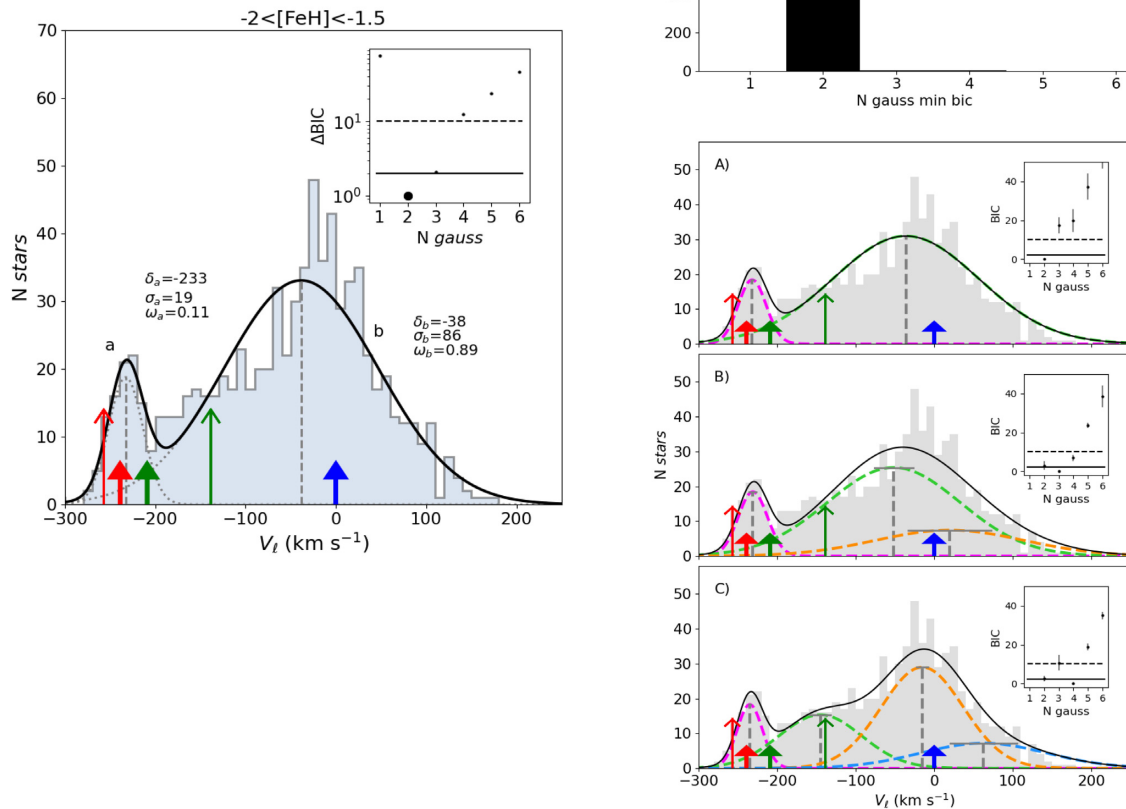
centred at, approximately, the typical  $V_\ell$  value of the LSR that we refer to as the metal-poor rotating stars.

It is also remarkable that this second Gaussian is always present even in the cases where the best fit comprises three or four Gaussians. In these cases, it is the Gaussian associated to the kinematical halo that splits into several components (a more prograde thick disc-like one and another more retrograde).

In summary, the evidence for the presence of a kinematical thin disc-like component is significantly robust at metallicities  $-2 < [\text{Fe}/\text{H}] < -1.5$ . We recall that the expected contamination of metal-rich stars at this metallicity range due to Pristine  $[\text{Fe}/\text{H}]$  error function is less than 1 per cent, based on our estimates (see Section 3.2). Obviously, this metal-rich contamination cannot account for the kinematical thin disc component detected. Also noticeable is the fact that this component does not sharply disappear at  $[\text{Fe}/\text{H}] = -2$  but extends at very low values down to  $[\text{Fe}/\text{H}] \sim -3.5$ .

On the other hand, our stellar sample at such low metallicities is dominated by a prograde halo (see Fig. 13). This component could be comprised by the sum of other substructures, like a thick disc and even a retrograde stellar population. However, the total number of stars in this component is too low to draw such conclusions with good statistical significance.

We perform the same exercise for stars at  $-4 < [\text{Fe}/\text{H}] < -2$ , obtaining that the data are well fit with a single Gaussian centred at  $-37 \text{ km s}^{-1}$  (see Fig. 14). Instead of a Monte Carlo, due to the low number of stars in this metallicity range, we applied a bootstrap technique, i.e. we implicitly consider that the uncertainties on  $V_\ell$  are smaller than the impact of low-number statistics on the sample realization. We in fact verified that the variance of the resulting Gaussian mixture is higher from the bootstrap technique than from the Monte Carlo on  $V_\ell$  uncertainties. We perform 1000 realizations, selecting each time a random sample of the same size of our data set at this metallicity range (141), and evaluating several Gaussian Mixture Models (comprising from one to six Gaussian components as for the previous metallicity ranges) for each of the 1000 realizations.



**Figure 13.** Left-hand panel: same as Fig. 12 but for the metallicity range  $-2 < [\text{Fe}/\text{H}] < -1.5$ . Right-hand panel: top plot shows the distribution of the number of components with the best fit obtained from 1000 Monte Carlo realizations; the bottom panels show the mean Gaussians when this best fit was a two-Gaussian model (panel A), three-Gaussian model (panel B), or four-Gaussian model (panel C). Horizontal lines above the Gaussians correspond to uncertainty on the centres of the distributions. On the right top corner of these panels, the mean difference of the BIC between each model and the best fit for the 1000 Monte Carlo realizations are plotted, as well as the standard deviations as error bars. The reference values of statistical significance for BIC differences of 2 (dashed black line) and 10 (continuous black line) are also indicated.

The bootstrapping technique confirms the previous result with 80 per cent of the realizations returning a single Gaussian as the best fit. The other 20 per cent fit the distribution with two, three, and even five Gaussians. As for the previous metallicity range, when comparing the BIC differences of the best-fitting model with the other models, the single Gaussian component is robustly the best fit when it has the minimum BIC. Noticeably, all the best-fitting models suggest that the component closest to the kinematical halo, i.e. with a mean  $V_\ell$  close to  $0 \text{ km s}^{-1}$ , has always a net prograde rotation. In the cases where the best fit is comprised by two components or more, there is always an additional more prograde component, revealing that the  $V_\ell$  distribution at these very low metallicities is shifted towards a net prograde rotation.

The extension of metal-poor rotating stars to the extremely metal-poor regime is not statistically identified as a separate Gaussian component (independently of the number of components), even though we do see it as a small excess around  $V_\ell \sim -230 \text{ km s}^{-1}$ .

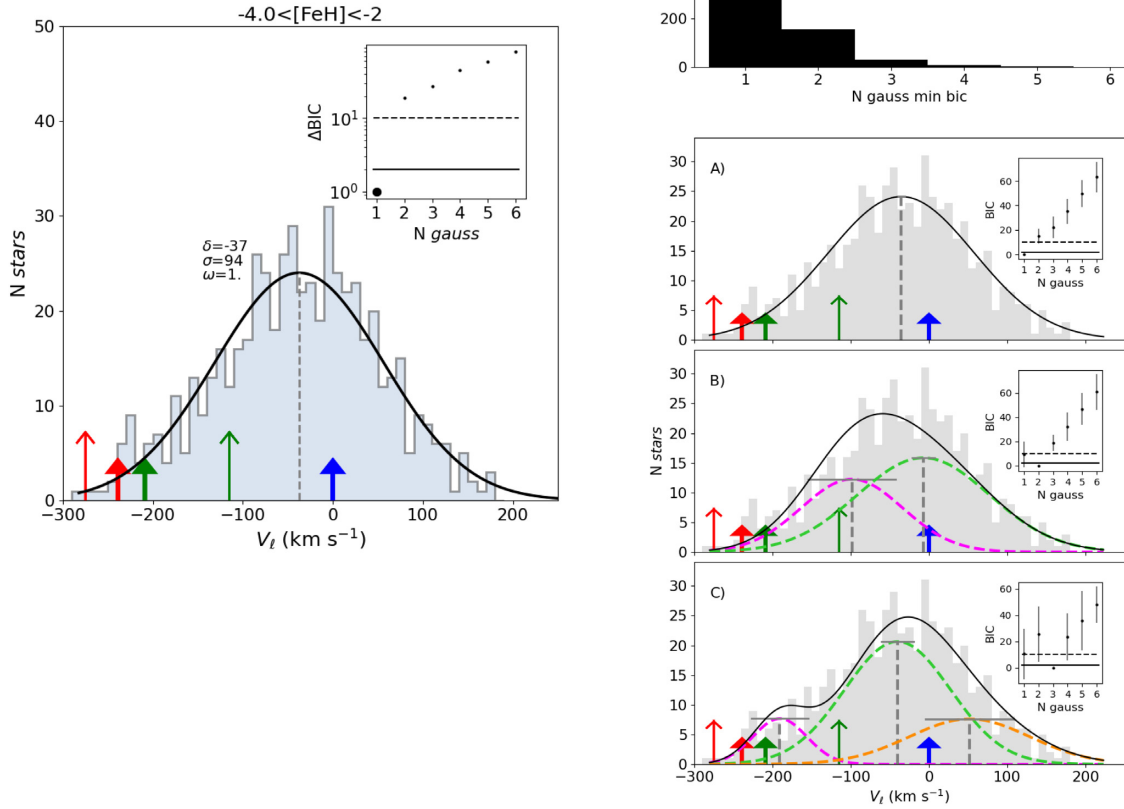
Finally, we present in Table 1 the different best-fitting Gaussians with their corresponding mean  $\langle V_\ell \rangle$ , standard deviations  $\sigma$ , and relative fractions per cent along with the Galactic populations with which we associate each of them.

## 4 DISCUSSION

In the present work, we offer an analysis of the distribution of kinematics along a wide metallicity range ( $-3.5 < [\text{Fe}/\text{H}] < -0.2$ ), from the analysis of the Pristine survey, a unique sample in terms of metallicity estimates (homogeneously derived), especially at the low-metallicity end. In this section, we discuss the implications of our findings in the context of Galaxy formation and the novelties that they uncover compared to previous analyses found in the literature.

### 4.1 Implications on the formation of the very and extremely metal-poor thin disc

Our findings strongly suggest the presence of a kinematical thin disc population down to  $[\text{Fe}/\text{H}] \sim -2$ . This population, at low metallicities, rotates significantly faster than the present-day thick disc, with velocities closer to the thin disc. Indeed, the  $V_\ell$  that we measure for this population are lower than the ones displayed by the thick disc detected at higher metallicities. There seems to be a continuous distribution of stars moving in thin disc-like orbits from very low ( $[\text{Fe}/\text{H}] \sim -2$ ) up to the highest metallicities. There are fast-rotating stars at even lower metallicities, down to  $[\text{Fe}/\text{H}] \sim$



**Figure 14.** Same as Fig. 13. In this case, the individual Gaussians were computed from the realizations with a best fit of one-Gaussian model (panel A), a two-Gaussian model (panel B), or a three-Gaussian model (panel C).

**Table 1.** Mean, standard deviation, and relative fraction of the Gaussian components that best fit the  $V_\ell$  distributions of Pristine stars towards the anticentre. We specify on the first column the Galactic component that we associate to each Gaussian based on how they compare to the literature values for the thin/thick discs and the halo. Since the velocity distributions of the Galactic components are not Gaussians, some of the resulting fitted Gaussians include stars of the non-Gaussian tail of two Galactic components. We indicate these cases by associating the corresponding Gaussian to more than one component.

	$-0.8 < [\text{Fe}/\text{H}] < -0.2$			$-1.5 < [\text{Fe}/\text{H}] < -0.8$			$-2.0 < [\text{Fe}/\text{H}] < -1.5$			$-4.0 < [\text{Fe}/\text{H}] < -2.0$		
	$\langle V_\ell \rangle$	$\sigma$	%	$\langle V_\ell \rangle$	$\sigma$	%	$\langle V_\ell \rangle$	$\sigma$	%	$\langle V_\ell \rangle$	$\sigma$	%
	(km s $^{-1}$ )			(km s $^{-1}$ )			(km s $^{-1}$ )			(km s $^{-1}$ )		
Thin disc	-245	13	37	-237	16	46	-233	19	11			
Thin disc + Thick disc	-220	13	28	-211	18	28						
Thick disc	-201	17	23	-166	30	14						
Thick disc + Halo/Splash	-172	29	10									
Halo/Splash	-107	75	2	-60	82	12	-38	86	89	-37	94	100

$-3.5$ , observed close to the plane and with low vertical motions. However, we did not find statistical evidence that these stars belong to a separate stellar population, linked to the thin disc-like one detected at higher metallicities, rather than the prograde rotating tail of the halo distribution centred at  $V_\ell$ , with a dispersion of around 100 km/s. In any case, it is striking to see the asymmetric shape of the  $V_\ell$  distribution for those stars at  $[\text{Fe}/\text{H}] > -2$ , with more fast rotators in the prograde rather than in the retrograde sense.

Extremely ( $[\text{Fe}/\text{H}] < -3$ ) and even ultra ( $[\text{Fe}/\text{H}] < -4$ ) metal-poor stars moving in disc-like orbits were recently discovered by Sestito et al. (2019), and, since then, several works have confirmed

their existence (Di Matteo et al. 2019, 2020; Sestito et al. 2020; Venn et al. 2020; Carter et al. 2021; Cordoni et al. 2021). From a full orbital characterization of all known ultra metal-poor stars with radial velocities, Sestito et al. (2019) found a large fraction of them confined to  $z_{\text{max}} < 3$  kpc to the plane (26 percent) and two of them (5 percent) moving in prograde circular motions. Sestito et al. (2020) also discovered, among stars with  $[\text{Fe}/\text{H}] < -2$  from the Pristine and LAMOST surveys, a significant number of them moving with thick disc-like orbits. In addition, Di Matteo et al. (2020), comparing stars from ESO’s Large Program ‘First Stars’ with other stellar samples, also discovered a subsample moving on thick disc-like orbits, i.e. sharing the locus of the thick disc in the

Toomre diagram. Comparing with other stellar samples covering all metallicity regimes from  $[\text{Fe}/\text{H}] < -4$  up to  $[\text{Fe}/\text{H}] > -0.5$ , they showed that, at every metallicity range, there is a kinematical disc and a halo population coexisting. In the present work, we confirm from a homogeneous sample that this is true, at least down to  $[\text{Fe}/\text{H}] \sim -2$ , except that our kinematical disc rotates more like a thin disc than a thick disc. The fact that our sample is mostly confined at low  $z$  implies that we probe this thin disc population preferentially compared to other studies.

Based on chemical evolution models that consider a disc formed in an inside-out fashion (Larson 1976; Cole et al. 2000), the outer regions form stars at a slower rate, due to a delayed accretion of gas at these radii. This gas dilutes the Galactic interstellar medium and decreases its metallicity. Thus, subsequent stars formed in the outer regions are more metal-poor than stars born in the inner disc at the same epoch, and, in addition, by that time the gas there may have settled already into a disc configuration. Inside-out chemical evolution models such the two-infall model (Chiappini, Matteucci & Gratton 1997; Chiappini, Matteucci & Romano 2001) have been applied to explain the age metallicity, age- $[\alpha/\text{Fe}]$ , the  $[\alpha/\text{Fe}]$  versus  $[\text{Fe}/\text{H}]$  dichotomy, or chemical abundance trends in Galactic stellar populations by several groups (Minchev, Chiappini & Martig 2013, 2016; Minchev et al. 2017; Grisoni et al. 2019, 2020a, b, c; Spitoni et al. 2019, 2020). Interestingly, Spitoni et al. (2021) reproduce a decrease of around 1 dex in metallicity at radius between 8 and 10 kpc, and their predicted  $[\text{Fe}/\text{H}]$  distributions show a metal-poor tail down to at least  $-2$ . Our disc stars down to  $[\text{Fe}/\text{H}] \sim -2$  would have formed, then, after this second infall of gas into the Galaxy. This scenario would explain their thin disc-like motion. It is, however, difficult to explain the existence of stars at  $[\text{Fe}/\text{H}] < -2$ . For these EMP stars, the formation in the early disc seems more feasible. Further analysis of the chemical abundances of these stars is needed, for instance, regarding the  $[\alpha/\text{Fe}]$  ratio to clarify whether there is a link between fast-rotating stars at metallicities higher and lower than  $[\text{Fe}/\text{H}] \sim -2$ .

Stars with metallicities  $[\text{Fe}/\text{H}] < -2$  are expected to be formed during the first 2–3 Gyr after the big bang (El-Badry et al. 2018), and at this epoch the Galaxy was still assembling through the hierarchical merger of smaller systems. Under this scenario, stars either accreted or born *in situ* in a proto-disc were expected to be pressure supported from the very beginning of their formation or heated into halo-like kinematics due to the impact of the mergers. Thus, finding stars at such low metallicities still moving in a disc configuration is puzzling. Di Matteo et al. (2019) pointed out that the earliest star formation in the Milky Way took place in a thick disc configuration that was partially heated due to mergers, based on the observational results evaluating the impact of the merger with Gaia–Enceladus in the kinematics of the chemically defined thick disc (see figs 12 and 13 in Di Matteo et al. 2019). On the other hand, Sestito et al. (2021) explored NIHAO–UHD cosmological zoom in simulations (Buck et al. 2020) and ended up with two possible and non-exclusive scenarios: (1) these UMP stars in planar orbits were accreted through minor mergers on to the Galactic plane and deposited in the disc by dynamical friction; and (2) they were born in gas-rich building blocks that led afterwards to the formation of the Milky Way proto-disc.

The recent work of cosmological simulations by Park et al. (2020) suggested that spatially defined thick discs originate first as thin discs and then evolve to a slower rotating and spatially thicker configuration due to orbital diffusion. The final thin disc observed at redshift zero is younger and rotates faster than the thick disc as a consequence of disc settling. In fact, their simulations (with the highest resolution obtained up to now) showed that both thin and thick discs are *two parts of a single continuous disc component that evolves with time as a result of the continued star formation*

*of thin disc stars and disc heating*. If our metal-poor thin disc was formed in the early disc, then these stars should have conserved their kinematics, not altered by the diffusion that formed the thick disc.

Further characterization is needed to clarify the origin of these stars and their chemical composition and ages would be very valuable. For example, the  $[\alpha/\text{Fe}]$  ratios could help to discriminate the different scenarios proposed. In the case where these stars formed in the early disc, the  $[\alpha/\text{Fe}]$  would be high, similar to the one characterizing the thick disc at  $[\text{Fe}/\text{H}] \sim -1$  ( $[\alpha/\text{Fe}] \sim +0.4$ ). However, if they were formed later from interstellar medium diluted by a second infall of gas, then they would present lower  $[\alpha/\text{Fe}]$  ratios, similar to the chemically metal-rich thin disc ( $[\alpha/\text{Fe}] \sim <0.2$ ). If the latter is true, then these stars could help to better constrain the intensity, initial time, and duration of this second infall.

## 4.2 The thick disc, the *Splash*, and the halo

The presence of the thick disc in the anticentre is necessary to explain the observed rotational velocity distribution. This detected thick disc component also follows the correlation of the mean rotational velocity with metallicity inferred in previous works as the metallicity decreases ( $47 \pm 10 \text{ km s}^{-1}$ , e.g. Kordopatis et al. 2017). Our analysis reveals the presence of the thick disc down to, at least,  $[\text{Fe}/\text{H}] \sim -1.5$ . In our sample, this thick disc component blurs kinematically into the halo when decreasing to metallicities lower than  $-1.5$ , but there are hints of its presence at  $-2 < [\text{Fe}/\text{H}] < -1.5$ . These are the values at which the metal-poor (or metal-weak) thick disc has been detected in previous works: Norris et al. (1985a), Morrison et al. (1990), Ruchti et al. (2011), Kordopatis et al. (2013a), Beers et al. (2014), and Carollo et al. (2019), but see subsequent discussions from Twarog & Anthony-Twarog (1994), Twarog & Anthony-Twarog (1996), Chiba & Beers (2000), Beers et al. (2002), Reddy & Lambert (2008), Brown et al. (2008), Kordopatis et al. (2013a), Hawkins et al. (2015), Li & Zhao (2017), and Hayes et al. (2018a).

The fraction of the thick disc at  $[\text{Fe}/\text{H}] < -1$  inferred in the previous works of An & Beers (2020a) and Di Matteo et al. (2019) was around 40 per cent. We recover a lower fraction: the weight of our fitted Gaussians with a mean close to the typical mean  $V_\phi$  of the thick disc is 14 per cent in the range between  $-1.5 < [\text{Fe}/\text{H}] < -0.8$  and 19 per cent at  $-2 < [\text{Fe}/\text{H}] < -1.5$ .

The extension of metal-rich stars at  $[\text{Fe}/\text{H}] > -1$  towards non-rotating and counter-rotating  $V_z$  values is compatible with the kinematical effect of a heated disc. Recently, signatures of a heated thick disc (the so-called *Splash*) have been detected in Gaia Collaboration (2018b), Fernández-Alvar et al. (2019), Belokurov et al. (2020a), Di Matteo et al. (2019), Gallart et al. (2019), and specifically in the anticentre by An & Beers (2020b), as a consequence of the merger with the dwarf satellite Gaia–Enceladus–Sausage (Belokurov et al. 2018; Haywood et al. 2018; Helmi et al. 2018; Di Matteo et al. 2019). Besides, previous work (Di Matteo et al. 2019) showed that the kinematical signature is more prominent at  $-1 < [\text{Fe}/\text{H}] < -0.3$ , in line with our results. Recent analysis of cosmological simulations (Grand et al. 2020) has found evidence of scatter of stars due to gas-rich mergers like GES leaving stars with a broad range of velocities that link the thick disc with the inner halo (although see also Amarante et al. 2020 for an alternative origin for the *Splash* due to clump scattering). On the other hand, there have also been detections of a heated thin disc, such as the overdensities at the Monoceros and TriAnd (Bergemann et al. 2018; Hayes et al. 2018b; Fernández-Alvar et al. 2019), likely due to the interaction with Sagittarius (Laporte et al. 2019; Thomas et al. 2019). A further analysis of such heated metal-rich stars is needed to identify the several processes among our sample, which is beyond the scope of this paper.



The non-rotating and retrograde overdensity of stars extends down to the lowest metallicities. An & Beers (2020b) in their analysis of anticentre stars detected a stellar overdensity distinct to GES, the *Splash*, and the thick disc, with a net prograde rotation. They called it the *in situ* halo. Our results are in agreement with this finding. This component would be different than the metal-poor thick disc and a retrograde population, as shown by the different Gaussians that fit the data. However, contrary to what was observed in An & Beers (2020b) at the low  $z$  range that we are studying, we see that this prograde halo dominates the distribution at such low-metallicities compared to the disc and retrograde components.

These results reinforce the power of photometric surveys such as Pristine. Right now, the Pristine metallicity estimates are linked to the SDSS photometry. There is, however, in preparation the conversion of the metallicity scale to *Gaia* photometry. This will hugely improve the Pristine analytical power: first, because there will be more stars with a metallicity estimate, not only stars with an SDSS photometric counterpart, and secondly, because *Gaia* photometry is more accurate than the SDSS one, and this will help to derive more precise metallicities.

## 5 CONCLUSIONS

We studied the rotational velocity distribution of Pristine stars towards the anticentre. We took advantage of the fact that at this direction the rotational velocity component  $V_\phi$  can be associated to the linear projection of the proper motion over the galactic longitude direction  $V_\ell$ . We estimated the rotational velocities from *Gaia* EDR3 proper motions and explored their distribution as a function of their metallicity. We inferred the metallicities from Pristine photometry, improving the calibration method by taking into account the stellar type of the stars.

Our analysis revealed the presence of very metal-poor stars ( $-3.5 < [\text{Fe}/\text{H}] < -1.5$ ) moving with rotational velocities typical of the thin disc. In addition, although our sample is dominated by a kinematical thin disc at metallicities  $[\text{Fe}/\text{H}] > -0.8$ , the presence of a thick disc component is also required to explain the observed velocity distribution. We also detect the signature of a heated disc at high metallicities,  $[\text{Fe}/\text{H}] > -1.5$ . At  $[\text{Fe}/\text{H}] < -1.5$ , our anticentre sample is dominated by a kinematical halo with a net prograde rotation.

Upcoming large-scale spectroscopic surveys such as WEAVE (Dalton et al. 2016) and 4-MOST (de Jong et al. 2019) at high and low resolution will significantly increase the number of stars in Pristine for which radial velocities will be measured. This will allow us to evaluate whether the rotational velocity distribution detected in this work shows the same pattern in other directions of the Galaxy and to fully characterize the kinematics with the vertical and radial velocity components. In addition, other chemical abundances will be measurable from spectroscopy. They will bring valuable information to constrain the scenarios that formed the observed galactic stellar populations.

## ACKNOWLEDGEMENTS

We thank Morgan Fouesneau for his valuable help providing a quality selection criteria to remove potential variable stars among *Gaia* data. We also gratefully thank Leticia Carigi for her insight in the interpretation of the present results based on chemical evolution models. EFA, GK, VH, and NFM gratefully acknowledge support from the French National Research Agency (ANR)-funded project ‘Pristine’ (ANR-18-CE31-0017). We gratefully thank the CFHT staff for performing the Pristine observations in queue mode, for their reactivity in adapting the schedule, and for answering our

questions during the data reduction process. ES acknowledges funding through VIDI grant ‘Pushing Galactic Archaeology to its limits’ (with project number VI.Vidi.193.093), which is funded by the Dutch Research Council (NWO). GT acknowledges support from the Agencia Estatal de Investigación (AEI) of the Ministerio de Ciencia e Innovación (MCINN) under grant with reference (FJC2018-037323-I). JIGH acknowledges financial support from the Spanish Ministry of Science and Innovation (MICINN) project AYA2017-86389-P and also from the Spanish MICINN under 2013 Ramón y Cajal program 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 Science de l’Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. The observations at the Canada–France–Hawaii Telescope were performed with care and respect from the summit of Maunakea, which is a significant cultural and historic site.

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

## DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

## REFERENCES

- Abazajian K. N. et al., 2009, *ApJS*, 182, 543  
 Adibekyan V. Z. et al., 2013, *A&A*, 554, A44  
 Aguado D. S. et al., 2019a, *MNRAS*, 490, 2241  
 Aguado D. S. et al., 2019b, *MNRAS*, 490, 2241  
 Alam S., Ho S., Vargas-Magaña M., Schneider D. P., 2015, *MNRAS*, 453, 1754  
 Allende Prieto C., Kawata D., Cropper M., 2016, *A&A*, 596, A98  
 Amarante J. A. S., Beraldo e Silva L., Debattista V. P., Smith M. C., 2020, *ApJ*, 891, L30  
 An D., Beers T. C., 2020a, *ApJ*, 897, 39  
 An D., Beers T. C., 2020b, *ApJ*, 897, 39  
 An D., Beers T. C., 2021, *ApJ*, 918, 74  
 Bailer-Jones C. A. L., 2015, *PASP*, 127, 994  
 Beers T. C., Christlieb N., 2005, *ARA&A*, 43, 531  
 Beers T. C., Drilling J. S., Rossi S., Chiba M., Rhee J., Führmeister B., Norris J. E., von Hippel T., 2002, *AJ*, 124, 931  
 Beers T. C., Norris J. E., Placco V. M., Lee Y. S., Rossi S., Carollo D., Masseron T., 2014, *ApJ*, 794, 58  
 Belokurov V., Erkal D., Evans N. W., Koposov S. E., Deason A. J., 2018, *MNRAS*, 478, 611  
 Belokurov V., Sanders J. L., Fattahi A., Smith M. C., Deason A. J., Evans N. W., Grand R. J. J., 2020a, *MNRAS*, 494, 3880  
 Belokurov V., Sanders J. L., Fattahi A., Smith M. C., Deason A. J., Evans N. W., Grand R. J. J., 2020b, *MNRAS*, 494, 3880  
 Bensby T., Feltzing S., Oey M. S., 2014, *A&A*, 562, A71  
 Bergemann M. et al., 2018, *Nature*, 555, 334  
 Binney J., Tremaine S., 2008, *Galactic Dynamics: Second Edition*. Princeton University Press, Princeton, NJ  
 Bonifacio P. et al., 2019, *MNRAS*, 487, 3797  
 Brown W. R., Beers T. C., Wilhelm R., Allende Prieto C., Geller M. J., Kenyon S. J., Kurtz M. J., 2008, *AJ*, 135, 564  
 Buck T., Obreja A., Macciò A. V., Minchev I., Dutton A. A., Ostriker J. P., 2020, *MNRAS*, 491, 3461

- Caffau E. et al., 2017, *Astron. Nachr.*, 338, 686
- Caffau E. et al., 2020, *MNRAS*, 493, 4677
- Cardelli J. A., Clayton G. C., Mathis J. S., 1989, *ApJ*, 345, 245
- Carollo D. et al., 2019, *ApJ*, 887, 22
- Carter C. et al., 2021, *ApJ*, 908, 208
- Chen X., Wang S., Deng L., de Grijs R., Liu C., Tian H., 2019, *Nature Astron.*, 3, 320
- Cheng X. et al., 2020, *ApJ*, 905, 49
- Chiappini C., Matteucci F., Gratton R., 1997, *ApJ*, 477, 765
- Chiappini C., Matteucci F., Romano D., 2001, *ApJ*, 554, 1044
- Chiba M., Beers T. C., 2000, *AJ*, 119, 2843
- Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, *MNRAS*, 319, 168
- Cordoni G. et al., 2021, *MNRAS*, 503, 2539
- Dalton G. et al., 2016, in Evans C. J., Simard L., Takami H., eds, *Proc. SPIE Conf. Ser.*, Vol. 9908, *Ground-Based and Airborne Instrumentation for Astronomy VI*, SPIE, p. 99081G
- de Jong R. S. et al., 2019, *Messenger*, 175, 3
- Di Matteo P., Haywood M., Lehnert M. D., Katz D., Khoperskov S., Snaith O. N., Gómez A., Robichon N., 2019, *A&A*, 632, A4
- Di Matteo P., Spite M., Haywood M., Bonifacio P., Gómez A., Spite F., Caffau E., 2020, *A&A*, 636, A115
- Doi M. et al., 2010, *AJ*, 139, 1628
- El-Badry K. et al., 2018, *MNRAS*, 480, 652
- Fermani F., Schönrich R., 2013, *MNRAS*, 432, 2402
- Fernández-Alvar E. et al., 2019, *MNRAS*, 487, 1462
- Frebel A., Norris J. E., 2015, *ARA&A*, 53, 631
- Freeman K., Bland-Hawthorn J., 2002, *ARA&A*, 40, 487
- Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, *AJ*, 111, 1748
- Gaia Collaboration, 2016, *A&A*, 595, A1
- Gaia Collaboration, 2018a, *A&A*, 616, A1
- Gaia Collaboration, 2018b, *A&A*, 616, A10
- Gaia Collaboration, 2018c, *A&A*, 616, A11
- Gaia Collaboration, Brown A. G. A., Vallenari A., Prusti T., de Bruijne J. H. J., Babusiaux C., Biermann M., 2020, *A&A*, 649, A1
- Gaia Collaboration, 2021a, *A&A*, 649, A8
- Gaia Collaboration, 2021b, *A&A*, 649, A1
- Gallart C., Bernard E. J., Brook C. B., Ruiz-Lara T., Cassisi S., Hill V., Monelli M., 2019, *Nature Astron.*, 3, 932
- Grand R. J. J. et al., 2020, *MNRAS*, 497, 1603
- Grisoni V., Matteucci F., Romano D., Fu X., 2019, *MNRAS*, 489, 3539
- Grisoni V., Matteucci F., Romano D., Fu X., 2020a, *Mem. Soc. Astron. Italiana*, 91, 163
- Grisoni V., Cescutti G., Matteucci F., Forsberg R., Jönsson H., Ryde N., 2020b, *MNRAS*, 492, 2828
- Grisoni V., Romano D., Spitoni E., Matteucci F., Ryde N., Jönsson H., 2020c, *MNRAS*, 498, 1252
- Gunn J. E. et al., 1998, *AJ*, 116, 3040
- Hawkins K., Jofré P., Masseron T., Gilmore G., 2015, *MNRAS*, 453, 758
- Hayden M. R. et al., 2014, *AJ*, 147, 116
- Hayes C. R. et al., 2018a, *ApJ*, 852, 49
- Hayes C. R. et al., 2018b, *ApJ*, 859, L8
- Haywood M., Di Matteo P., Lehnert M. D., Snaith O., Khoperskov S., Gómez A., 2018, *ApJ*, 863, 113
- Helmi A., Babusiaux C., Koppelman H. H., Massari D., Veljanoski J., Brown A. G. A., 2018, *Nature*, 563, 85
- Kielty C. L. et al., 2020, *MNRAS*, 506, 1438
- Kordopatis G. et al., 2011, *A&A*, 535, A107
- Kordopatis G. et al., 2013a, *MNRAS*, 436, 3231
- Kordopatis G. et al., 2013b, *A&A*, 555, A12
- Kordopatis G., Wyse R. F. G., Chiappini C., Minchev I., Anders F., Santiago B., 2017, *MNRAS*, 467, 469
- Kordopatis G., Recio-Blanco A., Schultheis M., Hill V., 2020, *A&A*, 643, A69
- Laporte C. F. P., Minchev I., Johnston K. V., Gómez F. A., 2019, *MNRAS*, 485, 3134
- Laporte C. F. P., Koposov S. E., Belokurov V., 2021, preprint ([arXiv:2103.12737](https://arxiv.org/abs/2103.12737))
- Larson R. B., 1976, *MNRAS*, 176, 31
- Lee Y. S. et al., 2011, *ApJ*, 738, 187
- Li C., Zhao G., 2017, *ApJ*, 850, 25
- Lindegren L., Dravins D., 2021, *A&A*, 652, A45
- Lindegren L. et al., 2020a, *A&A*, 649, A4
- Lindegren L. et al., 2020b, *A&A*, 649, A2
- López-Corredoira M., Cabrera-Lavers A., Garzón F., Hammersley P. L., 2002, *A&A*, 394, 883
- Minchev I., Chiappini C., Martig M., 2013, *A&A*, 558, A9
- Minchev I., Chiappini C., Martig M., 2016, *Astron. Nachr.*, 337, 944
- Minchev I., Steinmetz M., Chiappini C., Martig M., Anders F., Matijevic G., de Jong R. S., 2017, *ApJ*, 834, 27
- Momany Y., Zaggia S., Gilmore G., Piotto G., Carraro G., Bedin L. R., de Angeli F., 2006, *A&A*, 451, 515
- Morrison H. L., Flynn C., Freeman K. C., 1990, *AJ*, 100, 1191
- Myeong G. C., Evans N. W., Belokurov V., Sanders J. L., Koposov S. E., 2018, *ApJ*, 863, L28
- Myeong G. C., Vasiliev E., Iorio G., Evans N. W., Belokurov V., 2019, *MNRAS*, 488, 1235
- Norris J., Bessell M. S., Pickles A. J., 1985a, *ApJS*, 58, 463
- Norris J., Bessell M. S., Pickles A. J., 1985b, *ApJS*, 58, 463
- O'Donnell J. E., 1994, *ApJ*, 422, 158
- Park M. J. et al., 2020, *Astrophys. J. Suppl. Ser.*, 254, 2
- Pedregosa F. et al., 2011, *J. Mach. Learn. Res.*, 12, 2825
- Perryman M. A. C., 1997, in Bonnet R. M. et al., eds, *Hipparcos – Venice '97*, Vol. 402. ESA Special Publication, p. 1
- Poggio E., Drimmel R., Andrae R., Bailer-Jones C. A. L., Fouesneau M., Lattanzi M. G., Smart R. L., Spagna A., 2020, *Nature Astron.*, 4, 590
- Poleski R., 2013, preprint ([arXiv:1306.2945](https://arxiv.org/abs/1306.2945))
- Raftery A. E., 1995, *Sociol. Methodol.*, 25, 111
- Re Fiorentin P., Lattanzi M. G., Spagna A., 2019, *MNRAS*, 484, L69
- Recio-Blanco A. et al., 2014, *A&A*, 567, A5
- Recio-Blanco A. et al., 2016, *A&A*, 585, A93
- Reddy B. E., Lambert D. L., 2008, *MNRAS*, 391, 95
- Reylé C., Marshall D. J., Robin A. C., Schultheis M., 2009, *A&A*, 495, 819
- Ruchti G. R. et al., 2011, *ApJ*, 737, 9
- Schlegel D. J., Finkbeiner D. P., Davis M., 1998, *ApJ*, 500, 525
- Schönrich R., 2012, *MNRAS*, 427, 274
- Schönrich R., Binney J., Dehnen W., 2010, *MNRAS*, 403, 1829
- Sestito F. et al., 2019, *MNRAS*, 484, 2166
- Sestito F. et al., 2020, *MNRAS*, 497, L7
- Sestito F. et al., 2021, *MNRAS*, 500, 3750
- Sharma S., Bland-Hawthorn J., 2013, *ApJ*, 773, 183
- Skowron D. M. et al., 2019, *Science*, 365, 478
- Spagna A., Lattanzi M. G., Re Fiorentin P., Smart R. L., 2010, *A&A*, 510, L4
- Spitoni E., Silva Aguirre V., Matteucci F., Calura F., Grisoni V., 2019, *A&A*, 623, A60
- Spitoni E., Verma K., Silva Aguirre V., Calura F., 2020, *A&A*, 635, A58
- Spitoni E. et al., 2021, *A&A*, 647, A73
- Starkenburger E. et al., 2017, *MNRAS*, 471, 2587
- Thomas G. F. et al., 2019, *MNRAS*, 483, 3119
- Twarog B. A., Anthony-Twarog B. J., 1994, *AJ*, 107, 1371
- Twarog B. A., Anthony-Twarog B. J., 1996, *AJ*, 111, 220
- Venn K. A. et al., 2020, *MNRAS*, 492, 3241
- Wyse R. F. G., 2016, in Skillen I., Balcells M., Trager S., eds, *ASP Conf. Ser.*, Vol. 507, *Multi-Object Spectroscopy in the Next Decade: Big Questions, Large Surveys, and Wide Fields*. Astron. Soc. Pac., San Francisco, p. 13,
- Yanny B. et al., 2009, *AJ*, 137, 4377
- Youakim K. et al., 2017, *MNRAS*, 472, 2963
- Youakim K. et al., 2020, *MNRAS*, 492, 4986
- Yusifov I., 2004, in Uyaniker B., Reich W., Wielebinski R., eds, *The Magnetized Interstellar Medium. Proceedings of the conference, held in Antalya, Turkey, September 8 - 12, 2003*, p. 165