

# A slow spin to win - the gradual evolution of the proto-Galaxy to the old disc

Akshara Viswanathan<sup>1\*</sup>, Danny Horta<sup>2</sup>, Adrian M. Price-Whelan<sup>2</sup>, and Else Starkenburg<sup>1</sup>

<sup>1</sup> Kapteyn Astronomical Institute, University of Groningen, Landleven 12, 9747 AD Groningen, The Netherlands

<sup>2</sup> Center for Computational Astrophysics, Flatiron Institute, 162 5th Ave., New York, NY 10010, USA

Received November 20, 2024; accepted xxx

## ABSTRACT

**Context.** Observational studies are identifying stars thought to be remnants from the earliest stages of the Milky Way's hierarchical mass assembly, referred to as the proto-Galaxy.

**Aims.** We use red giant stars with kinematics from *Gaia* DR3 RVS data and  $[\alpha/M]$  and  $[M/H]$  estimates from low-resolution *Gaia* XP spectra to investigate the relationship between azimuthal velocity and metallicity, aiming to understand the transition from a chaotic proto-Galaxy to a well-ordered, rotating (old) disc-like population.

**Methods.** To analyze the structure of the data in  $[M/H]$ - $v_\phi$  space for both high- and low- $\alpha$  samples with carefully defined  $\alpha$ -separation, we develop a model with two Gaussian components in  $v_\phi$ : one representing a disc-like population and the other a halo-like population. This model is designed to capture the conditional distribution  $P(v_\phi \mid [M/H])$  with a 2-component Gaussian Mixture Model with fixed azimuthal velocities means and standard deviations. To quantify the spin-up of the high- $\alpha$  disc population, we extend this two-component model by allowing the mean velocity and velocity dispersion to vary between the spline knots across the metallicity range used. We also compare our findings with existing literature using traditional Gaussian Mixture Modelling in bins of  $[M/H]$  and investigate using orbital circularity instead of azimuthal velocity.

**Results.** Our findings show that the metal-poor high- $\alpha$  disc gradually spins up across  $[M/H] \sim -1.7$  to  $-1$ , while the low- $\alpha$  sample exhibits a sharp transition at  $[M/H] \approx -1$ . This latter result is due to GES debris dominating the metal-poor end, underscoring the critical role of  $[\alpha/M]$  selection in studying the Milky Way's (old) disc evolution.

**Conclusions.** These results indicate that the proto-Galaxy underwent a slow, monotonic spin-up phase rather than a rapid, dramatic spin-up at  $[M/H] \sim -1$ , as previously inferred.

## 1. Introduction

A key objective in Galactic astronomy is to construct a cohesive formation history of the Milky Way down to its earliest times. More broadly, we aim to understand the extent to which a galaxy retains its formation history over cosmic timescales and to investigate the physical processes shaping disc galaxies using the Milky Way as an example (i.e., Galactic archaeology). The Milky Way is the perfect cosmic backyard for this, offering detailed, unparalleled, star-by-star 6D kinematics and chemical information.

The rotationally supported disc contains most of the stars in the Galaxy. It is well established that the disc consists of multiple components or populations, distinguished by their chemistry, kinematics, spatial extent, and age (Norris et al. 1985; Gilmore & Reid 1983; Chiba & Beers 2000; Nissen & Schuster 2010; Bovy et al. 2012; Haywood et al. 2013; Hayden et al. 2015). Structurally, there is evidence of thin and thick discs with different scale heights. Chemically, there is a distinct dichotomy and bimodality in the abundance of  $\alpha$ -elements relative to iron ( $[\alpha/\text{Fe}]$ ); the Milky Way disc manifests a high- $\alpha$  and low- $\alpha$  sequence at fixed  $[\text{Fe}/\text{H}]$  (Hayden et al. 2015; Imig et al. 2023). Generally, the high- $\alpha$  population is older and has a larger scale height than the low- $\alpha$  population. While the connection between high- $\alpha$  and thick disc and low- $\alpha$  and thin disc is fairly strong in the solar vicinity, this connection weakens significantly at larger galactocentric distances (e.g., Hayden et al. 2015).

On the other hand, the Galactic halo is crucial for understanding the formation and evolution of our Galaxy, as it holds the remnants of ancient cosmic events and provides a window into the processes that shaped the Milky Way. Our current understanding of Galactic stellar halo formation suggests a dual process: (i) gas accretion from cosmic filaments that drives secular evolution and in-situ star formation, and (ii) the accretion of various mass building blocks, which contribute their baryonic material and dark matter to the larger host galaxy, adding to its mass as these building blocks are consumed (see recent reviews by Helmi 2020; Deason & Belokurov 2024). Thanks to the combination of *Gaia* astrometry and extensive spectroscopic surveys, our understanding of stars on halo-like orbits has advanced significantly in recent years. A notable finding was the discovery of stars with chemistry identical to the high- $\alpha$  disc but with halo-like orbits (e.g., Bonaca et al. 2017; Koppelman et al. 2018; Helmi et al. 2018; Belokurov et al. 2020). One interpretation of these stars is that they were born *in situ* and later dynamically heated to halo-like orbits, as some simulations predicted (e.g., Grand et al. 2020). The age distribution of these '*in situ*' halo stars is cut off at lookback times of 8 to 11 Gyr (Gallart et al. 2019; Xiang & Rix 2022), suggesting that the heating event, possibly a merger, occurred at  $z \sim 1$  to 2 (see also Montalbán et al. 2021). Numerous other accreted systems have also been identified as part of the accreted stellar halo (Ibata et al. 1994; Helmi & White 1999; Belokurov et al. 2018; Helmi et al. 2018; Koppelman et al. 2019; Myeong et al. 2019; Yuan et al. 2020; Horta et al. 2021). The most significant of these is the *Gaia*-Enceladus-Sausage (GES) system, which contributes most of the accreted

\* Corresponding author e-mail: astroakshara97@gmail.com

halo within  $\sim 6 - 30$  kpc of the Galactic center (Naidu et al. 2020). There is also evidence for another major building block in the inner Galaxy (*Heracles*: Horta et al. 2021), whose stellar mass could possibly have been as massive as the GES (Horta & Schiavon 2024). It has been speculated that there may be a link between *Heracles* and a grouping of globular cluster populations classified in terms of their orbital and/or age-metallicity properties (Massari et al. 2019; Kruijssen et al. 2019; Forbes 2020). Based on the chemical-dynamical properties of this population and a comparison of these with cosmological simulations such as FIRE (Horta et al. 2024), it is likely that *Heracles* coalesced with the primordial Galaxy before the GES ( $z \gtrsim 2$ ). However, deciphering if such populations arise from one single building block or multiple is still an open question.

While the characterization and origins of the main disc and halo populations are becoming clearer, the very earliest epochs of the Galaxy and the emergence of the disc remain poorly understood. The physical mechanism behind disc formation remains a subject of active debate. Recent studies discover very and extremely metal-poor stars (VMP,  $[Fe/H] < -2.0$ , and EMP,  $[Fe/H] < -3.0$ ) on disc-like orbits (Sestito et al. 2019, 2020; Di Matteo et al. 2020; Viswanathan et al. 2024), driving an important question. *When and how did the Milky Way's old disc form?* Metal-poor stars are old and therefore, studying them can provide insights into the early epochs. The orbit and abundance distributions of old, metal-poor stars today reflect the earliest phases of the Milky Way's star formation and enrichment history (Beers & Christlieb 2005; Frebel & Norris 2015; Starkenburg et al. 2018; Lucey et al. 2019; Horta et al. 2021; Ardern-Arentsen et al. 2024; Horta et al. 2024; McCluskey et al. 2024). Belokurov & Kravtsov (2022) used  $[Al/Fe]$  from APOGEE (Majewski et al. 2017) to distinguish between *in situ* and accreted stars, identifying *in situ* stars down to  $[Fe/H] \sim -1.5$ . They found that the average tangential velocity of the *in situ* stars increased rapidly at  $[Fe/H] \sim -1$ . They interpreted this transition as the epoch when the Galaxy transitioned from a relatively disordered state to well-ordered rotation. Conroy et al. (2022) used abundances and ages from the H3 Survey, finding that this transition coincides with a non-monotonic rise in  $[\alpha/Fe]$  abundances. This suggests a near-instantaneous change in star formation efficiency or gas inflow (see also Chen et al. 2023). In the metal-poor end, Rix et al. (2022) revealed a significant concentration of metal-poor stars ( $[M/H] < -1.5$ ) near the Galactic center, corroborating earlier studies of the very metal-poor component in the Galactic bulge region (e.g., Lucey et al. 2019; Arentsen et al. 2020). This also corroborates with other works such as Tumlinson (2010); Starkenburg et al. (2017); El-Badry et al. (2018); Horta et al. (2021); Ardern-Arentsen et al. (2024). The exact relationship between the proto-Galaxy and the metal-poor stars within the Solar radius is still unclear, although Belokurov & Kravtsov (2023) provide evidence that the proto-Galaxy's density follows a steep power-law with Galactocentric radius. Horta et al. (2024) used Milky Way-analogs from high-resolution FIRE simulations, showing that the proto-Galactic populations in all their Milky Way analogs exhibit weak net rotation aligned with the present-day disc. Semenov et al. (2024) used the Illustris TNG50 simulations to interpret evidence of the Milky Way's early disc formation. They suggest that rapid early mass growth was key. However, pin-pointing the exact driver is difficult, as both hot halo formation and gravitational potential steepening occur alongside disc spin-up, indicating that a concentrated potential might be a result, not a cause, of disc formation (Dillamore et al. 2024).

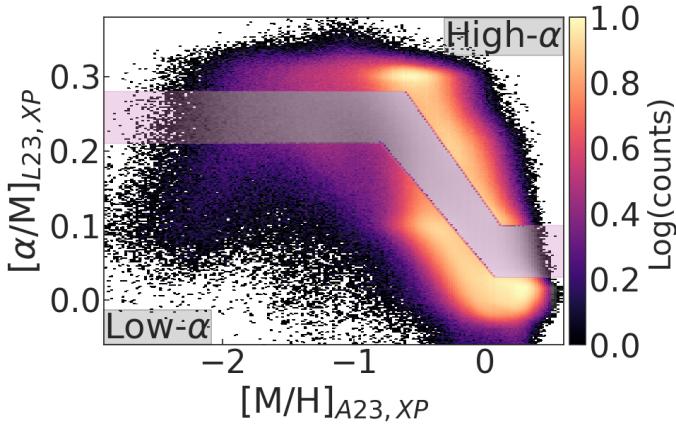
Recently, Zhang et al. (2023) analyzed the *Gaia* DR3 RVS kinematic sample of the metal-poor population from Andrae et al. (2023a) using XGBoost algorithm on *Gaia* XP spectra in the Milky Way and identified two kinematic groups in velocity-[M/H] space: one stationary accreted halo and the other with a net prograde rotation of 80 km/s associating it with the proto-Galaxy. Chandra et al. (2023) used a similar *Gaia* RVS+XP sample with an additional  $\alpha$ -separation to study the MW evolution in a similar parameter space (orbital circularity-[Fe/H]) instead of velocities ( $[V_r, V_\phi, V_z]$ -[Fe/H]). They found that the old high- $\alpha$  disc begins at  $[Fe/H] \sim -1$  dex, with a dramatic transition from a proto-Galaxy to old disc.

In this work, we build on these studies in more detail and set out to study the transition of a proto-Galactic population to a high- $\alpha$  old disc using kinematics from *Gaia* DR3 RVS sample crossmatched with *Gaia* XP spectra [ $\alpha/M$ ] and [M/H] abundances from Andrae et al. (2023a) and Li et al. (2024). Section 2 presents the underlying data used in this work, along with introducing the  $\alpha$ -separation. In section 3, we introduce the azimuthal velocity versus metallicity space where we study the evolution of the high- $\alpha$  disc. We also compare the results from this space with APOGEE chemical abundances to see if our inferences hold well with high-resolution spectroscopy. Section 4 presents a simple two component gaussian mixture model corresponding to a disc-like and halo-like population using knots across the metallicity range to carry information between different metallicities, to interpret the transition between the proto-Galactic component to a rotation-dominated high- $\alpha$  disc. In section 6, we discuss the conclusions derived in this work in the context of what we already know about the proto-Milky Way and the old high- $\alpha$  disc. We also compare our results with the recent literature and discuss some limitations and future prospects in this regard. Section 7 presents a summary of our conclusions.

## 2. Gaia DR3 XP+RVS Data

The *Gaia* DR3 catalogue provides low-resolution spectroscopy (XP) for approximately 200 million stars brighter than  $G = 17.65$ , with radial velocities (RVS) available for a subset of around 30 million (Gaia Collaboration et al. 2023; De Angeli et al. 2023; Katz et al. 2023). Subsequent studies have derived precise metallicity and  $\alpha$ -abundance estimates from these spectra (Andrae et al. 2023a; Zhang et al. 2023; Martin, Starkenburg et al. 2023; Li et al. 2024). This dataset is exceptional due to its immense size, homogeneity, all-sky coverage, and relatively straightforward selection function. It is therefore an ideal resource to comprehensively investigate our Galaxy's formation and evolution.

Andrae et al. (2023a) developed a powerful tool using a machine learning model called *XGBoost* – to estimate a star's metallicity ([M/H]) from its low-resolution XP spectrum. They trained this model using a robust dataset from the Apache Point Observatory Galactic Evolution Experiment (APOGEE: Majewski et al. 2017), further enhanced with metal-poor stars from Li et al. (2022). The metallicities used for training are highly accurate, having been validated against established spectroscopic surveys. For more details on how they chose specific features and validated their catalog, we refer the reader directly to Andrae et al. (2023a). We leverage their metallicity data to create a refined sample of red giant branch (RGB) stars with line-of-sight velocities from *Gaia* DR3 radial velocity spectrometer (RVS). The following selection is made to the Andrae et al. (2023a) catalogue similar to their vetted RGB sample released as Table 2 in their work:



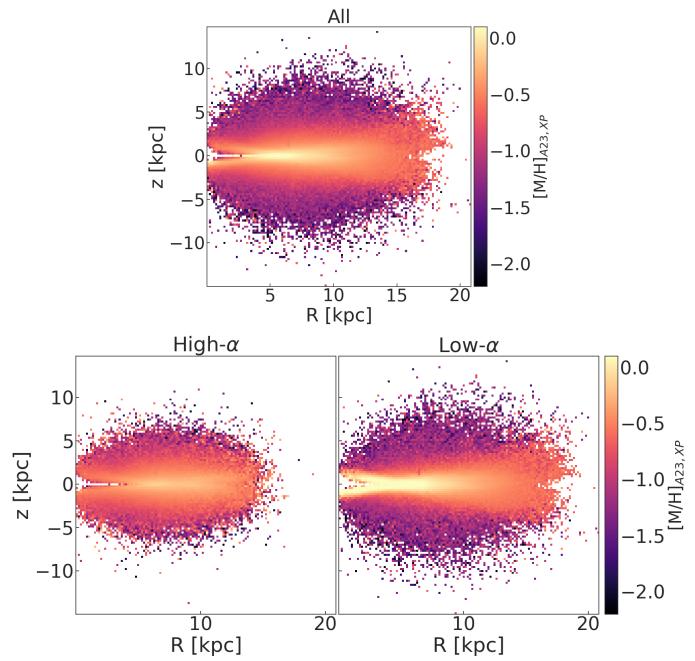
**Fig. 1.** Logarithmic density of  $[\alpha/\text{M}]$  versus  $[\text{M}/\text{H}]$ . The purple band represents the high- and low- $\alpha$  sequence separation defined in this work (see text for details). Stars in the purple band are excluded. The bulk of accreted last major merger (GES) is primarily restricted to the low- $\alpha$  population with our selection.

- We focused on bright stars ( $\text{phot\_g\_mean\_mag} < 16$ ) with highly reliable parallaxes ( $\varpi/\sigma_\varpi > 5$ ). This ensures sufficient data quality for robust metallicity estimates.
- We excluded hot stars ( $\text{teff\_xgboost} > 5200$  K) as their measured metallicities can be misleadingly low.
- We applied a series of cuts based on colour and broad-band magnitudes ( $\text{logg\_xgboost}$ ,  $M_{W1}$ ,  $G-W2$ ,  $G_{BP}-W1$ ) to select genuine red giants on the Hertzsprung-Russell (HR) diagram. These are listed here:
  - $\text{logg\_xgboost} < 3.5$
  - $M_{W1} > -0.3 - 0.006 \times (5500 - \text{teff\_xgboost})$
  - $M_{W1} > -0.01 \times (5300 - \text{teff\_xgboost})$ , where  $M_{W1} = W1 + 5 \times \log_{10}(\pi/100)$
  - $(G - W2) < 0.2 + 0.77 \times (G_{BP} - W1)$
- We removed stars with a high probability (>90%) of belonging to globular clusters (GCs). This probability comes from a catalogue by [Vasiliev & Baumgardt \(2021\)](#).

This resulted in a sample of approximately 11 million RGB stars with *Gaia* XP metallicities and RVS radial velocities.

[Li et al. \(2024\)](#) created a catalogue of alpha-over-iron abundance ( $[\alpha/\text{M}]$ ) values derived from *Gaia* XP spectra using a neural network model. This model learns to use XP spectra to predict stellar labels but also uses and predicts the high-resolution APOGEE spectra that lead to the same stellar labels. This catalogue, too, has been cross-checked against existing surveys for accuracy, demonstrating a median error of only  $\sim 0.05$  dex in  $[\alpha/\text{M}]$  for stars within our sample. We cross-matched our clean RGB sample with this catalogue, yielding a final sample of 9,645,972 RGB stars with metallicities and  $\alpha$ -abundances from *Gaia* XP spectra, and full 6D phase space coordinates from *Gaia* astrometry and RVS line-of-sight velocities.

Figure 1 shows the logarithmic density distribution of  $[\alpha/\text{M}]$  versus  $[\text{M}/\text{H}]$  of our final sample. While Figure 1 showcases a clear separation between high- $\alpha$  and low- $\alpha$  stars in terms of their chemical abundance, this distinction does not necessarily translate directly to their height above the Milky Way's midplane. The chemical difference between these populations maps more strongly onto their radial distribution (distance from the Galactic Centre) than their vertical thickness ([Hayden et al. 2015](#)). Therefore, we leverage this chemical separation (high- $\alpha$  versus low- $\alpha$ ) as a tool to differentiate the two chemically distinct disc popu-



**Fig. 2.** Distribution of stars in cylindrical Galactic coordinates (R-z plane) colour-coded by their mean metallicities for all the stars in our sample (top), high- $\alpha$  selection (bottom left), and low- $\alpha$  selection (bottom right). We can see that dust near the Milky Way's midplane significantly impacts our survey selection. We see a sharp negative metallicity gradient w.r.t height above the disc plane in *all stars* and low- $\alpha$  stars. High- $\alpha$  stars have a shallower negative metallicity gradient.

lations and then analyze their spatial and kinematic properties (positions and velocities) independently.

## 2.1. High- and low- $\alpha$ sequences

The purple band in Figure 1 shows the selection cut used to separate high-/low- $\alpha$  populations. Stars in the purple band are cut out to ensure we have purer samples of high-/low- $\alpha$  stars. This is especially important for high- $\alpha$  stars since we use the high- $\alpha$  sample as the sample of stars tracing the evolution of the proto-Galaxy to high- $\alpha$  disc over metallicity and any contamination from the denser low- $\alpha$  disc can affect the purity of the high- $\alpha$  sample. Typically, the high-alpha sequence is attributed to stars formed *in situ*. However, in practise, the *in situ* versus *accreted* separation is not as simple as using an  $\alpha$ -separation, the consequences of which are discussed in section 6.

The selection is defined using the following equations:

$$\text{High-}\alpha = \begin{cases} [\text{M}/\text{H}] < -0.6 \& [\alpha/\text{M}] > 0.28 \\ [\text{M}/\text{H}] \in [-0.6, 0.125] \& [\alpha/\text{M}] > -0.25 \times [\text{M}/\text{H}] \\ & + 0.13 \\ [\text{M}/\text{H}] > 0.125 \& [\alpha/\text{M}] > 0.1 \end{cases} \quad (1)$$

$$\text{Low-}\alpha = \begin{cases} [\text{M}/\text{H}] < -0.8 \& [\alpha/\text{M}] < 0.21 \\ [\text{M}/\text{H}] \in [-0.8, 0.07] \& [\alpha/\text{M}] < -0.21 \times [\text{M}/\text{H}] \\ & + 0.045 \\ [\text{M}/\text{H}] > 0.07 \& [\alpha/\text{M}] < 0.03 \end{cases} \quad (2)$$

This selection is somewhat different from the one implemented by Chandra et al. (2023) who use the same sample and abundances from *Gaia* XP spectra. The main difference between Chandra et al. (2023) and our selection are: (i) we use a more stringent selection for high- $\alpha$  to make sure that the bulk of accreted *Gaia*-Enceladus-Sausage (GES) merger (Belokurov et al. 2018; Helmi et al. 2018) does not contaminate the high- $\alpha$  population (see Figure 1), and (ii) the purple band is also quantitatively larger to ensure a sample with highest purity. We test the validity of this separation by using the  $\alpha$ -selection on *APOGEE* DR17 data in the [M/H]-[ $\alpha$ /M] space and using [Al/Fe]-[Mg/Mn] space to see if they occupy accreted, low- $\alpha$  or high- $\alpha$  disc stars regions as described by the tracks used in Horta et al. (2021). The contamination rate from this validation is as low as 5% and 8% for high- and low- $\alpha$  stars respectively. The high- $\alpha$  selection used here is stricter than the selection used in the literature, in order to remove as much accreted low- $\alpha$  stars as possible (mainly, GES, the last major merger). Even though we show the full sample in Figure 1, we restrict the analysis in the rest of this work to metallicities above  $-2.5$ .

## 2.2. Positions and kinematics

To estimate the distance to each RGB star, we use photo-geometric distance provided by Bailer-Jones et al. (2021), which is a Bayesian approach that incorporates both *Gaia* parallax measurements and a prior model of the Milky Way's stellar density distribution. While using a simpler method based solely on zero-point-corrected parallax (Lindegren et al. 2021) yields similar results, the approach by Bailer-Jones et al. (2021) offers a more robust framework for estimating distances across stars with varying signal-to-noise ratios. To ensure using reliable astrometry, we focused on stars with a low renormalised unit weight error ( $\text{ruwe} < 1.4$ ). This metric indicates good astrometric quality, potentially filtering out binary star systems. We adopted several standard assumptions: a Local Standard of Rest velocity ( $V_{LSR}$ ) of 232 km/s, a distance of 8.2 kpc between the Sun and the Milky Way's center (GRAVITY Collaboration et al. 2018), and the Sun's peculiar motion of  $(U_\odot, V_\odot, W_\odot) = (11.1, 12.24, 7.25)$  km/s (Schönrich et al. 2010). This allows us to calculate the positions and velocities for all the stars in our final sample.

Figure 2 shows the distribution of our sample in cylindrical Galactocentric radius,  $R$ , versus height above the midplane,  $z$ , colour-coded by the mean metallicity value for every grouping of stars in each pixel. The top panel of Figure 2 shows this distribution for all the stars in our sample, while bottom left and right panels show this distribution for high-/low- $\alpha$  samples, respectively. It is worth noting that the high- $\alpha$  stars probe a smaller range in  $R$ - $z$  compared to low- $\alpha$  stars; most of the low- $\alpha$  stars at larger scale heights are expected to come from accretion events, and this is likely why the low- $\alpha$  sample covers a wider scale height. We see a strong and steep negative metallicity gradient with increasing  $z$  for the low- $\alpha$  sample; we also see this in the full sample, likely because it is dominated by stars in our low- $\alpha$  selection, while the high- $\alpha$  sample has a shallower negative metallicity gradient with increasing  $z$ . The high- $\alpha$  disc has a larger scale height compared to the low- $\alpha$  disc and we also see a clear disc flaring for the low- $\alpha$  disc (as seen for the Galactic thin disc) with a metallicity gradient towards larger  $R$ , reminiscent of radial migration (see for e.g., Haywood et al. 2013; Hayden et al. 2015; Ratcliffe et al. 2024).

## 3. Milky Way populations in the $v_\phi$ -[M/H] plane

This section summarises the evolution of azimuthal velocity <sup>1</sup> versus metallicity for our high-/low- $\alpha$  samples. In examining this plane, we aim to gain an intuition of when and how quickly metal-poor populations (that for high- $\alpha$  likely contain components of the proto-Galaxy) go from a low net spin and kinematically hotter configuration to a more rotationally-supported high- $\alpha$  disc population.

### 3.1. Azimuthal velocity versus metallicity tracks

Figure 3 illustrates the formation and evolution of the Milky Way using azimuthal velocity,  $v_\phi$ , as a function of metallicity, [M/H], for our full sample (top), and our high-/low- $\alpha$  samples (bottom). For all panels, each [M/H] bin has a size 0.01 dex, and we show the sum normalised histogram of azimuthal velocity, plotted as a 2D column-normalised histogram<sup>2</sup> for all, high-, and low- $\alpha$  stars.

Although it is intuitive to read metallicity as an age sequence, we know that the age-metallicity relation (AMR) of the Milky Way is not monotonic (Bensby et al. 2014; Xiang & Rix 2022; Gallart et al. 2024; Xiang et al. 2024). This is true for *all stars* (top panel, which has a mix of *in situ* and accreted stars), and low- $\alpha$  stars (bottom right panel), as they are made of two different stellar populations (thin disc and accreted mergers). However, Xiang & Rix (2022) have shown that the stars in high- $\alpha$  sequence have a relatively consistent and monotonic AMR. They also show that the high- $\alpha$  disc reached solar metallicities at around  $\sim 8$  Gyr in lookback time. Therefore, the high- $\alpha$  panel should trace the first  $\sim 5\text{-}6$  Gyr of the Milky Way's evolution, which is also proposed by Chandra et al. (2023) in their orbital circularity versus metallicity plane. All the 2D histogram in Figure 3 have 16<sup>th</sup>, 50<sup>th</sup>, and 84<sup>th</sup> percentile tracks of  $v_\phi$  versus [M/H] overlaid. The median and percentile tracks look very similar if we restrict our analysis to solar neighbourhood stars ( $d < 3$  kpc), as the velocities are less position dependent in a small volume around the Sun. However, in order to preserve the number statistics, especially in the low-metallicity end, we use the entire sample in the rest of this paper instead of restricting only to solar neighbourhood stars. As an additional check, we use orbital circularity versus metallicity instead of azimuthal velocity, which is expected to be position independent in section 5. We will discuss the different panels separately in the following subsections.

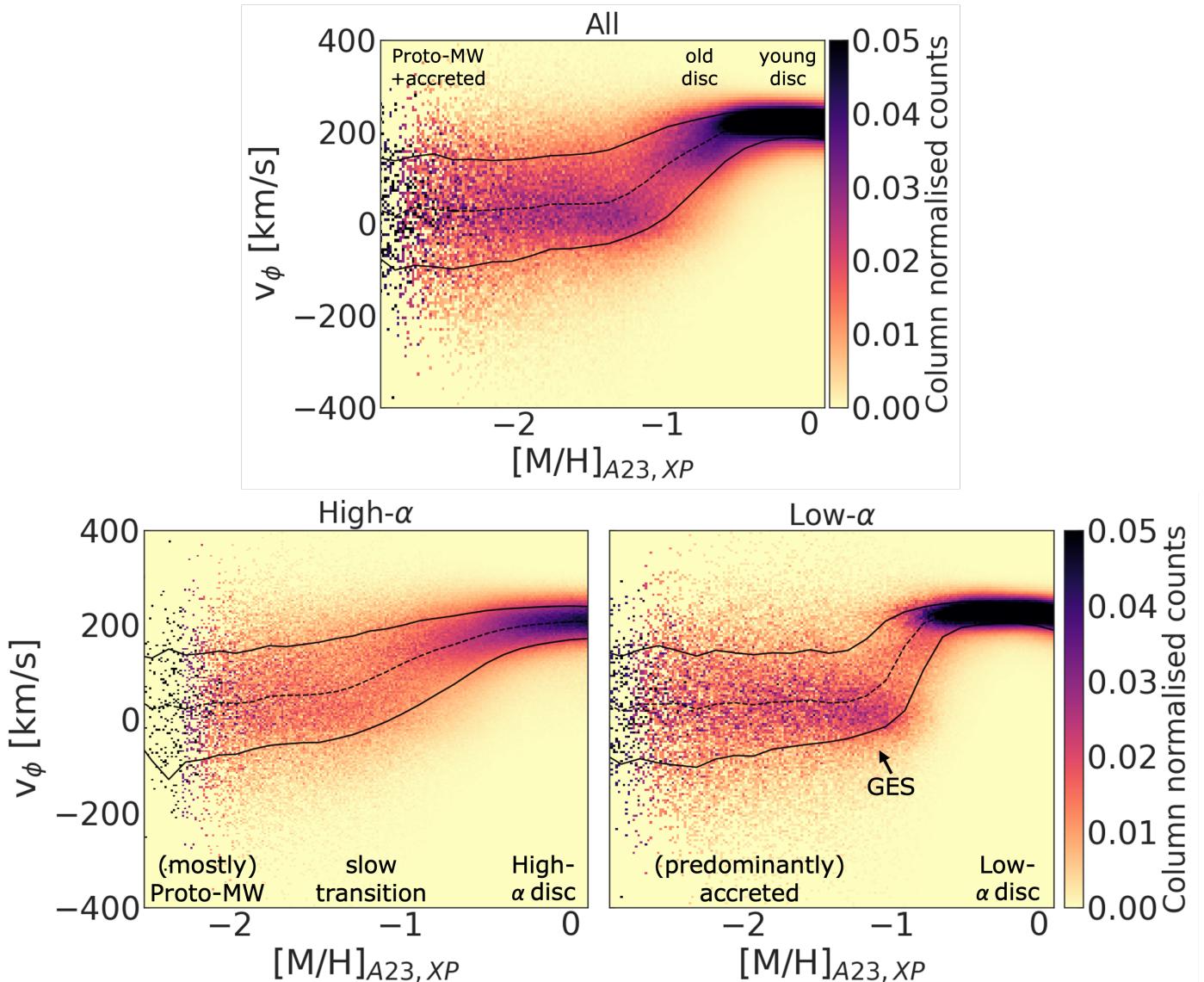
#### 3.1.1. All stars

In Figure 3 top panel, we see strongly rotational ( $v_\phi \sim 200$  km/s), kinematically cold thin disc dominating higher metallicities down to  $[\text{M}/\text{H}] \sim -0.7$ . Below these metallicities, we see the kinematically hotter high- $\alpha$  disc population ( $v_\phi \sim 180$  km/s) which then has a disconnection at  $[\text{M}/\text{H}] \sim -1.0$ . Furthermore, it is possible to see the GES merger (and other accretion events) dominating radial orbits ( $v_\phi \sim 0$  km/s) at metallicities below  $[\text{M}/\text{H}] < -1.0$ . This creates a step function behavior (also seen in the running median tracks) around the metallicities between  $-1.3$  and  $-0.9$  which has been reported in the literature as the spin-

<sup>1</sup> In the rest of this paper, we use azimuthal and rotational velocities interchangeably.

<sup>2</sup> In this work, 2D column-normalised histogram automatically means column normalisation by sum, such that the sum under the histogram curve equals 1. This is directly proportional to the probability density function of azimuthal velocity at each metallicity bin.

## Metallicity trends with azimuthal velocity from *Gaia* XP metallicities



**Fig. 3.** Column-normalised (by sum) 2D histogram of stars in the  $[M/H]$ - $v_\phi$  plane (azimuthal velocity versus metallicity) for all the stars (top), high- $\alpha$  selection (bottom left), and low- $\alpha$  selection (bottom right). The running median track is shown as dashed black line and the 16<sup>th</sup> and 84<sup>th</sup> percentile tracks are shown as black lines in all panels. The running median tracks for the low- $\alpha$  panel look more like a step-function, while high- $\alpha$  tracks is shallower, supporting the gradual monotonic increase, that can be interpreted as a gradual spin-up from old proto-Galactic populations to the present day high- $\alpha$  disc.

up phase, where we have the dramatic transition from what has been dubbed a “proto-Galactic” population with low net rotation to a rotation-supported disc component (Belokurov & Kravtsov 2022; Chandra et al. 2023; Zhang et al. 2023; Kurbatov et al. 2024).

### 3.1.2. Low- $\alpha$ stars

In Figure 3 bottom right panel, we see strongly rotation-supported thin disc ( $v_\phi \sim 200$  km/s) dominating higher metallicities ( $[M/H] > -1$ ) and is almost fully disjoint to the accreted population at lower metallicities ( $[M/H] < -1$ ) that are radial and present low rotational velocities ( $v_\phi \sim 0$  km/s). Here, we can clearly see a step function behaviour at  $[M/H] \sim -1.0$ , which demarks the transition from the accreted populations to the low- $\alpha$  disc. We see little-to-no evolution in the azimuthal velocity and

velocity dispersions for the thin disc low- $\alpha$  stars with increasing metallicities. This is also because metallicity clearly does not trace age for low- $\alpha$  disc stars, but instead the birth radius of the star (e.g., Ratcliffe et al. 2024), i.e., metallicities for low- $\alpha$  stars cannot be read as a temporal axis.

### 3.1.3. High- $\alpha$ stars

In Figure 3 bottom left panel, we see high- $\alpha$  stars with their azimuthal velocity evolving across the metallicity range monotonically. Both the 2D column-normalised histogram and the median tracks show the evolution of a less rotationally supported (kinematically hotter) metal-poor component (likely the proto-Galaxy), with a small but non-negligible mean azimuthal velocity value ( $v_\phi \sim 50$  km/s, McCluskey et al. (2024); Semenov et al. (2024); Horta et al. (2024)). As metallicity increases, we

see the high- $\alpha$  sample increasing in  $v_\phi$  towards more prograde orbits, gradually transitioning from a hotter less rotating component to a rotation-dominated high- $\alpha$  disc ( $v_\phi \sim 180$  km/s). The velocity dispersion for high- $\alpha$  disc orbits becomes smaller with increasing metallicities. Therefore, Figure 3 likely illustrates the emergence of the high- $\alpha$  disc, revealing the evolution of proto-Galactic populations gradually spinning-up with increasing metallicities.

It is important to note that our  $[\alpha/M]$  selection would not fully remove accreted stars in the high- $\alpha$  sample, especially in the metal-poor end,  $[\text{M}/\text{H}] < -1.3$ , where GES and other debris like *Heracles* (Horta et al. 2021) appear. Therefore, accreted halo stars would still be present (although probably contributing a small fraction in the metal-poor end). Additionally, the mean velocities and velocity dispersion that we show are present day velocity distribution. Therefore, Fig 3 shows the instantaneous velocity information for these populations, and not the velocity profile at formation; for these metal-poor (old) populations, these two could be drastically different due to dynamical heating over time, for example. It is also important to note that the hotter high- $\alpha$  disc population, is still present in the high- $\alpha$  sample, but smaller in number than the rotation-supported high- $\alpha$  disc stars. These "hot" high- $\alpha$  disc stars can be easily seen in row-normalised  $[\text{M}/\text{H}]\text{-}v_\phi$  plane as shown in Appendix C.

From the bottom left panel of Figure 3, we find evidence in support of a metal-poor high- $\alpha$  population (likely part of the proto-Galaxy) gaining rotation at a slower pace across wider range of metallicities (between -2 and -0.7) that eventually settles into a high- $\alpha$  disc population.

### 3.2. Azimuthal velocity versus metallicity tracks using high-resolution APOGEE abundances

Figure 4 shows 2D column-normalised histograms of  $[\text{M}/\text{H}]\text{-}v_\phi$  plane for high- $\alpha$  (left), low- $\alpha$  (center), and *all stars* (right) for APOGEE DR17 stars with the same  $\alpha$ -selection described by equation 1 and 2. In APOGEE DR17, we select stars with  $\log g < 3.5$ , excluding potentially problematic data based on quality flags such as ASPCAPFLAG, and WARNING, or BAD flags defined in ASPCAP for  $T_{\text{eff}}$ ,  $\log g$ ,  $[\text{M}/\text{H}]$ , and  $[\alpha/\text{M}]$ . The running 16<sup>th</sup>, 50<sup>th</sup>, and 84<sup>th</sup> percentile tracks for the APOGEE data are shown as black lines and the tracks from our final sample in Figure 3 are shown in gray in all three panels. We find very good resemblance between the track in the *APOGEE* and *Gaia* XP data, especially for high- $\alpha$  stars stars. It shows that a scenario of a the gradual spin-up of proto-Galaxy to the high- $\alpha$  disc population is not due to large errors, as it is supported by a much more reliable high-resolution chemical abundance data. The low- $\alpha$  density distribution (and median tracks) are the ones that vary the most between the two samples. This would be expected if contamination from *in situ* centrally concentrated stars from the proto-Galaxy were present in our low- $\alpha$  sample, while APOGEE has a much cleaner and purer low- $\alpha$  selection, due to higher quality of the chemical abundances. This can also be seen by the thin disc and halo populations being completely disjoint in the low- $\alpha$  2D histograms. This is discussed more in detail in Appendix A using  $[\text{M}/\text{H}]\text{-}v_\phi$  plane towards and away from the inner Galaxy and noting the differences for low- $\alpha$  2D histograms. Lastly, the *all stars* panel in APOGEE also resembles what we see in our *Gaia* XP sample, with stars at lower metallicities dominated by accreted stars and proto-Galaxy populations with low net spin, and higher metallicities dominated by the kinematically hotter high- $\alpha$  disc and kinematically colder low- $\alpha$  disc at even higher metallicities.

Given these results, we are confident that our selection presented in Figure 1 is efficient to separate different stellar populations in the Milky Way, where we can analyse the  $[\text{M}/\text{H}]\text{-}v_\phi$  plane. In the following section, we will set out to model these data using a new mixture model, with the aim of quantifying the point at which metal-poor populations transition into the more rotationally supported disc populations; we also aim to assess the fraction of halo/disc populations at different metallicity bins.

## 4. Modelling the high-/low- $\alpha$ stars in the $v_\phi\text{-}[\text{M}/\text{H}]$ plane

To decipher the underlying structure of the data in  $[\text{M}/\text{H}]\text{-}v_\phi$  space for both the high-/low- $\alpha$  samples, we create a simple model described by two Gaussian components in  $v_\phi$  (one that is more disc-like and other that is more halo-like) the probabilities of which is conditioned on the value of the metallicity, thereby carrying over the 2-component Gaussian Mixture Model information across metallicities. This model is implemented using `numpyro` (Phan et al. 2019; Bingham et al. 2019), a lightweight probabilistic programming library that provides a `numpy` backend for `pyro`.

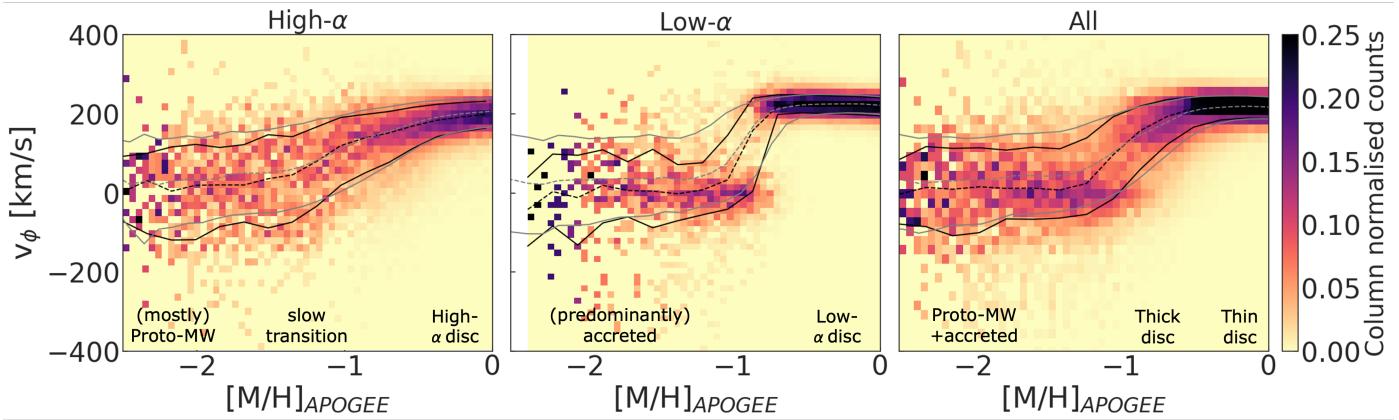
We argue that this model is more well-suited for this problem than a Gaussian mixture Model (GMM) decomposition (even if the number of components is set free and determined using the Bayesian Information Criteria) in bins of  $[\text{M}/\text{H}]$ , because the underlying Gaussians are independent of each other in between different  $[\text{M}/\text{H}]$  bins. Therefore, instead of manually linking different Gaussians across metallicity, our model  $P(v_\phi \mid [\text{M}/\text{H}])$  is able to connect the two Gaussians across the metallicities as a continuous conditional distribution. This is also advantageous as it allows us to follow and interpret the different Gaussian components and thus examine how they vary across metallicity space. An underlying assumption of our method is that we can model the low- $\alpha$  and high- $\alpha$  discs as a Gaussian distribution in  $v_\phi$  space. To first order, this is a good approximation (see middle panels of Figure 9); it is also a good approximation for the stellar halo, as our sample is dominated by the debris of one major accreted satellite (i.e., GES).

Moreover, our model allows us to interpret how different stellar populations evolve in  $v_\phi$  across  $[\text{M}/\text{H}]$ . It also enables us to measure the transition from a halo-like non-rotating population to a disc-like rotating one for both the high-/low- $\alpha$  samples. We reason that the two-component fit is also physically motivated as we know that the Galaxy has a halo and a disc in the high- and low- $\alpha$  regime. While simple in model form, we show here how this model is able to accurately capture the data.

In practice, the model takes in  $[\text{M}/\text{H}]$  and  $v_\phi$  as input parameters, and uses priors on two (separate) Gaussian components to split the data into a halo-like and disc-like population across  $[\text{M}/\text{H}]$ . In detail, we choose 16 evenly spaced knots on a log scale<sup>3</sup> across the metallicity space from  $-2.5 < [\text{M}/\text{H}] < 0.01$ . For each batch of data in these bins, our model computes the fraction of the data can be modeled by one Gaussian compared to the other (i.e., the fractional contribution of each Gaussian); this allows us to quantify how much of the data is better fitted by a halo-like populations vs a disc-like one in each  $[\text{M}/\text{H}]$  bin. We set the priors of each Gaussian component as a normal distribution with  $\mathcal{N}(\mu, \sigma)$ , where  $\mu = 0$  km/s and  $\sigma = 150$  km/s. For the priors on the weights (or fractional contribution) of each Gaussian at the locations of each  $[\text{M}/\text{H}]$ -knot in the spline, we use

<sup>3</sup> We space the knots in the spline on a log scale to account for the fact that there are many more metal-rich stars than metal-poor stars.

## Metallicity trends with azimuthal velocity from APOGEE metallicities



**Fig. 4.** Column-normalised (by sum) 2D histogram of stars in the  $[M/H]$ - $v_\phi$  plane for APOGEE DR17 giants (with the same high- and low- $\alpha$  separation as shown in Figure 1) for all the stars (left), high- $\alpha$  selection (center), and low- $\alpha$  selection (right). The running median track is shown as dashed black line and the 16<sup>th</sup> and 84<sup>th</sup> percentile tracks are shown as black lines in all panels. The median, 16<sup>th</sup> and 84<sup>th</sup> percentile tracks for our *Gaia* XP sample are shown in gray for comparison. The tracks between *Gaia* XP and APOGEE samples show remarkable resemblance in high- $\alpha$  and small differences for low- $\alpha$  and *all* stars due to lower contamination in low- $\alpha$  selection in APOGEE (see Appendix A).

a Dirichlet prior with a constant concentration value of 0.5 for both the components<sup>4</sup>. The means, standard deviations, and relative weights for the two Gaussian components are then sampled using a Hamiltonian Monte Carlo (HMC) inference, using the No U-Turn Sampler (NUTS). We run the sampler for 200 warm-up steps, and 100 sampling steps. This results in the Markov Chain Monte Carlo fits for high- $\alpha$  and low- $\alpha$  stars in our sample. We show the graphical model representation of this model in the left panel of Figure 5 and a list of model parameters and their functional forms are provided in the second column of Table 4.1.

When running this method on our high-/low- $\alpha$  samples, we find that the  $r_{\text{hat}}$  split Gelman Rubin diagnostic parameter is less than 1.1 for the majority of the  $[M/H]$  bins, indicating that the chains have converged. Moreover, the chains look stable, and  $n_{\text{eff}}$ , the number of effective samples, is at least 30 or more for the means, standard deviations, and relative weights. We do note however that for the most  $[M/H]$ -poor bins, the  $r_{\text{hat}}$  values are higher for high- $\alpha$  stars (up to 1.72 for the knots in the metal-poor regime, even though the chains are converged) than for the low- $\alpha$  stars (oscillate below 1.1 regardless of the warm up steps and sample size for the MCMC chains). We tested varying the number of warm-up and sampling steps, but found that this did not affect our results. This result indicates that the samples we generate are a fair representation of the posterior probability distribution over the model parameters for the low- $\alpha$  stars. This is also the case for the metal-rich bins in the high- $\alpha$  sample. However, at lower  $[M/H]$  for high- $\alpha$  stars, our model requires more flexibility to generate a fair representation of posterior distributions. The low- $\alpha$  sample is comprised of two clearly distinct populations, the low- $\alpha$  disc and GES debris at different metallicities. Conversely, the high- $\alpha$  sample is comprised of stars that appear to follow a single trajectory across the entire metallicity range. Moreover, metal-poor high- $\alpha$  populations are likely an amalgamation of stellar populations (i.e., *Heracles*, old *in situ* material, etc). Thus, our results hint that our two-Gaussian component is too simple to be able to capture the distribution of this data perfectly. We can see the results from this model in Figure 3. Here, the low- $\alpha$  stars cleanly separate into two distinct stellar popu-

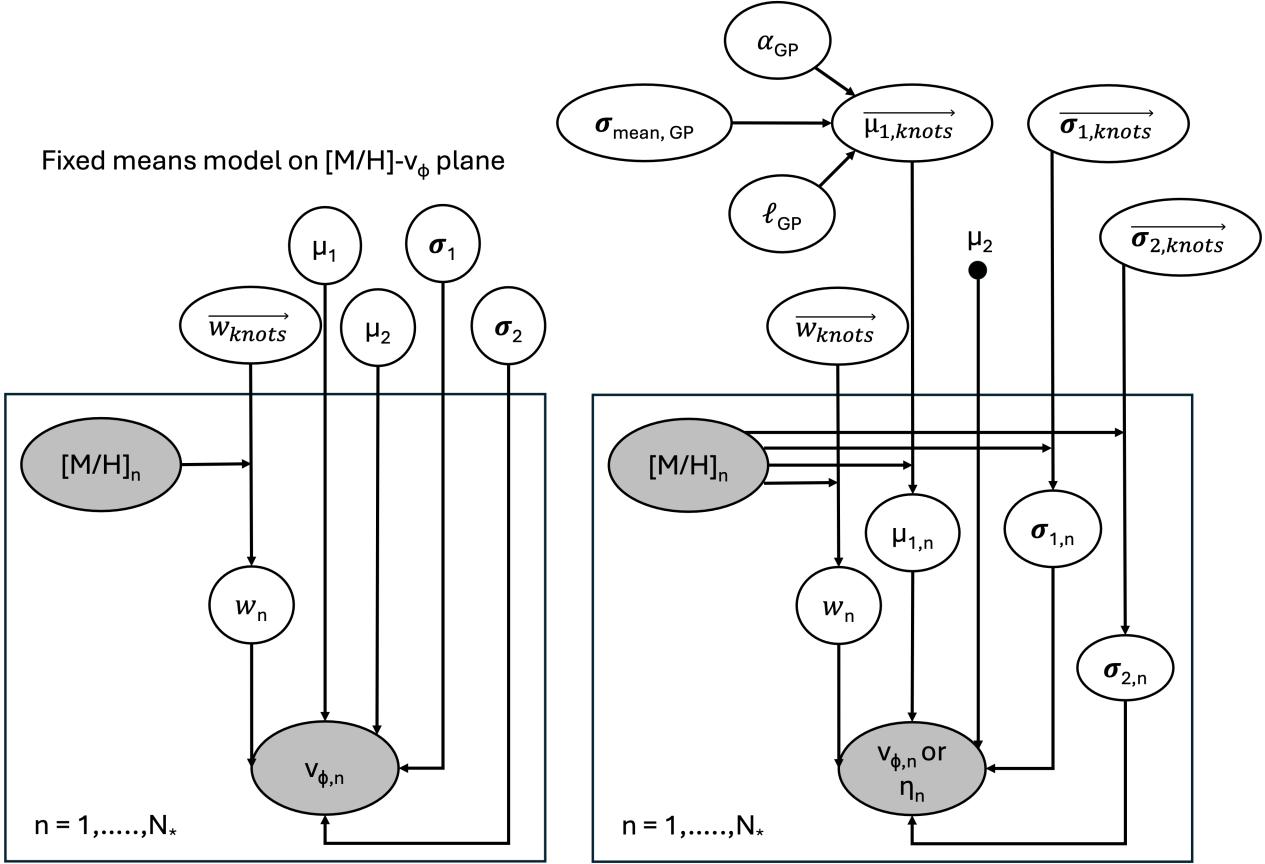
lations, each of which dominates a different metallicity range (note the median tracks displaying almost a step function). At high (low)  $[M/H]$ , the low- $\alpha$  disc (GES debris) dominates. This is not the case for the high- $\alpha$  star sample, that displays a monotonically increasing behaviour in  $v_\phi$  for increasing metallicity. Therefore, a model with the means and standard deviations of the azimuthal velocity varying across the metallicities may be more suited for the high- $\alpha$  stars.

However, despite this limitation, it is useful to compare the results from our model for the high- and low- $\alpha$  population. Thus, we will proceed with this simple two-component model for both high-/low- $\alpha$  populations to assess the fractional contribution of disc-like and halo-like components in each sample.

Figure 6 shows the fractional contribution of a halo-like component in orange and disc-like component in purple with the  $1\sigma$  uncertainty in the fractions as orange and purple shaded bands, respectively. The knots in metallicity chosen to run the model are shown as scatter points. The converged velocity means for high-/low- $\alpha$  stars are  $188 \text{ km s}^{-1}$  and  $223 \text{ km s}^{-1}$ , respectively. These values match well the mean rotational velocity of the high- $\alpha$  (thick) and low- $\alpha$  (thin) disc. The corresponding velocity dispersions are  $39 \text{ km s}^{-1}$  and  $25 \text{ km s}^{-1}$  for high-/low- $\alpha$  discs, showing that the high- $\alpha$  disc is kinematically hotter than the low- $\alpha$  one. Furthermore, the halo-like components have a more radial mean velocity and hotter velocity dispersion when compared to their respective high-/low- $\alpha$  disc samples. Here, the high- $\alpha$  halo-like component has a smaller velocity dispersion ( $74 \text{ km s}^{-1}$ ) than low- $\alpha$  stars ( $107 \text{ km s}^{-1}$ ). This could simply be due to the fact that the two-component fit with fixed mean velocities and standard deviations might not be the best representation of the underlying data for high- $\alpha$  stars. Conversely, it could also imply that the high- $\alpha$  halo component is kinematically colder than the low- $\alpha$  (GES) one.

The main difference we can see between the fractional contribution of disc-like and halo-like stars in high-/low- $\alpha$  subsamples is that the low- $\alpha$  stars completely switch between halo and disc around a very narrow range in metallicities between -1.1 and -0.8. This step function behaviour is clearly seen in the right panel of Figure 6. The small fall in halo-like fractions at lower metallicities could simply be due to noisy data in the metal-poor end. For the low- $\alpha$  stars, within the uncertainties in the

<sup>4</sup> We choose a Dirichlet distribution to ensure that the the sum of the two weights can never be above 1.0 or below 0, as expected for the fractional contribution parameter.

Evolving means model on  $[M/H]$ - $v_\phi$  or  $[M/H]$ - $\eta$  plane

**Fig. 5.** Graphical model representation of the evolving means model, modeling the conditional distribution  $P(v_\phi | [M/H])$  with a 2-component GMM, wherein the means and standard deviations are fixed (left) and varying (right) with respect to the underlying metallicities. The evolving/varying means model is also implemented for circularity as the conditional distribution or  $P(\eta | [M/H])$ . The subscript 1 and 2 stands for the disc-like and halo-like components respectively, with  $\mu$  and  $\sigma$  as the Gaussian mean and standard deviations in  $v_\phi$  or  $\eta$ , conditioned on the value of  $[M/H]$ .  $w$  stands for the relative contribution of the disc-like component with the relative contribution of the halo-like component defined as  $(1-w)$ . A list of the model parameters and their priors/functional forms are given in Table 4.1.

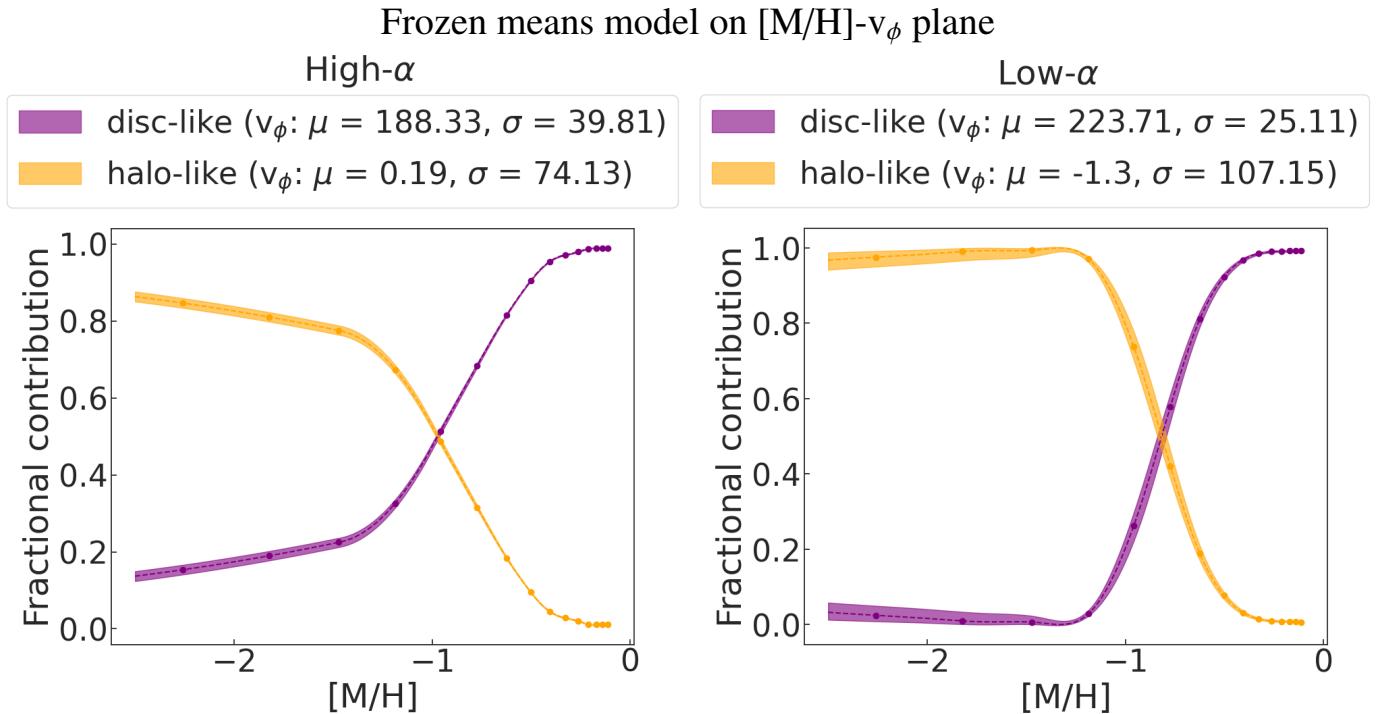
weights, we can clearly see that the metal-poor end is fully composed of halo-like stars ( $[M/H] \leq -1.1$ ) and the metal-rich end is fully composed of disc-like stars ( $[M/H] \geq -0.6$ ). Conversely for high- $\alpha$  stars, we do not see such a steep turn-over between the halo-like and disc-like samples. In contrast, the disc-like component gradually increases with increasing metallicity in its relative fraction, while the halo-like component gradually decreases. The disc-like component is present with 18% relative contribution in the very metal-poor end ( $[M/H] < -2$ ) in this simple model. This result favours a more gradual spin-up scenario. However, as discussed above, our model is not able to fully capture the distribution of the high- $\alpha$  stars in the metal-poor regime. To improve the underlying model to better capture the data in order to quantify the spin up of the high- $\alpha$  disc, we run a separate model in the following section that is able to let the mean velocity and velocity dispersion evolve over increasing metallicity bins.

#### 4.1. Quantifying the evolution of the high- $\alpha$ disc with metallicity

In order to quantify the spin-up of the high- $\alpha$  disc population, we modify the simple two-component model from the previous section to allow the mean velocity and velocity dispersion to vary between the spline knots across the metallicity range.

We fix the mean velocity of the halo component to be  $0 \text{ km s}^{-1}$  across the metallicity knots (the value converged from the HMC run using fixed means and standard deviations). This is fixed in order to separate the metal-poor high- $\alpha$  disc population (likely components of the proto-Galaxy, with a small but non-negligible prograde rotation) from the more radial halo population, as they heavily overlap. However, we let the standard deviation of the halo-like component vary across the metallicity knots with a truncated normal prior with a mean of  $100 \text{ km s}^{-1}$  and standard deviation of  $50 \text{ km s}^{-1}$ , with low/high limits restricted between  $50$  and  $200 \text{ km s}^{-1}$ .

It is important to note that the high- $\alpha$  selection is not perfect, and still captures the metal-poor end of accreted mergers (like GES). Therefore, fitting the halo-like component is still important in our high- $\alpha$  sample. The mean and standard deviation of the velocity of the disc-like component are set to vary with increasing metallicity. This will allow us to measure the spin-up of a proto-Galactic population to a rotation-dominated high- $\alpha$  disc. The metallicity range used for high- $\alpha$  stars with this model is between  $-2.0$  and  $+0.1$ . For the more metal-rich bins, we undersample the data to have an equal number of stars in each bin. The number of stars is set to the number in the lowest metallicity bin. This downsampling reduces the sample size for the most metal-rich bins, that in turn reduces the computation time.



**Fig. 6.** Fractional contribution of disc-like (purple) and halo-like (orange) GMM components as a function of metallicity for high-/low- $\alpha$  stars. The scatter points show the location of the knots chosen to run our model. The bands are  $1\sigma$  uncertainties on the converged weights. These fractions are computed for a fixed Gaussian for halo and disc-like stars when the MCMC is converged. The converged azimuthal velocity and velocity dispersion is shown in the legend, shown in units of km/s. The low- $\alpha$  model clearly shows two separate populations (thin disc and accreted) shown by the step function that could be misclassified as a rapid spin-up, while the high- $\alpha$  model shows a shallower monotonically increasing(decreasing) profile for the disc(halo) with  $[M/H]$ .

This also allows us to have equidistant knots, allowing the sampling of the metal-poor end as well as the metal-rich end (as opposed to the log-space knots used in the previous subsection, which sample the metal-poor end less than the metal-rich end). We choose 8 metallicity knots (equidistant in linear space) for the relative fraction and 9 metallicity knots (equidistant in linear space) for the velocity means and standard deviations. The mean velocity, velocity dispersion, and relative weights are computed for each knot in metallicity and spline interpolated for the data in-between the knots.

In order to measure the the spin-up of the disc-like component, and to enable the information of the disc-like component to be carried across metallicity bins, we implement a Gaussian process such that every finite collection of the azimuthal velocities (measuring the mean azimuthal velocity of the disc-like component) indexed by its metallicity has a multivariate normal distribution described by a Rational Quadratic Kernel function<sup>5</sup> described below:

$$\text{cov}_{\text{spin-up}}([M/H]_i, [M/H]_j) = \sigma^2 \left( 1 + \frac{([M/H]_i - [M/H]_j)^2}{2\alpha\ell^2} \right)^{-\alpha} \quad (3)$$

where  $\sigma^2$  is the overall variance,  $\ell$  is the characteristic length scale of the covariance function, that describes the range in  $[M/H]$  to which the covariance function typically holds, and  $\alpha$  is the positive scale-mixture parameter of the covariance function ( $\alpha > 0$ ), that in simple terms, describes the curvature of the covariance function. This function models the covariance between

<sup>5</sup> This is because Gaussian processes can be seen as an infinite-dimensional generalization of multivariate normal distributions.

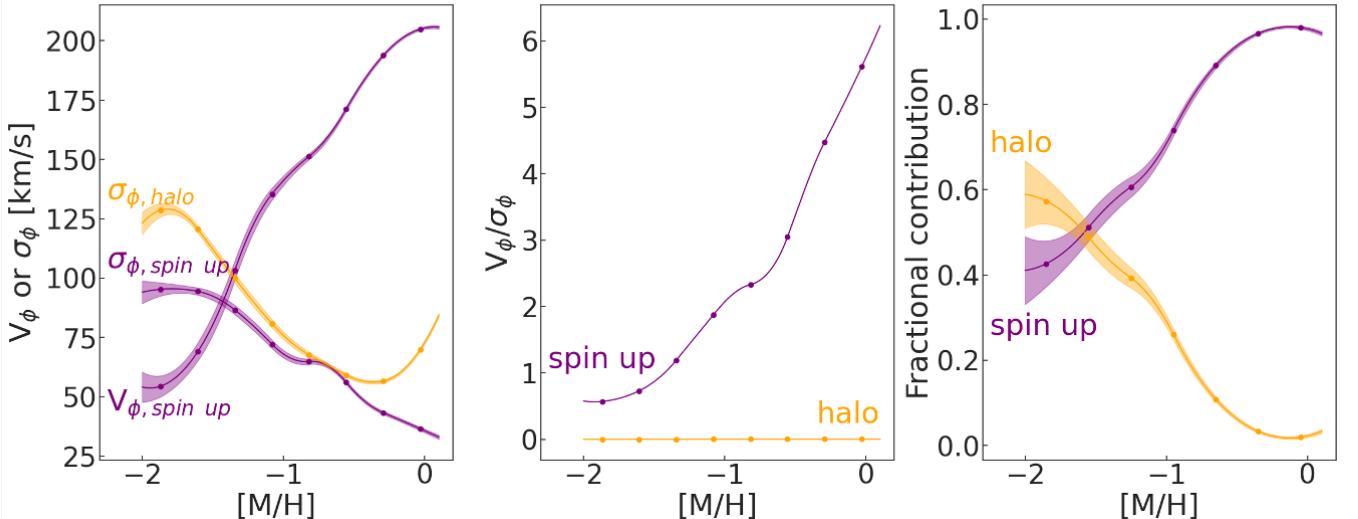
each pair of  $([M/H]_i, [M/H]_j)$ . The standard deviation on the mean of the azimuthal velocity describing the disc-like component has a uniform prior between 50 and 200, the characteristic length scale has a uniform prior between 0 and 1, and the scale-mixture parameter has a uniform prior between 0 and 4. All these parameters are sampled with the MCMC along with the means, standard deviations, and weights of the disc-like and halo-like components across the metallicity bins. The mean azimuthal velocity of the disc-like component has a multivariate normal distribution prior centered at  $120 \text{ km s}^{-1}$  described by the rational quadratic kernel function covariance matrix shown in equation 3. The mean azimuthal velocity is therefore defined as the mean function together with the kernel function that define the Gaussian process distribution of the azimuthal velocity of the disc-like component with varying metallicity. The standard deviation is described by a half normal prior centered at  $150 \text{ km s}^{-1}$ . The relative weight of the disc-like component is forced to be monotonically increasing using the `positive_ordered_vector` constraint on the argument, which automatically forces the halo-like component to be monotonically decreasing, removing noisy fits in the low-metallicity end.

The NUTS sampler with a HMC inference is run on the sampled data with the model described above for 100 warm-up steps, and 50 number of samples to generate from the Markov Chain for the high- $\alpha$  stars. We show the graphical model representation of this model in the right panel of Figure 5 and a list of model parameters and their functional forms are provided in the third column of Table 4.1. The high- $\alpha$  sample chains look much more stable and converged with this evolving velocity means and standard deviations model, they have  $r_{\text{hat}}$  below 1.1 (between 0.98 and 1.03), and  $n_{\text{eff}}$ , the number of effective sample is at least 30 or more for the means, standard deviations and relative

Parameter	Prior / Functional form		
	Frozen means model on [M/H]- $v_\phi$ plane	Evolving means model on [M/H]- $v_\phi$ plane	Evolving means model on [M/H]- $\eta$ plane
$w_{\text{knots}}$	Dirichlet( $0.5 \cdot 2$ )	Dirichlet( $0.5 \cdot 2$ )	Dirichlet( $0.5 \cdot 2$ )
$\sigma_{\text{mean,GP}}$	-	$\mathcal{U}(50 \text{ km/s}, 200 \text{ km/s})$	$\mathcal{U}(0.1, 0.9)$
$\ell_{\text{GP}}$	-	$\mathcal{U}(0, 1)$	$\mathcal{U}(0, 1)$
$\alpha_{\text{GP}}$	-	$\mathcal{U}(0, 4)$	$\mathcal{U}(0, 4)$
$\mu_{1,\text{knots}}$	-	$\mathcal{N}(80 \text{ km/s}, \text{cov}_{\text{spin up}})$	$\mathcal{N}(0.3, \text{cov}_{\text{spin up}})$
$\sigma_{1,\text{knots}}$	-	$\mathcal{HN}(150 \text{ km/s})$	$\mathcal{HN}(0.3)$
$\sigma_{2,\text{knots}}$	-	$\mathcal{TN}(100 \text{ km/s}, 50 \text{ km/s}; 50 \text{ km/s}, 150 \text{ km/s})$	$\mathcal{TN}(0.4, 0.2; 0.4, 0.7)$
$w$	$S_{\text{interp}}([M/H]   \overrightarrow{\text{knots}_w}, \overrightarrow{w_{\text{knots}}})$	$S_{\text{interp}}([M/H]   \overrightarrow{\text{knots}_w}, \overrightarrow{w_{\text{knots}}})$	$S_{\text{interp}}([M/H]   \overrightarrow{\text{knots}_w}, \overrightarrow{w_{\text{knots}}})$
$\mu_1$	$\mathcal{N}(0 \text{ km/s}, 150 \text{ km/s})$	$S_{\text{interp}}([M/H]   \overrightarrow{\text{knots}_\mu}, \overrightarrow{\mu_{1,\text{knots}}})$	$S_{\text{interp}}([M/H]   \overrightarrow{\text{knots}_\mu}, \overrightarrow{\mu_{1,\text{knots}}})$
$\sigma_1$	$\log_{10} \frac{\sigma_1}{\text{km/s}} \sim \mathcal{U}(0, 2.5)$	$S_{\text{interp}}([M/H]   \overrightarrow{\text{knots}_\mu}, \overrightarrow{\sigma_{1,\text{knots}}})$	$S_{\text{interp}}([M/H]   \overrightarrow{\text{knots}_\mu}, \overrightarrow{\sigma_{1,\text{knots}}})$
$\mu_2$	$\mathcal{N}(0 \text{ km/s}, 150 \text{ km/s})$	Fixed, 0 km/s	Fixed, 0
$\sigma_2$	$\log_{10} \frac{\sigma_2}{\text{km/s}} \sim \mathcal{U}(0, 2.5)$ $w \cdot \mathcal{N}(\mu_1, \sigma_1) +$ $(1-w) \cdot \mathcal{N}(\mu_2, \sigma_2)$	$S_{\text{interp}}([M/H]   \overrightarrow{\text{knots}_\mu}, \overrightarrow{\sigma_{2,\text{knots}}})$ $w \cdot \mathcal{N}(\mu_1, \sigma_1) +$ $(1-w) \cdot \mathcal{N}(\mu_2, \sigma_2)$	$S_{\text{interp}}([M/H]   \overrightarrow{\text{knots}_\mu}, \overrightarrow{\sigma_{2,\text{knots}}})$ $w \cdot \mathcal{FN}(\mu_1, \sigma_1; -1, 1) +$ $(1-w) \cdot \mathcal{FN}(\mu_2, \sigma_2; -1, 1)$
$v_\phi$ or $\eta$			

**Table 1.** Model parameters, their priors and functional forms for the three different models presented in this work. Acronyms -  $\mathcal{HN}$  = Half-normal distribution,  $\mathcal{TN}$  = Truncated normal distribution,  $S_{\text{interp}}$  = spline interpolation,  $\mathcal{FN}$  = Folded normal distribution. The parameters of  $\mathcal{TN}$  and  $\mathcal{FN}$  are  $(\mu, \sigma; \min, \max)$ . The subscript 1 and 2 stands for the disc-like and halo-like components respectively, with  $\mu$  and  $\sigma$  as the Gaussian mean and standard deviations in  $v_\phi$  or  $\eta$ , conditioned on the value of [M/H].  $w$  stands for the relative contribution of the disc-like component with the relative contribution of the halo-like component defined as  $(1-w)$ .  $\text{knots}_w$  and  $\text{knots}_\mu$  stands for the [M/H] spline knots location for the relative contribution parameter ( $w$ ) and the velocity/circularity means and standard deviations ( $\mu, \sigma$ ) respectively. The graphical representation of the models are shown in 5.

### Evolving means model on [M/H]- $v_\phi$ plane High- $\alpha$

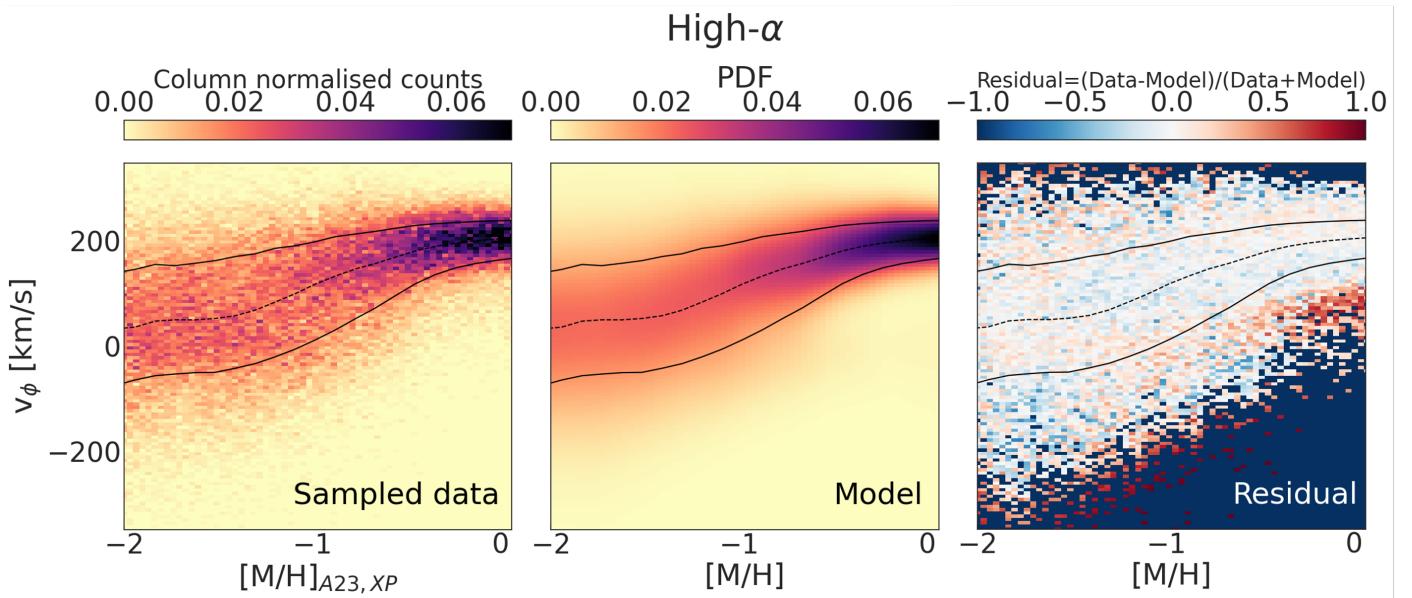


**Fig. 7.** High- $\alpha$  model with evolving means and standard deviation for azimuthal velocity of the disc-like component. Left: Evolution of azimuthal velocity and azimuthal velocity dispersion as a function of metallicity, capturing a monotonic transition of a chaotic dispersion-dominated state (predominantly made of proto-Galactic populations) to a rotationally supported state with disc-like motion in purple; the velocity dispersion evolution of halo is shown in orange. Center: Evolution of azimuthal velocity over azimuthal velocity dispersion,  $V_\phi/\sigma_\phi$  as a function of metallicity for the disc phase and halo component is shown in purple and orange respectively. Right: Fractional contribution of disc phase (proto-Galaxy to high- $\alpha$  disc) (purple) and halo-like (orange) GMM components as a function of metallicity. In all the panels, the scatter points show the location of the knots chosen to run our model. The bands are  $1\sigma$  errors on the converged weights, velocities and velocity dispersions. The fractions are computed for an evolving Gaussian for the disc phase and fixed Gaussian for the halo phase when the MCMC is converged.

weights, and about 20 for the covariance function parameters. The converged characteristic length is 0.75 (showing larger correlation between the velocities at different metallicities), and the scale-mixture parameter is 1.95. In summary, these results sug-

gest that this model provided posterior distributions that better describe the underlying high- $\alpha$  sample.

Figure 7 shows the converged parameters as a function of metallicity. As in Fig 6, the metallicity knots are shown as scatter points and  $1\sigma$  uncertainties on the converged parameters are



**Fig. 8.** Distribution of sampled data column normalised by sum (left), model colour-coded by the probability density function (center) and the normalised residual (right) in the  $[M/H]$ - $v_\phi$  plane (azimuthal velocity versus metallicity) for high- $\alpha$  evolving means model. The flat distribution of red stars at higher metallicities in the residuals show that the model is not able to fit the non-Gaussianity (long tail towards lower velocities) of the high- $\alpha$  disc. Systematic scatter in the periphery is due to random error in the data. However, for  $[M/H] < -0.7$ , the lack of strong systematic patterns, the low-amplitude and the small scatter of the residuals in the central regions (with the major bulk of the data) verifies the validity of the model. The median, 16<sup>th</sup> and 84<sup>th</sup> percentile tracks for our data are shown in black lines to guide the eye towards the bulk of the data.

shown as coloured bands. The disc-like component is shown in purple and the halo-like component is shown in orange. The left panel of Figure 7 shows the evolution of the mean azimuthal velocity and azimuthal velocity dispersion for the high- $\alpha$  disc-like component. Here, the trend gradually increases from  $v_\phi \sim 50$   $\text{km s}^{-1}$  at  $[M/H] \sim -2$  to  $v_\phi \sim 200$   $\text{km s}^{-1}$  at  $[M/H] \sim -0$ . This result can be interpreted as the high- $\alpha$  disc spinning up from a proto-Galactic population to a rotation dominated disc. While previous work have alluded to this in the literature (Chandra et al. 2023; Zhang et al. 2023), this is the first time that high- $\alpha$  proto-Galaxy-to-disc population has been quantitatively measured. In more detail, our results suggest that the proto-Galactic phase (pre-disc) lasts for approximately 0.5 dex (between  $-2$  and  $-1.5$  dex). After this, the proto-Galaxy populations gain azimuthal velocities as metallicity increases—in an almost linear fashion—until they settle into a disc at around  $-0.5$  dex. In summary, our results show that the so-called “spin-up” phase of the Galaxy happens gradually across a large range of  $[M/H]$ , starting from metallicities as low as  $[M/H] \sim -1.7$ .

Furthermore, throughout this phase, the velocity dispersion of the disc-like component decreases with metallicity, which also highlights how as the population gains  $v_\phi$ , it becomes less dispersion dominated. The velocity dispersion of the halo component (orange, that has fixed mean of  $v_\phi = 0$   $\text{km s}^{-1}$ ) is also decreasing with increasing metallicities. This could be due to different sub-structures dominating different metallicities. For example, we know that the GES merger has a lower velocity dispersion (of about 50  $\text{km s}^{-1}$ ) compared to the rest of the halo resting at about 100  $\text{km s}^{-1}$  velocity dispersion (see Figure 3). The rise in standard deviation at higher metallicities (when the high- $\alpha$  disc is in place) is not physical, but is rather caused by the model trying to make a broad halo component to fit the asymmetric tail of the high- $\alpha$  disc. This is one of the disadvantages of using a Gaussian distribution.

The middle panel of Figure 7 shows the ratio of  $v_\phi$  to  $\sigma_\phi$ , measuring how rotationally supported or ‘discy’ the stars are.

This ratio is basically zero (extremely pressure-supported) for the halo-like component, as we set the mean of the halo-like component to be close to 0  $\text{km s}^{-1}$ . However, the disc-like component has a clear rise in  $v_\phi/\sigma_\phi$ , reaching up to a ratio of 6 at solar metallicity. Moreover, the right panel of Figure 7 shows the relative contribution of halo-like (orange) and disc-like (purple) components at different metallicities. The accreted halo contribution decrease quickly with increasing metallicities. It dominates the distribution only at the lowest metallicity bins, below  $[M/H] \lesssim -1.5$ . In the very metal-poor end ( $[M/H] < -2$ ) of our high- $\alpha$  selection, the halo-like to disc-like component (i.e., halo-like to proto-Galaxy-like) ratio is 40%:60%. It is worth noting that we do not call this component “disc”, but simply refer to this component as “disc-like” for the modeling purpose. This component captures the proto-Galaxy to spin-up phase to high- $\alpha$  disc. The disc-like component’s fractional contribution increases slowly from  $[M/H] \sim -2$ , with an approximately constant gradient, up to  $[M/H] \sim -0.5$ . Upon reaching this point, the disc-like component (i.e., high- $\alpha$  disc) dominates the sample. All these results are highly in favour of a gradual spin-up for the disc-like component. In terms of the evolution of the Galaxy, our results highly favour the scenario where a proto-Galaxy population with low (but non-negligible)  $v_\phi$  profile spins up into a fully rotation-dominated high- $\alpha$  disc.

In Figure 8, we show the  $[M/H]$ - $v_\phi$  plane for high- $\alpha$  sample, with the sampled data in 2D column-normalised histogram on the left, the fitted model normalised with the integral under the curve equal to 1 (similar to column normalisation by sum) in the middle, and normalised residual (i.e., data – model) on the right. All the panels have the running 16<sup>th</sup>, 50<sup>th</sup>, and 84<sup>th</sup> percentile tracks from Figure 3 shown as black lines. We can clearly see that the sampled data follows the median tracks very well. We construct the model using the splines on the mean velocities, velocity dispersions, and fractional contributions for the two-component fits. The normalised residual is constructed by subtracting the probability density function of the sampled data

with the model's probability density function (PDF) in each cell (with bin sizes of  $6.25 \text{ km s}^{-1}$  in  $v_\phi$  and 0.04 in [M/H]). The residuals are scaled by the sum of sampled data and the model's PDF. This way, the residual only goes from -1 to 1 in value. If the residual is less than 0 the model predicts more stars than the data shows, and if the residual is greater than 0 the data has more stars in that region than the model predicts.

When inspecting Figure 8, the first thing one catches by eye in the residuals (right panel) is the horizontal patch of red stars at higher metallicities (between  $-0.5 < [\text{M}/\text{H}] < 0$ ). Here, the model predicts less stars than what is present in reality. We conjecture this is because the  $v_\phi$  distribution of high- $\alpha$  disc stars is asymmetric probably caused by the mechanism of asymmetric drift and is strongly non-Gaussian with a long tail towards lower  $v_\phi$  for this metallicity bin. Thus, the model is not able to capture well these stars. This asymmetric drift is stronger in high- $\alpha$  disc than the low- $\alpha$  disc. However, it is present in both populations (Anguiano et al. 2020). The same effect can also be seen in the model with frozen means and standard deviation for azimuthal velocity in the low- $\alpha$  sample. Due to this asymmetric tail towards lower velocities, we find that the model is underfitting the data by 22–26% at higher metallicities ( $[\text{M}/\text{H}] > -0.5$ ), while the under/overfitting is as low as 2–5% at lower metallicities.

The evolving mean high- $\alpha$  model presented here models well the bulk of the stars (also represented by the percentile tracks). However, we note that it struggles to fit the stars in the periphery of the  $v_\phi$  distribution (edge of the grid values in  $[\text{M}/\text{H}]$ - $v_\phi$  plane), mostly due to Poisson error. When compared to the residuals for the model with frozen means and standard deviation for azimuthal velocity in the high- $\alpha$  sample, the evolving means model has much smaller systematic effects. This is due to the frozen means model's underlying assumption of frozen mean velocity and velocity dispersion. Therefore, this improved model with evolving means is a much better representation of the high- $\alpha$  sample. Furthermore, it is important to note that in the residuals, the model is fully represented by a gradual (almost linear) rise in azimuthal velocity over the entire range of metallicities. If the data/model were better represented by a rapid and more exponential growth in  $v_\phi$  across a narrow range of metallicity – as reported in the literature—, we would find large systematic effects in our residuals (that are not seen). The absence of such systematic effects is more supporting evidence that the metal-poor high- $\alpha$  (or proto-Galaxy population) gradually spins-up to a rotation-supported high- $\alpha$  disc over a wide range of metallicities.

To our knowledge, this model is a first attempt at a simplified, physically motivated, and easily interpretable representation of the azimuthal velocity versus metallicity plane in the high- $\alpha$  regime, capturing the evolution of the first 5–6 Gyr of the of proto-Milky Way populations to the high-alpha disc. However, given the consequences of the simplified Gaussian distribution assumption, this model is a more qualitative representation in the metal-rich end ( $[\text{M}/\text{H}] > -0.8$ ).

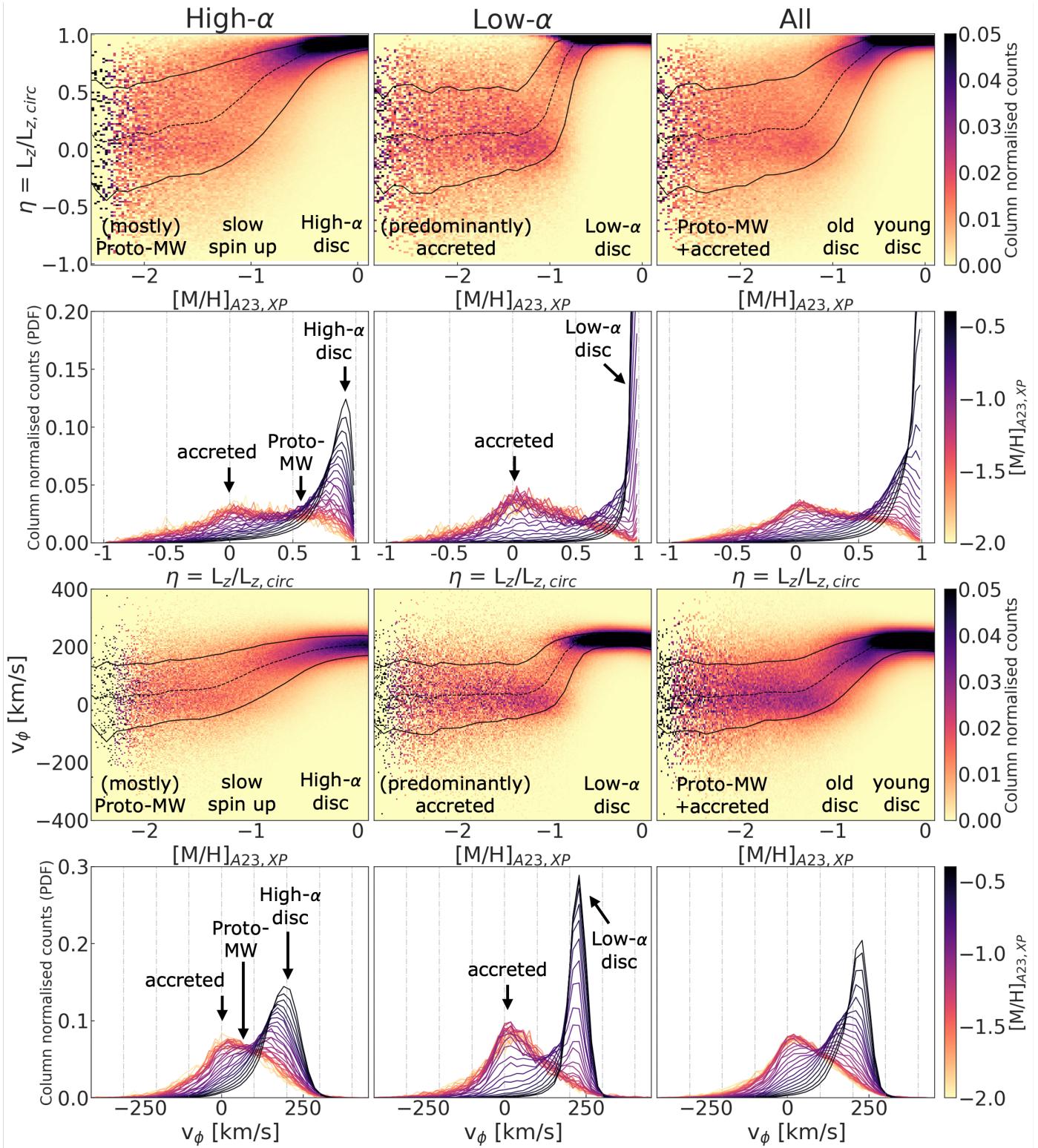
## 5. Evolution of orbital circularity with $[\text{M}/\text{H}]$

The orbital circularity metric,  $\eta$ , defined with respect to the current disc, ranging from -1 (perfectly retrograde, in-plane) to 1 (perfectly prograde, in-plane), interprets a star's orbital properties within the Galaxy. Intermediate values represent elliptical (closer to 0, isotropic, radial). This quantitative measure of orbital shape is very useful in understanding the dynamical processes that shaped the formation of the high- $\alpha$  disc. We compute orbital circularities for our sample of stars using the `gala` soft-

ware package (Price-Whelan 2017). We integrate the orbits of stars within a realistic Milky Way model for the Galactic potential, `MilkyWayPotential2022` (Price-Whelan et al. 2022), which incorporates four crucial components: a dense central core, a surrounding bulge of stars, a flattened disc of stars and gas, and a vast, spherical dark matter halo. This model is also calibrated to match observations of the Milky Way's rotation curve by Eilers et al. (2019) and a compilation of Milky Way's total mass enclosed (see Hunt et al. 2022). Using the adopted gravitational potential, we calculate the azimuthal angular momentum ( $L_{z,circ}$ ) and total energy ( $E$ ) for a perfectly circular orbit. By interpolating the resulting ( $L_{z,circ}(E)$ ) curve for each observed star's total energy, we determine the orbital circularity as  $\eta = L_z/L_{z,circ}$ .

The evolution of orbital circularity with respect to increasing metallicities is shown as 2D column normalised (by sum) histograms in the top panels of Figure 9 for high- $\alpha$ , low- $\alpha$ , and *all stars*. All the 2D histograms have 16<sup>th</sup>, 50<sup>th</sup>, and 84<sup>th</sup> percentile tracks of  $\eta$  versus  $[\text{M}/\text{H}]$  overlaid. To compare the orbital circularity with the azimuthal velocities used in this work, the 1D histograms of azimuthal velocities in bins of metallicities are shown in the middle panels of Figure 9 for the high- $\alpha$ , low- $\alpha$ , and all samples. We also show the 1D histograms of orbital circularity,  $\eta$ , in bins of metallicities in the bottom panels of Figure 9. The use of orbital circularity is very important because, unlike  $v_\phi$ , orbital circularity is position independent (especially at larger distances away from the solar neighbourhood). From the high- $\alpha$  panel of Figure 9, we can confirm the gradual evolution of circularity from a non-zero median circularity ( $\eta \sim 0.1$ ) metal-poor (halo-like) population to a rotation-dominated ( $\eta \sim 0.9$ ) disc-like component over a broad range of metallicities, ranging between  $-2.5 < [\text{M}/\text{H}] < -0.7$  (see the median tracks overlaid), if they are composed of a single stellar population. This is in favour of the gradual spin-up phase of the Milky Way's high- $\alpha$  disc. This can also be seen in the 1D histograms of  $v_\phi$  (second row) and  $\eta$  (third row), with the lighter colours (lower metallicities) having a bimodal distribution; this bimodal feature is likely attributed to the superposition of a halo-like population with no rotation and a proto-Galaxy-like population with small rotation (Horta et al. 2024). The bimodal distribution at lower metallicities is strikingly clear for orbital circularities as annotated in the bottom panels of Figure 9. Furthermore, the low- $\alpha$  panel of Figure 9 reveals that the low- $\alpha$  stars are made of two (almost) disjoint populations: an accreted halo and low- $\alpha$  (thin) disc. The halo dominates the lower metallicity bins, whilst the disc dominates the higher ones, as to be expected. This can be seen also in the median tracks (top panel of 9), which delineate a trajectory similar to a step function. The low- $\alpha$  1D histograms of  $v_\phi$  also show that the accreted halo (lighter colours, lower metallicities) dominates at  $v_\phi \sim 0 \text{ km s}^{-1}$  from  $[\text{M}/\text{H}] \sim -2$  to  $[\text{M}/\text{H}] \sim -1$ , without a decrease in the relative weight (i.e., the peak in the distribution stays relatively the same). After  $[\text{M}/\text{H}] \sim -1$ , the  $v_\phi$  distribution rapidly changes into a highly-rotating, disc dominated, population, that spans over 0.3 dex in metallicity (darker colours, higher metallicities).

In the *all stars* bottom panel of Figure 9, we see a compilation of the high- $\alpha$  and low- $\alpha$  components: accreted halo+proto-Galactic populations dominating the lower metallicities, high- $\alpha$  disc populations emerging from  $-1 < [\text{M}/\text{H}] < -0.7$ , and low- $\alpha$  disc populations taking over the distribution at metallicities higher than  $[\text{M}/\text{H}] \sim -0.6$ . The same picture can be deduced from the 1D histograms of  $v_\phi$ , that trace the same distribution as the circularity. Moreover, the median tracks on the 2D histograms show a rapid rise in circularity at around  $[\text{M}/\text{H}] \sim -1.0$ ,



**Fig. 9.** Top: Column-normalised (by sum) 2D histogram of stars in the  $[M/H]$ - $\eta$  plane (orbital circularity versus metallicity, top) and  $[M/H]$ - $v_\phi$  plane (azimuthal velocity versus metallicity, middle bottom) for all the stars (right), high- $\alpha$  selection (left), and low- $\alpha$  selection (center). The running median track is shown as dashed black line and the 16<sup>th</sup> and 84<sup>th</sup> percentile tracks are shown as black lines in all panels. The running median tracks for the *all stars* and low- $\alpha$  panels look more like a step-function, while high- $\alpha$  tracks is shallower, supporting the more gradual spin-up phase. (middle top) Probability density functions (PDF) of  $\eta$  in bins of  $[M/H]$  for all the stars (right), high- $\alpha$  selection (left), and low- $\alpha$  selection (center). (bottom) Probability density functions (PDF) of  $v_\phi$  in bins of  $[M/H]$  for all the stars (right), high- $\alpha$  selection (left), and low- $\alpha$  selection (center). Note the second peak with low net spin in high- $\alpha$  panel slowly gains rotation and spins-up into the old high- $\alpha$  disc in both middle top and bottom panels.

which is driven by a transition from the accreted (mostly GES) debris to the high-/low- $\alpha$  discs, similar to what is seen in the

low- $\alpha$  sample. However, in the *all stars* panel, the jump is not as sharp due to the presence of high- $\alpha$  disc populations. The conclusions are therefore consistent with the use of orbital circularity or azimuthal velocity. Given the position dependence on the azimuthal velocity, the use of orbital circularity brings more confidence that the gradual rise in rotational support for the high- $\alpha$  population is real.

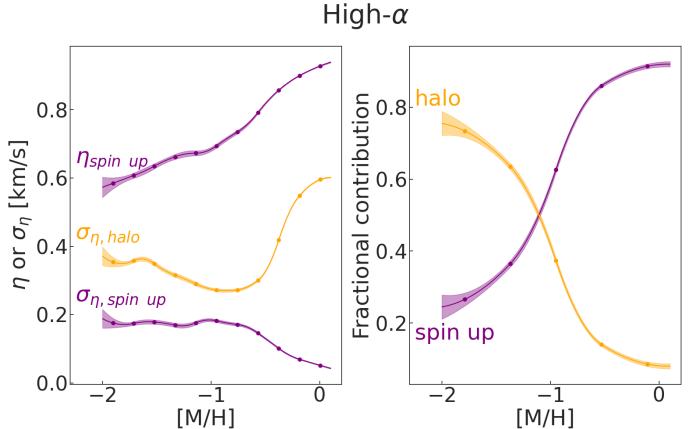
This result is important. If one does not account for the  $\alpha$ -separation like done in this work, a conclusion of the disc spin up in the  $v_\phi$ -[M/H] diagram could be interpreted as rapid, which we find is not the case (see Figure 9). The rapid transition from radial ( $v_\phi \sim 0$  km s $^{-1}$ ) orbits to circular ones ( $v_\phi \sim 200$  km s $^{-1}$ ) is caused by the presence of accreted populations and the low- $\alpha$  disc, and is only seen when examining low- $\alpha$  populations. On the contrary, when looking solely at high- $\alpha$  populations—which should trace directly the spin up of the old disc to the younger high- $\alpha$  disc—it is immediately clear that the relation in this diagram is much more gradual.

Previous work has attempted to look at the transition between hot/radial orbits and cold/circular ones using this  $\alpha$ -selection (Chandra et al. 2023). Thus, it is important that we compare our results to theirs. We argue that the reason these [M/H]- $\eta$  2D column-normalised histograms look strikingly different (especially for the high- $\alpha$  samples) from the study by Chandra et al. (2023) is because of the way the 2D histograms are plotted (both studies use the same data and similar  $\alpha$ -separation curves). Chandra et al. (2023) column-normalises their histograms by amplitude (tracing the mode of the distribution), while we column-normalise by their sum (tracing the underlying PDF). Column-normalisation by amplitude traces the mode of the distribution, which makes the whole 2D histograms more noisy (as the distribution is scaled by a factor of the standard deviation of the curve). Tracing the mode also means that in a bimodal distribution across a large range of metallicities (like for the high- $\alpha$  sample), the mode switches between one and the other rapidly within a small bin size in metallicity. This could lead to the data manifesting a sharp increase from  $\eta \sim 0$  to  $\eta \sim 1$ , as seen in Chandra et al. (2023), when in reality the data shows a more gradual increase (this work). Therefore, it is important to know how different normalisation methods can give rise to different interpretations of the underlying data and choose the normalisation method that best represents the science question that needs solving. In our case, as we aim to understand how the low-[M/H] regime transitions into the high-[M/H] regime for both high-/low- $\alpha$  populations, we choose to plot the sum column-normalised distribution. Furthermore, the differences with Chandra et al. (2023)'s results also arise from a small difference in the  $\alpha$ -separation. These normalisation and  $\alpha$ -selection differences and their implications are discussed in detail in Appendix B.

### 5.1. Modelling the high- $\alpha$ stars in the [M/H]- $\eta$ plane

Even though orbital circularity depends on the adopted Galactic potential, it is position independent as opposed to azimuthal velocity. This makes it valuable to model the circularity evolution across metallicities. Due to the non-Gaussian nature of orbital circularity (as it abruptly cuts at -1 and 1 for a perfectly circular retrograde and prograde orbit respectively), the models described earlier in this work using azimuthal velocities cannot be directly applied to orbital circularities. Therefore, the underlying Gaussians are modelled as folded normal distributions (folded at -1 and 1). The modelling is only performed on high- $\alpha$  stars because they trace the transition of the proto-Galaxy to rotation-

### Evolving means model on [M/H]- $\eta$ plane



**Fig. 10.** High- $\alpha$  model with evolving means and standard deviation for orbital circularity of the disc-like component. Left: Evolution of orbital circularity, and orbital circularity dispersion as a function of metallicity, capturing a gradual transition of a chaotic state (predominantly made of proto-Galactic populations) to a rotationally supported disc-like motion in purple and the orbital circularity dispersion evolution of halo is shown in orange. Right: Fractional contribution of disc phase (proto-Galaxy to high- $\alpha$  disc) (purple) and halo-like (orange) GMM components as a function of metallicity. In all the panels, the scatter points show the location of the knots chosen to run our model. The bands are  $1\sigma$  uncertainties on the converged weights, circularity and circularity dispersions. The fractions are computed for an evolving Gaussian for the disc phase and fixed Gaussian for the halo phase when the MCMC is converged.

supported high- $\alpha$  disc more clearly as seen in Figure 9, while the low- $\alpha$  stars can be simply described by two separate components similar to the [M/H]- $v_\phi$  plane.

The halo-like component has a fixed mean of 0 while varying standard deviation with a truncated normal prior with a mean of 0.4 and standard deviation of 0.2, with low/high limits restricted between 0.2 and 0.7 across the metallicity knots. We use the same metallicity range as in subsection 4.1, with 12 and 6 metallicity knots (equidistant in linear space and downsampled between each bins to have the same number of stars) for relative fractions and orbital circularity means and standard deviation. The definition of a covariance matrix to enable the information of between each components across the metallicity is unchanged in this circularity model, with the standard deviation on the mean of the circularity has a uniform prior between 0.1 and 0.9. The mean orbital circularity of the disc-like component is described by a multivariate normal distribution centered at 0.3 and the standard deviation is described by a half normal prior centered at 0.3. The final distribution of both the components are converted to a folded normal distribution to account for the folding at -1 and 1.

The model is run with 250 warm-up steps and 50 sample chains with the chains converged and  $r_{\text{hat}}$  less than 1.1 and effective samples above 30. We show the graphical model representation of this model in the right panel of Figure 5 and a list of model parameters and their functional forms are provided in the fourth column of Table 4.1. Figure 10 shows the converged parameters as a function of metallicity after the spline interpolation between the metallicity knots. We see that the orbital circularity of the disc-like component is steadily increasing with increasing metallicities over a large range of metallicities,  $[\text{M}/\text{H}] \sim -1.7$  to -1 and is still steadily increasing up to -0.5. However, the interpretation of metal-rich stars ( $[\text{M}/\text{H}] > -1.0$ ) is trickier given that the long asymmetric tail of the disc-like population is poorly fit

due to the assumption of the underlying distribution as a simple (folded) Gaussian distribution. For the metal-poor stars, we can more confidently say that the spin up phase is more extended over a large range of metallicities from  $[M/H] \sim -1.7$  to  $-1$ . The relative fractional contribution from the spin up (disc-like) and halo-like population also supports the slower evolution of circularity, as opposed to the rapid spin up shown in the literature (see for e.g., Chandra et al. 2023; Kurbatov et al. 2024). The major difference between the azimuthal velocity and orbital circularity evolution from the modelling perspective is (i) the metallicity at which the halo and spin up component crossover in the relative contribution, which is much more metal-rich in circularity ( $[M/H] \sim -1.2$ ) compared to velocity ( $[M/H] \sim -1.6$ ), and (ii) the mean orbital circularity is more circular ( $\eta \sim 0.57$ ) at lower metallicities ( $[M/H] < -1.5$ , also seen in the 1D histogram in Figure 9) than previously reported and more circular than what mean azimuthal velocities show at the same metallicities ( $v_\phi \sim 50$  km/s). Both of these differences can be explained due to the fact that circularity has less position dependence than velocities (given that our sample has stars outside the solar neighbourhood, 49% of our stars are at heliocentric distances larger than 3 kpc). This is because the mean azimuthal velocity reduces close to the inner Galaxy when compared to solar neighbourhood, making the mean azimuthal velocity smaller in value. This affects the metal-poor stars more, as metal-poor stars are centrally concentrated (Rix et al. 2022), especially in the high- $\alpha$  sample. The latter difference could also arise from the fact that in the literature, the metal-poor high- $\alpha$  stars are not modeled with both an accreted and *in situ* population, whereas we can clearly see a bimodal distribution in the 1D histograms in Figure 9, justifying our choice of modelling the leftover accreted halo stars in the high- $\alpha$  regime along with the proto-Milky Way-like population. Therefore, the evolution of orbital circularity over increasing metallicities shows that the transition from a slowly rotating population to high- $\alpha$  disc is more gradual and not as rapid at  $[M/H] \sim -1$ . However, an important caveat to mention is that the orbital circularity represents how circular the orbit of a star is compared to the present-day disc orientation, which does not necessarily trace the circularity at formation.

## 6. Discussion

In this section, we discuss the summary of our results and the implications of the gradual spin-up phase. We furthermore perform a simple GMM decomposition in bins of metallicities, to support the gradual spin-up phase scenario. Finally, we present the limitations and future scope of this work.

### 6.1. Summary of results

In this work, we have set out to model the azimuthal velocity and orbital circularity evolution of high-/low- $\alpha$  stars across metallicity space using *Gaia* XP element abundances, DR3 astrometry, and RVS radial velocities, to understand the transition phase of the proto-Galactic population to the high- $\alpha$  disc. At first, we model the conditional distribution  $P(v_\phi | [M/H])$  using a 2-component (disc-like and halo-like) Gaussian Mixture Model with frozen means and standard deviations for the azimuthal velocities for the high- and low- $\alpha$  stars. We find the inferred posterior matches better with the data for low- $\alpha$  stars better than high- $\alpha$  stars. The fractional contribution from disc-like and halo-like components look like a step function at  $[M/H] \sim -1$  for low- $\alpha$  stars, while the high- $\alpha$  stars have a relatively gradual rise over increasing metallicities in the fractional contribution

(Figure 6). Secondly, given that the high- $\alpha$  stars are not modeled well by the frozen means model, we model the conditional distribution  $P(v_\phi | [M/H])$  using a 2-component (disc-like and halo-like) Gaussian Mixture Model with evolving means and standard deviations for the azimuthal velocities for the high- $\alpha$  stars. From this exercise, we see that both the mean azimuthal velocity and fractional contribution of the disc-like component is gradually increasing starting from a  $v_\phi \sim 50$  km/s over increasing metallicities ( $-1.7 < [M/H] < -1.0$ ). Thirdly, we perform the same exercise for high- $\alpha$  stars in orbital circularity versus metallicity plane,  $[M/H] - \eta$ , given that the orbital circularity is position independent, as opposed to azimuthal velocity. From this exercise, we also see a gradual rise in orbital circularity starting from an  $\eta \sim 0.57$  over increasing metallicities for the disc-like component.

Using different flavours of these mixture models, we have provided several lines of evidence that the metal-poor high- $\alpha$  disc increases its average azimuthal velocity, orbital circularity and rotational support gradually and monotonically across a wide range of  $[M/H]$ , spanning approximately  $-1.7 < [M/H] < -1$ . These data favour the scenario of a gradual spin-up of the metal-poor high- $\alpha$  disc (likely the proto-Galaxy) to a rotationally-supported high- $\alpha$  disc at higher  $[M/H]$ . On the contrary, due to the superposition of the GES debris and the low- $\alpha$  disc in the low- $\alpha$  sample, the transition from metal-poor (halo) populations to metal-rich (disc) ones is much sharper, appearing almost like a step-function at  $[M/H] \sim -1$ . Due to the GES debris dominating the metal-poor sample for *all stars* in our data set, this yields a similar profile when inspecting the *Gaia* XP sample without any  $\alpha$  selection. Thus, our results highlight the importance of the  $[\alpha/M]$  selection for studying the azimuthal velocity evolution of the old Milky Way disc and to avoid the GES debris. This also suggests that the proto-Galactic debris gained rotation gradually, eventually settling into a high- $\alpha$  disc and the disc formation was not as rapid or dramatic across a narrow range of metallicities as previously reported.

### 6.2. On the spin up of the Milky Way disc

At the earliest cosmic times, the systems that seeded the formation of the Milky Way are conjectured to be manifested as a chaotic, unstructured, “proto-Galactic” population (e.g., Mowla et al. 2024), shaped by continual merging of building blocks that gave birth to the Galaxy’s ancient stars (Tumlinson 2010; El-Badry et al. 2018; Horta et al. 2024; Semenov et al. 2024; Xiang et al. 2024). Evidence of stellar populations associated with the proto-Milky Way have been proposed in the literature (Lucey et al. 2019; Arentsen et al. 2020; Reggiani et al. 2020; Horta et al. 2021; Belokurov & Kravtsov 2022; Ardern-Arentsen et al. 2024), with a clear picture now emerging thanks to the vast *Gaia* data (Rix et al. 2022). Expectations from cosmological simulations corroborate these observational findings (e.g., Horta et al. 2024; McCluskey et al. 2024; Semenov et al. 2024); these suggest that at present day, the remnants of proto-Milky Way fragments should present weak but systematic net rotation w.r.t. the Galactic disc ( $v_\phi \sim 50$  km/s), should primarily be concentrated towards the innermost Galactic regions ( $r \lesssim 10$  kpc), and should host stars with high  $[\alpha/Fe]$  abundance patterns, similar to the old Milky Way disc.

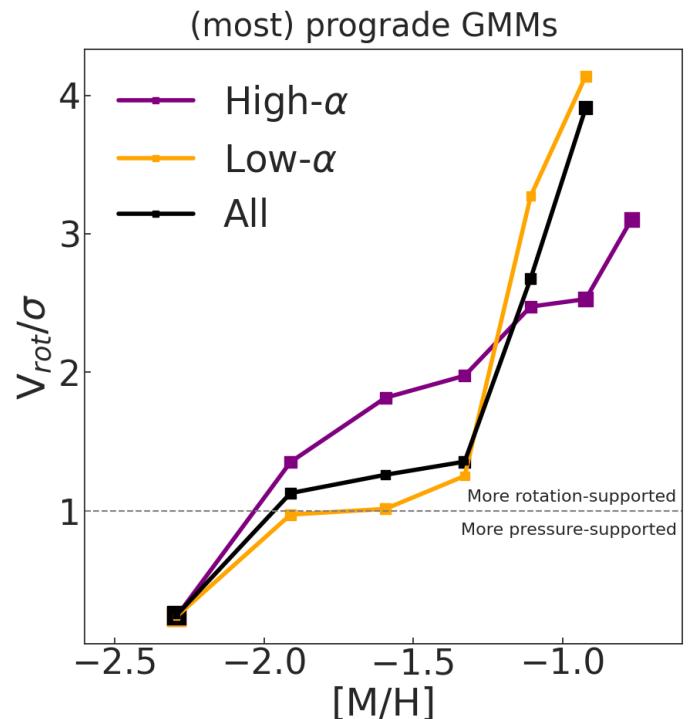
Ascertaining the transition from a dispersion dominated population to a rotationally supported one is key to unravel the genesis of the Galactic disc. Due to the inability to currently measure stellar ages precisely for metal-poor stars, metallicity ( $[M/H]$ ), a much more reliable stellar parameter estimate to determine, is

often used instead. While  $[M/H]$  is not a direct tracer of age, the reason why  $[M/H]$  is used is because, under the assumption that a stellar population chemically evolves in a progressive manner, stars that form later are more enriched in metals than their previous generations; in other words, old populations tend to be metal-poor while younger populations tend to be metal-rich. Thus,  $[M/H]$  has been shown to be a useful metric to study the evolution of kinematic samples (e.g., Belokurov & Kravtsov 2022; Xiang & Rix 2022; Xiang et al. 2024). This is especially the case in the current era of large-scale stellar surveys. For example, the vast *Gaia* data set of  $[M/H]$  (Andrae et al. 2023a) and now also  $[\alpha/M]$  (Li et al. 2024) for over two million stars has given an unprecedented (qualitative) view on the transition from the proto-Milky Way to the high- $\alpha$  disc (e.g., Chandra et al. 2023; Zhang et al. 2023). However, most of these studies either: (i) examine the profile of azimuthal velocity,  $v_\phi$ , and/or orbital circularity,  $\eta$ , with  $[M/H]$  agglomerating stars with different  $[\alpha/M]$ . This is known to be an issue due to the inclusion of stellar halo populations in the sample, that have different  $[M/H]$  and kinematic distributions; (ii) model the transition of dispersion dominated to rotationally dominated populations in the velocities- $[M/H]$  plane using disconnected Gaussian Mixture Models (see section 6.3), making it difficult to link components across  $[M/H]$ ; (iii) do not provide a quantitative measure of how the Milky Way transitioned from being dispersion dominated to rotationally dominated.

Belokurov & Kravtsov (2022) used high-resolution spectroscopic data from the APOGEE survey to study the evolution of stars around the Solar neighborhood they deem *in situ* in the  $v_\phi$ -metallicity plane; these authors found that the Milky Way “spun up” rapidly between  $-1.3 < [\text{Fe}/\text{H}] < -0.9$  (see also Zhang et al. (2023) for a similar result using *Gaia* XP metallicities). Along those lines, Chandra et al. (2023) used the available  $[\alpha/\text{Fe}]$  for the *Gaia* XP sample to study the profile of  $[\text{M}/\text{H}]\text{-}\eta$ , finding that the high- $\alpha$  disc emerges rapidly at  $[\text{M}/\text{H}] \sim -1$ .

In this work, we have found that when performing a similar  $[\alpha/M]$  cut to Chandra et al. (2023), our transition from proto-Milky Way populations to the high- $\alpha$  disc (Figure 9) shows a much more gradual increase when compared to these previous works. In more detail, both when examining and modelling the  $v_\phi$ - $[\text{M}/\text{H}]$  and  $[\text{M}/\text{H}]\text{-}\eta$  planes, we have found that the spin up of the old Milky Way disc occurs across a much wider range in  $[\text{M}/\text{H}]$ , namely between  $-1.7 < [\text{M}/\text{H}] < -1$ . The reason for the discrepancy between our results and previous work is because: (i) Belokurov & Kravtsov (2022) do not go down to lower metallicities ( $[\text{M}/\text{H}] < -1.5$ ) and have lower number statistics compared to our work. The 1D histograms of azimuthal velocities in Belokurov & Kravtsov (2022) already shows preliminary hints of a gradual spin up (see their Figure 5); (ii) Zhang et al. (2023) perform a traditional GMM without any  $\alpha$ -selection, resulting in the GES debris driving the rapid spin up inference at a narrow range in metallicities; (iii) (Chandra et al. 2023) shows a qualitative view of  $[\text{M}/\text{H}]\text{-}\eta$  plane, while their high- $\alpha$  sample is contaminated by GES debris more than ours, and their column normalisation is performed by amplitude instead of sum. Therefore, the results presented in this work, in comparison with the recent literature, shows that the proto-Galaxy to high- $\alpha$  disc transition has likely been gradual over increasing metallicities and not as dramatic over a narrow range of metallicities as previously reported.

We finish this section by raising a speculative, yet interesting, point. In Figure 7 and Figure 10, we find that the profile corresponding to the “spin up” displays a bump in the profile at  $[\text{M}/\text{H}] \approx -0.8$ . We have tested to ensure this bump is not ar-



**Fig. 11.**  $V_\phi/\sigma_z$  ratio (azimuthal velocity over vertical velocity dispersion) for the most prograde Gaussian component in each metallicity bin (spaced evenly on a log scale) from our GMM runs in subsection 6.3. Square size indicates the relative weight of the most prograde Gaussian in each metallicity bin. This ratio is a unitless quantity indicating the level of rotational support within a disc, essentially quantifying how dynamical cold it is at  $z = 0$ . The dashed line at  $V_\phi/\sigma_z = 1$  marks the boundary between pressure-supported (halo-like) versus rotation-dominated kinematics (disc-like).

tificially caused by our fitting methods. Interestingly, the location of this bump in  $[\text{M}/\text{H}]$  coincides with the end of the low- $\alpha$  accreted sample (primarily the GES merger debris). It is postulated that this population is the debris from a major merger in the Milky Way that dominates the local stellar halo ( $6 \lesssim r \lesssim 30$  kpc; e.g., Belokurov et al. 2018; Helmi et al. 2018; Koppelman et al. 2018). This merger possibly catalyzed the formation of the in-situ halo, or “hot high- $\alpha$  disc” comprising stars on halo-like orbits with chemistry similar to the high- $\alpha$  disc (Di Matteo et al. 2019; Bonaca et al. 2020; Belokurov et al. 2020). Thus, we speculate that this bump could be a sign of the impact of the GES with the old Milky Way disc.

### 6.3. Gaussian Mixture Model on 3D velocities in bins of metallicities

In this subsection, we perform the traditional  $N$ -component Gaussian Mixture Model (GMM) decomposition of high- $\alpha$ , low- $\alpha$ , and *all stars* subsamples in bins of  $[\text{M}/\text{H}]$  on the 3D velocity space in cylindrical coordinates -  $v_r$ ,  $v_\phi$ , and  $v_z$ . We set out to perform this exercise to understand how this method compares to our method presented in Section 4.

Typically, within any GMM framework, each sub-population is characterized by three key parameters. The first is a weighting factor, which reflects the relative proportion of stars belonging to that particular group. The second is the mean velocity, representing the first moment of the velocity distribution, and the average velocity along each of the three dimensions for stars within that

group. Finally, the second moment of the velocity, the velocity dispersion, captures the variation in velocities around the average for stars within that group. We use the same method outlined in Zhang et al. (2023) for the GMM decomposition, to be able to compare the effects of  $\alpha$ -separation in this type of GMM decomposition. To perform the GMM analysis, we utilized a software package named pyGMMs (Melchior & Goulding 2016). This software incorporates the sophisticated "Extreme Deconvolution" technique developed by Bovy et al. (2011) to account for the inherent uncertainties in stellar velocity measurements and the covariances between the input parameters.

Our analysis focuses specifically on metal-poor stars. We restricted the GMM decomposition to stars with a metallicity range of  $-2.5$  to  $-0.7$ , expressed as [M/H]. Below this we are hampered by low number statistics and above these metallicities we are affected by the non-Gaussian nature of the low- $\alpha$  and high- $\alpha$  discs distribution in velocity space that could compromise the accuracy of the GMM results. The [M/H] bins are spaced equidistantly on a log scale with seven(six) bins for the high-(low-) $\alpha$  samples. The *all stars* sample is fit using the same six bins as in the low- $\alpha$  sample<sup>6</sup>. Zhang et al. (2023) excluded stars located further than  $|z| > 2.5$  kpc from the Galactic plane, which is a cut that we do not use to be able to study the entire underlying data. We argue that as we are modelling the halo and disc populations separately, this cut is unnecessary. For each metallicity bin, we applied the GMM to the three-dimensional velocity space defined by the cylindrical Galactocentric coordinates: radial velocity ( $v_r$ ), azimuthal velocity ( $v_\phi$ ), and vertical velocity ( $v_z$ ). Moreover, each star's measurement uncertainty was incorporated as a covariance matrix with the diagonals representing the variances for each velocity component and the off-diagonals representing the covariance between a pair of velocities. To determine the optimal number of Gaussian components at each metallicity bin, we employed a standard approach based on the Bayesian Information Criterion (BIC), defined in the same way as Zhang et al. (2023). Here, lower BIC values indicate a better fit, balancing model complexity with data fidelity. To mitigate the risk of getting stuck in a local minima during the optimisation (sub-optimal solutions), we repeated the GMM fitting process 100 times with different starting conditions and recorded the BIC value for each trial. The fit with the lowest BIC value (in the median BIC curve) was considered the optimal solution, suggesting a high likelihood of finding the globally best fit. Increasing the complexity of the model by adding an extra component, while keeping the log-likelihood unchanged, raises the BIC value by  $\sim 100$ . Given this order of magnitude, some  $N$ -component GMMs have very similar BIC values, especially in relatively metal-rich regimes, where the rotating component is fit with multiple Gaussians instead of 1. When BIC values are similar, we prefer models with fewer components, as they offer a more straightforward physical interpretation. This occurs for at least 1 bin in high- $\alpha$  and 2 bins in low- $\alpha$  and *all stars* sub-samples. We do not compute the uncertainty in the GMM parameters as Zhang et al. (2023) found them to be generally around  $\sim 0.1$  kms<sup>-1</sup>. Thus, uncertainties in the fits can be treated as negligible.

In Figure 11, we show the ratio of azimuthal velocity to vertical velocity dispersion ( $V_\phi/\sigma_z$ ), which is commonly used as a measure of how rotation-supported or 'disc-like' a sample of stars is. We restrict our analysis to the most prograde GMM component (largest mean  $v_\phi$ ) per [M/H] bin to trace the evolution of

the Milky Way's disc; these are also the component with an almost zero mean  $v_r$  and  $v_z$ . In addition, to ensure that our model is capturing well the halo and disc components, we verify that the other GMM components the model finds fit for a halo-like substructure ( $V_\phi/\sigma_z \ll 1$ ). However, as we are only interested in the evolution of the disc-like GMM, we restrict our analysis to the most prograde GMM component in this work in Figure 11.

From Figure 11, we can see that the high- $\alpha$  stars gradually increase in their  $V_\phi/\sigma_z$  (in purple) towards higher metallicity bins while low- $\alpha$  (in orange) and *all stars* (in black) have a more rapid increase in  $V_\phi/\sigma_z$  between two [M/H] bins,  $\sim -1.3$  and  $\sim -1$ . The size of the scatter points are directly proportional to their relative weights in each bin. Therefore, our results suggest that not accounting for the  $\alpha$ -separation could make the spin-up seem more rapid and exponential over a small range of metallicities. In our case, as we have been able to distinguish high-/low- $\alpha$  populations, we find that the high- $\alpha$  prograde GMMs favour a gradual spin-up of a metal-poor high- $\alpha$  population (likely the proto-Galaxy) to a metal-rich high- $\alpha$  disc. The GMMs from Zhang et al. (2023) show a two-component fit in the most metal-poor bin while our model fits only one component. This is due to the fact that our *all stars* sample is composed of all stars with reliable metallicities from Andrae et al. (2023b), while having an  $\alpha$  estimate from Li et al. (2024), whereas Zhang et al. (2023) only used the metallicity estimates and therefore, we end up having much less VMP stars than theirs. Zhang et al. (2023)'s results also suggest a more rapid growth in  $V_\phi/\sigma_\phi$  over a shorter range in metallicity that is slightly more metal-poor than ours ( $-1.7 < [\text{M}/\text{H}] < -1.3$ ), which is explained by the difference in bin sizes and the scale height  $|z| < 2.5$  kpc cut. GMM decomposition can be highly dependent on the choice of the [M/H] bins. However, this GMM decomposition is performed in order to make a qualitative comparison between high- and low- $\alpha$  samples, and is not a one-to-one comparison to the literature results. The evolution of the velocity components and its dispersion focused on the high- $\alpha$  stars tracing the evolution of the proto-galaxy is shown in Appendix C.

In summary, our results reveal that the high- $\alpha$  subsample shows a gradual rise in  $V_\phi/\sigma_z$ , favouring a gradual spin-up phase for the Milky Way's high- $\alpha$  disc.

#### 6.4. Limitations and future scope

In this work, we use an  $\alpha$ -separation on *Gaia* XP based [ $\alpha/\text{M}$ ] and [M/H] abundances to separate high-/low- $\alpha$  stars and use the high- $\alpha$  stars to trace the azimuthal velocity ( $v_\phi$ ) evolution of the old Milky Way disc over a range of metallicities ( $-2.5 < [\text{M}/\text{H}] < 0.1$ ). The high- $\alpha$  selection implemented in this work is a simple piece-wise function (as described by equations 1 and 2), and is supposed to select only *in situ* stars. However, at lower metallicities  $[\text{M}/\text{H}] < -1.5$ , accreted mergers other populations (e.g., *Heracles* and the high-alpha tail of the GES and possibly other merger remnants) overlap, so a simple  $\alpha$ -cut can no longer establish a clean separation between *in situ* versus accreted population purely. This is modelled with our mixture model with evolving mean velocities and velocity dispersion for the spin-up phase and a background halo model that is isotropic as described in section 4.1 and 5.1. However, this model assumes a Gaussian distribution of azimuthal velocity at any given metallicity, which is not strictly true for the high- $\alpha$  disc, which has an asymmetric drift (long tail towards lower  $v_\phi$ , see Figure 9). Because of this, the model tries to inflate the halo velocity dispersion at higher metallicities to fit this tail of the high- $\alpha$  disc which is nonphysical. This also makes the model yield strong systematic residuals

<sup>6</sup> For the *all stars* and low- $\alpha$  stars samples, we neglect the most metal-rich bin as it is affected by the disc's asymmetry and suggests up to 8 components as the optimal fit, which is un-physical.

when compared to the data in the metal-rich end, where the high- $\alpha$  disc displays a strong prograde profile. This is a consequence of a simple Gaussian assumption that the model is based on. In the future, it would be possible to extend this model by using a more sophisticated dynamically-motivated distribution function that represents more accurately the Galaxy's disc/halo in order to fully measure the evolution of velocities from the chaotic proto-Galactic state to an ordered high- $\alpha$  disc state. In doing so, we could also use the total velocity distributions to constrain the enclosed mass and in turn, measure the mass growth of the Milky Way from a proto-Galactic state at high-redshift to the old high- $\alpha$  disc state as a function of increasing metallicity. We reserve this exploration for future work.

It is also important to note that the measured mean velocities and velocity dispersions with respect to decreasing metallicities are all present-day velocities and not the velocity they had at formation. From cosmological simulations, there have been clues that Galaxy must have undergone post-formation dynamical heating which could change the net rotation that we see at present-day (McCluskey et al. 2024; Horta et al. 2024).

One other factor to keep in mind is the ability to precisely capture the subtle trends in  $[\alpha/\text{Fe}]$  with our high- $\alpha$  sample, given our  $\alpha$  separation. Recent studies such as Belokurov & Kravtsov (2022); Conroy et al. (2022) used *APOGEE* and *H3* surveys, respectively, to show that the  $[\alpha/\text{Fe}]$  ratio gradually declines between  $-3 > [\text{Fe}/\text{H}] > -1.3$ , and then rises to a higher value to meet the high- $\alpha$  disc population. It is very likely that we are not catching this dip in  $[\alpha/\text{Fe}]$  with our simple  $\alpha$ -separation. However, this should not affect the conclusions of this work.

## 7. Conclusions and outlook

In this work, we set out to model the azimuthal velocity evolution of high- and low- $\alpha$  stars across metallicity space using *Gaia* XP element abundances, DR3 astrometry, and RVS radial velocities. By employing various mixture models, we have uncovered several lines of evidence that provide new insights into the kinematics of the metal-poor high- $\alpha$  disc and its evolution within the Milky Way.

Our analysis reveals that the metal-poor high- $\alpha$  disc shows a gradual and monotonic increase in its average azimuthal velocity distribution over a broad range of  $[\text{M}/\text{H}]$ , spanning approximately  $-1.7 < [\text{M}/\text{H}] < -1$ . This finding supports the scenario of a gradual spin-up of the metal-poor high- $\alpha$  disc, likely representing the proto-Galaxy, evolving into a rotationally-supported high- $\alpha$  disc as  $[\text{M}/\text{H}]$  increases. This gradual transition underscores the dynamic evolution of the early Milky Way, reflecting a continuous and progressive change in the rotational characteristics of the metal-poor high- $\alpha$  stellar population.

In contrast, the low- $\alpha$  sample presents a different picture due to the superposition of the *Gaia*-Enceladus-Sausage (GES) and other debris and the low- $\alpha$  disc. The transition from metal-poor (halo) populations to metal-rich (disc) populations in this sample is much sharper, resembling a step-function at  $[\text{M}/\text{H}] \sim -1$ . The dominance of the GES debris in the metal-poor sample for *all stars* in our dataset results in a similar profile when inspecting the *Gaia* XP sample without any  $\alpha$ -selection. This sharp transition highlights the distinct formation and evolutionary histories of these stellar populations compared to the more gradual evolution observed in the high- $\alpha$  disc.

Our results emphasize the critical importance of  $[\alpha/\text{M}]$  selection for studying the azimuthal velocity (or the orbital circularity) evolution of the old Milky Way disc. By distinguishing between high- and low- $\alpha$  populations, we can better understand

the complex interplay between different stellar populations and their contributions to the overall dynamics of the Galaxy. This distinction allows us to help disentangle the effects of accreted debris and in-situ star formation, providing a clearer picture of the Milky Way's formation history and its subsequent evolution. However, we note that this separation, while useful, does not lead to a purely in situ population.

In conclusion, our study provides strong evidence for the gradual spin-up of the metal-poor high- $\alpha$  disc, suggesting a continuous and progressive evolution from the proto-Galaxy to a rotationally-supported high- $\alpha$  disc, which has been recently reported in Horta & Schiavon (2024) using *APOGEE-Gaia* data's detailed chemical abundances and velocity distributions. The sharp transition observed in the low- $\alpha$  sample underscores the impact of accreted material on the Galaxy's dynamical structure. These findings contribute to our understanding of the early evolutionary processes of the Milky Way and highlight the need for detailed chemical and kinematic studies to unravel the complex history of our Galaxy. With the advent of more sophisticated distribution functions and increased availability of abundances for a larger number of stars, we will be better equipped to perform detailed analyses of in-situ stars, further unraveling the intricate history of the Milky Way and refining our understanding of its formation and evolution.

*Acknowledgements.* This work was developed for the CCA Pre-Doctoral Program in 2023 at the Flatiron Institute. We thank them for their generous support. AV thanks Vasily Belokurov for suggesting the test using high-resolution spectroscopy data. AV also thanks Ewoud Wempe for the helpful discussions on Gaussian processes, Tom Callingham for the helpful discussions on orbital circularity, Amina Helmi for the useful comments on this work, that helped improve our model, and Alis Deason for the insightful conversations on this topic. 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). This work has been partially supported by a Spinova Prize from NWO. 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. AV also thanks the availability of the following packages and tools that made this work possible: vaex (Breddels & Veljanoski 2018), pandas (Reback et al. 2022), astropy (Astropy Collaboration et al. 2022), NumPy (Oliphant 2006; Van Der Walt et al. 2011), SciPy (Jones et al. 2001), matplotlib (Hunter 2007), seaborn (Waskom et al. 2016), agama (Vasiliev 2019), gala (Price-Whelan 2017), galpy (Kluyver et al. 2016), healpy (Zonca et al. 2019), gaiadr3-zero point (Lindegren et al. 2021), jax (Schoenholz & Cubuk 2019), numpyro (Phan et al. 2019), scikit-learn (Pedregosa et al. 2011), pyGMMs (Melchior & Goulding 2016), JupyterLab (Kluyver et al. 2016), and topcat (Taylor 2018).

## References

- Andrae R., Rix H.-W., Chandra V., 2023a, *ApJS*, **267**, 8
- Andrae R., et al., 2023b, *A&A*, **674**, A27
- Anguiano B., et al., 2020, *AJ*, **160**, 43
- Ardern-Arentsen A., et al., 2024, *MNRAS*, **530**, 3391
- Arentsen A., et al., 2020, *MNRAS*, **491**, L11
- Astropy Collaboration et al., 2022, *ApJ*, **935**, 167
- Bailer-Jones C. A. L., Rybizki J., Fouesneau M., Demleitner M., Andrae R., 2021, *AJ*, **161**, 147
- Beers T. C., Christlieb N., 2005, *ARA&A*, **43**, 531
- Belokurov V., Kravtsov A., 2022, *MNRAS*, **514**, 689
- Belokurov V., Kravtsov A., 2023, *MNRAS*, **525**, 4456
- 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., 2020, *MNRAS*, **494**, 3880
- Bensby T., Feltzing S., Oey M. S., 2014, *A&A*, **562**, A71
- Bingham E., et al., 2019, *J. Mach. Learn. Res.*, **20**, 28:1
- Bonaca A., Conroy C., Wetzel A., Hopkins P. F., Kereš D., 2017, *ApJ*, **845**, 101
- Bonaca A., et al., 2020, *ApJ*, **897**, L18

- Bovy J., et al., 2011, [ApJ, 729, 141](#)
- Bovy J., Rix H.-W., Liu C., Hogg D. W., Beers T. C., Lee Y. S., 2012, [ApJ, 753, 148](#)
- Breddels M. A., Veljanoski J., 2018, *Astronomy & Astrophysics*, 618, A13
- Chandra V., et al., 2023, [arXiv e-prints](#), p. arXiv:2310.13050
- Chen B., Ting Y.-S., Hayden M., 2023, [arXiv e-prints](#), p. arXiv:2308.15976
- Chiba M., Beers T. C., 2000, [AJ, 119, 2843](#)
- Conroy C., et al., 2022, [arXiv e-prints](#), p. arXiv:2204.02989
- De Angelis F., et al., 2023, [A&A, 674, A2](#)
- Deason A. J., Belokurov V., 2024, [New A Rev.](#), 99, 101706
- 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](#)
- Dillamore A. M., Belokurov V., Kravtsov A., Font A. S., 2024, [MNRAS, 527, 7070](#)
- Eilers A.-C., Hogg D. W., Rix H.-W., Ness M. K., 2019, [ApJ, 871, 120](#)
- El-Badry K., et al., 2018, [MNRAS, 480, 652](#)
- Forbes D. A., 2020, [MNRAS, 493, 847](#)
- Frebel A., Norris J. E., 2015, [ARA&A, 53, 631](#)
- GRAVITY Collaboration et al., 2018, [A&A, 615, L15](#)
- Gaia Collaboration et al., 2023, [A&A, 674, A1](#)
- Gallart C., Bernard E. J., Brook C. B., Ruiz-Lara T., Cassisi S., Hill V., Monelli M., 2019, [Nature Astronomy](#), 3, 932
- Gallart C., et al., 2024, [A&A, 687, A168](#)
- Gilmore G., Reid N., 1983, [MNRAS, 202, 1025](#)
- Grand R. J. J., et al., 2020, [MNRAS, 497, 1603](#)
- Hayden M. R., et al., 2015, [ApJ, 808, 132](#)
- Haywood M., Di Matteo P., Lehnert M. D., Katz D., Gómez A., 2013, [A&A, 560, A109](#)
- Helmi A., 2020, [ARA&A, 58, 205](#)
- Helmi A., White S. D. M., 1999, [MNRAS, 307, 495](#)
- Helmi A., Babusiaux C., Koppelman H. H., Massari D., Veljanoski J., Brown A. G. A., 2018, [Nature, 563, 85](#)
- Horta D., Schiavon R. P., 2024, [arXiv e-prints](#), p. arXiv:2410.16374
- Horta D., et al., 2021, [MNRAS, 500, 1385](#)
- Horta D., et al., 2024, [MNRAS, 527, 9810](#)
- Hunt J. A. S., Price-Whelan A. M., Johnston K. V., Darragh-Ford E., 2022, [MNRAS, 516, L7](#)
- Hunter J. D., 2007, Computing in science & engineering, 9, 90
- Ibata R. A., Gilmore G., Irwin M. J., 1994, [Nature, 370, 194](#)
- Imig J., et al., 2023, [ApJ, 954, 124](#)
- Jones E., Oliphant T., Peterson P., et al., 2001, SciPy: Open source scientific tools for Python, <http://www.scipy.org/>
- Katz D., et al., 2023, [A&A, 674, A5](#)
- Kluyver T., et al., 2016, Jupyter Notebooks-a publishing format for reproducible computational workflows.. IOS Press
- Koppelman H., Helmi A., Veljanoski J., 2018, [ApJ, 860, L11](#)
- Koppelman H. H., Helmi A., Massari D., Price-Whelan A. M., Starkenburg T. K., 2019, [A&A, 631, L9](#)
- Kruijssen J. M. D., Pfeffer J. L., Crain R. A., Bastian N., 2019, [MNRAS, 486, 3134](#)
- Kurbatov E. P., et al., 2024, [arXiv e-prints](#), p. arXiv:2410.22250
- Li H., et al., 2022, [ApJ, 931, 147](#)
- Li J., Wong K. W. K., Hogg D. W., Rix H.-W., Chandra V., 2024, [ApJS, 272, 2](#)
- Lindgren L., et al., 2021, [A&A, 649, A4](#)
- Lucey M., et al., 2019, [MNRAS, 488, 2283](#)
- Majewski S. R., et al., 2017, [AJ, 154, 94](#)
- Martin, Starkenburg et al., 2023, [arXiv e-prints](#), p. arXiv:2308.01344
- Massari D., Koppelman H. H., Helmi A., 2019, [A&A, 630, L4](#)
- McCluskey F., Wetzel A., Loebman S. R., Moreno J., Faucher-Giguère C.-A., Hopkins P. F., 2024, [MNRAS, 527, 6926](#)
- Melchior P., Goulding A. D., 2016, pyGMMis: Mixtures-of-Gaussians density estimation method, *Astrophysics Source Code Library*, record ascl:1611.013
- Montalbán J., et al., 2021, [Nature Astronomy](#), 5, 640
- Mowla L., et al., 2024, [arXiv e-prints](#), p. arXiv:2402.08696
- Myeong G. C., Vasiliev E., Iorio G., Evans N. W., Belokurov V., 2019, [MNRAS, 488, 1235](#)
- Naidu R. P., Conroy C., Bonaca A., Johnson B. D., Ting Y.-S., Caldwell N., Zaritsky D., Cargile P. A., 2020, [ApJ, 901, 48](#)
- Nissen P. E., Schuster W. J., 2010, [A&A, 511, L10](#)
- Norris J., Bessell M. S., Pickles A. J., 1985, [ApJS, 58, 463](#)
- Oliphant T. E., 2006, A guide to NumPy. Trelgol Publishing USA
- Pedregosa F., et al., 2011, *Journal of Machine Learning Research*, 12, 2825
- Phan D., Pradhan N., Jankowiak M., 2019, arXiv preprint arXiv:1912.11554
- Price-Whelan A. M., 2017, [The Journal of Open Source Software](#), 2, 388
- Price-Whelan A., et al., 2022, [adrn/gala: v1.6.1, doi:10.5281/zenodo.7299506](#)
- Ratcliffe B., Minchev I., Cescutti G., Spitoni E., Jönsson H., Anders F., Queiroz A., Steinmetz M., 2024, [MNRAS, 528, 3464](#)
- Reback J., et al., 2022, pandas-dev/pandas: Pandas 1.4.0rc0, doi:10.5281/zenodo.5824773
- Reggiani H., Schlaufman K. C., Casey A. R., Ji A. P., 2020, [AJ, 160, 173](#)
- Rix H.-W., et al., 2022, [ApJ, 941, 45](#)
- Schoenholz S. S., Cubuk E. D., 2019, [arXiv e-prints](#), p. arXiv:1912.04232
- Schönrich R., Binney J., Dehnen W., 2010, [MNRAS, 403, 1829](#)
- Semenov V. A., Conroy C., Chandra V., Hernquist L., Nelson D., 2024, [ApJ, 962, 84](#)
- Sestito F., et al., 2019, [MNRAS, 484, 2166](#)
- Sestito F., et al., 2020, [MNRAS, 497, L7](#)
- Starkenburg E., et al., 2017, [MNRAS, 471, 2587](#)
- Starkenburg E., et al., 2018, [MNRAS, 481, 3838](#)
- Taylor M., 2018, arXiv preprint arXiv:1811.09480
- Tumlinson J., 2010, [ApJ, 708, 1398](#)
- Van Der Walt S., Colbert S. C., Varoquaux G., 2011, Computing in science & engineering, 13, 22
- Vasiliev E., 2019, [MNRAS, 482, 1525](#)
- Vasiliev E., Baumgardt H., 2021, [MNRAS, 505, 5978](#)
- Viswanathan A., et al., 2024, [A&A, 683, L11](#)
- Waskom M., et al., 2016, Seaborn: V0.7.0 (January 2016), doi:10.5281/zenodo.45133
- Xiang M., Rix H.-W., 2022, [Nature, 603, 599](#)
- Xiang M., Rix H.-W., Yang H., Liu J., Huang Y., Frankel N., 2024, [Nature Astronomy](#)
- Yuan Z., et al., 2020, [ApJ, 891, 39](#)
- Zhang H., Ardern-Arentsen A., Belokurov V., 2023, [arXiv e-prints](#), p. arXiv:2311.09294
- Zonca A., Singer L., Lenz D., Reinecke M., Rosset C., Hivon E., Gorski K., 2019, [Journal of Open Source Software](#), 4, 1298

## Appendix A: [M/H]- $v_\phi$ plane towards and away from the inner Galaxy

Examining the [M/H]- $v_\phi$  plane for low- $\alpha$  stars in the *APOGEE* (Figure 4) data, we can see that the halo-like (isotropic) population at low metallicity is almost fully disconnected from the higher metallicity disc-like component. However, this feature is not as immediately clear when inspecting the same stellar populations in the *Gaia* XP (Figure 3) data (there appears to be more of a connection between the halo-like and disc-like population for low- $\alpha$  stars). In this Appendix, we set out to investigate if this difference is due to the sample selection difference between *APOGEE* and *Gaia*, that could lead to more high- $\alpha$  stars contaminating our low- $\alpha$  star sample, by looking at the difference between the [M/H]- $v_\phi$  plane towards and away from the Galactic Center.

One possible reason for the discrepancy is that proto-Galactic fragments tend to be more centrally concentrated (Horta et al. 2024). As *APOGEE* is a near-infrared survey that can penetrate through prevalent dust extinction in the central regions of the Galaxy more easily, it is likely that we are probing the proto-Galaxy better with *APOGEE* than with the (optical) *Gaia* survey. Thus, if our  $\alpha$  cut was designed well, the low- $\alpha$  star sample should not be as centrally concentrated as the high- $\alpha$  population, that hosts both more centrally concentrated (high- $\alpha$ ) disc stars (Hayden et al. 2015; Imig et al. 2023) and proto-Galactic populations (Horta et al. 2024).

Figure A.1 shows the [M/H]- $v_\phi$  plane for all, low- $\alpha$ , and high- $\alpha$  stars towards Galactic longitudes in the direction of the Milky Way Centre (i.e.,  $|\ell| < 30^\circ$ , top panels), and towards Galactic longitudes away from the Galactic Centre (i.e., the Galactic anti-center,  $|\ell| > 30^\circ$ , bottom panels) for our *Gaia* XP sample. For low- $\alpha$  stars, we can see that the halo-like population at lower metallicities is fully disjoint from the disc-like population at higher-metallicities for stars towards the Galactic anti-center (similar to what is seen in *APOGEE*, Figure 4), whereas the transition in  $v_\phi$  across different metallicities is smoother towards the Galactic center, similar to high- $\alpha$  median tracks. We reason that this is because, when looking towards the Galactic anti-centre, the two dominant populations contributing to the low- $\alpha$  regime are: 1) the (outer) low- $\alpha$  disc that is highly rotating; 2) the debris from the GES merger, which are highly radial. This also affects the *all stars* panel for stars towards the Galactic center, as seen in the top panels of Figure A.1.

This leads us to infer that there exists a slowly rotating *in situ* proto-Galaxy-like population contaminating our low- $\alpha$  sample that is centrally concentrated, which is likely not present in the *APOGEE* stars. This arises mainly due to the higher abundance precision in *APOGEE* (high resolution spectroscopic survey) compared to *Gaia*-XP inferred  $\alpha$ -abundances.

In Figure A.2, we show the on-sky distribution of stars in Galactic coordinates in *Gaia* XP sample with a healpix level of 7 (top panels) and *APOGEE* sample with a healpix level of 4 (bottom panels) for both high- and low- $\alpha$  selection between the metallicities of  $-0.7$  and  $-1.3$ . We choose this metallicity range because this is the range between which we see the difference in the  $v_\phi$  median tracks between low- $\alpha$  stars towards and away from the Galactic center as shown in Figure A.1. In both *APOGEE* and *Gaia* XP high- $\alpha$  stars, we see centrally concentrated distribution of metal-poor stars, reminiscent of the poor old heart (Rix et al. 2022) a.k.a the proto-Galactic *in situ* population. However, we also see a higher density of stars towards the inner Galaxy in the *Gaia* XP low- $\alpha$  panel, which is not as clear an overdensity in the *APOGEE* low- $\alpha$  panel. Within the inner Galaxy ( $|\ell| < 30^\circ$  and

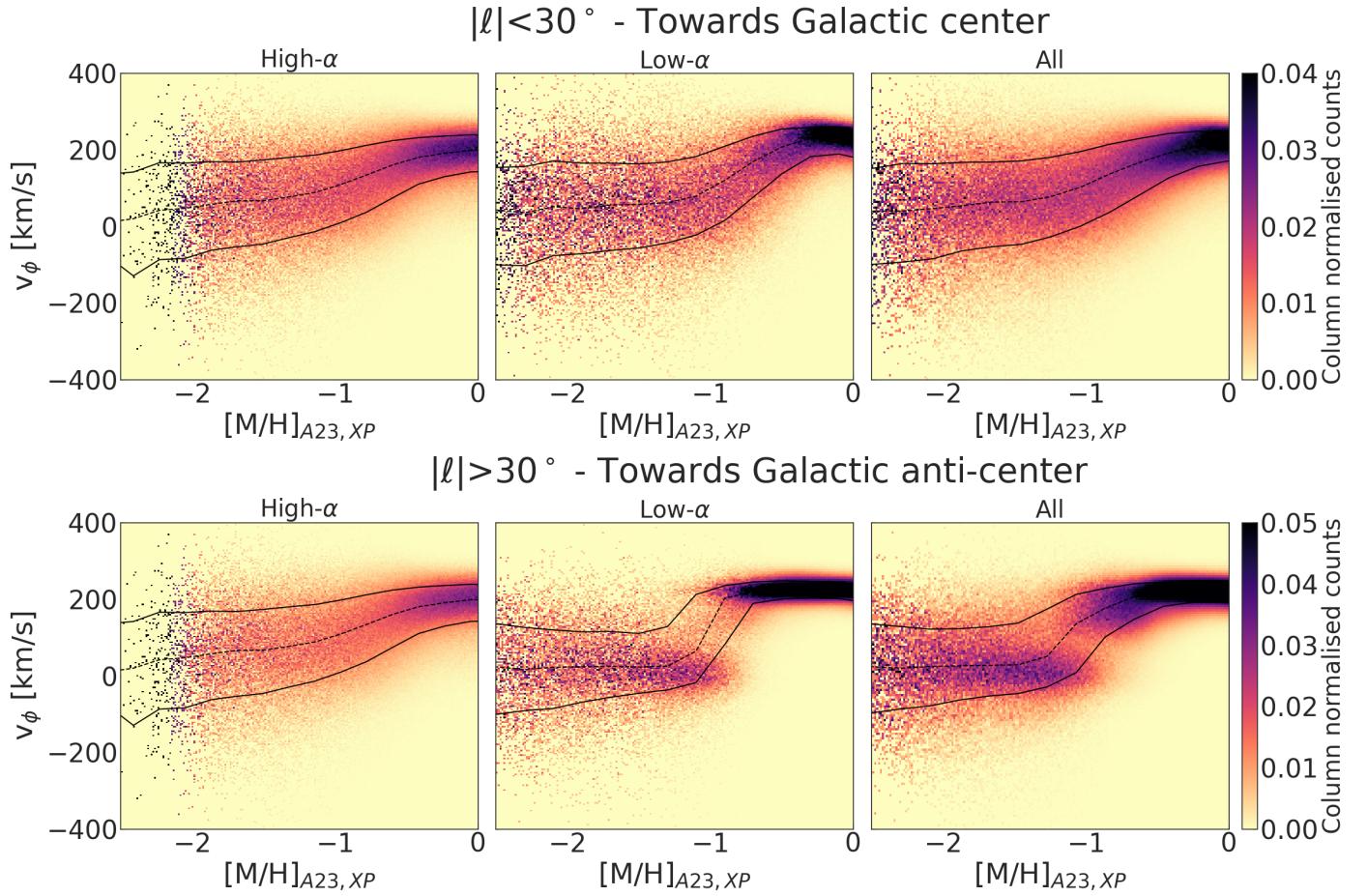
$|b| < 30^\circ$ ), the high- $\alpha$  stars are overdense compared to the low- $\alpha$  stars by a factor of 8 for *APOGEE* sample while the high- $\alpha$  stars are overdense compared to the low- $\alpha$  stars by a factor of 4 for the *Gaia* XP sample. Therefore, this mismatch between *APOGEE* and *Gaia* XP is not just due to lower number statistics in *APOGEE*. This difference could be due to unreliable  $\alpha$  estimates and simple definition of  $\alpha$ -separation. Therefore, *in situ* versus accreted separation using *Gaia* XP  $\alpha$  estimates is not as reliable and the low- $\alpha$  selection using *Gaia*-XP sample is more contaminated than the *APOGEE* sample, which is most likely the reason for a shallower step function in the low- $\alpha$   $v_\phi$  median tracks in *Gaia* XP (bottom left panel of Figure 3) compared to the steeper one in *APOGEE* (middle panel of Figure 4). However, we do not expect our conclusions to be affected by this contamination as our high- $\alpha$  (mostly *in situ*) selection is still pure.

## Appendix B: Comparison of orbital circularity versus metallicity with Chandra et al. (2023) results

In this section, we compare the evolution of orbital circularity versus metallicity in our work with the results from Chandra et al. (2023). Both the works use the same input catalogue based on *Gaia* XP spectra. The main differences between our approaches are the difference in the  $\alpha$ -separation, and the way the column normalisation is performed. In Figure B.1 top panels, we compare the column-normalised by amplitude 2D histograms of orbital circularity versus metallicity and their corresponding  $v_\phi$  1D histograms for Chandra et al. (2023)  $\alpha$ -selection and the  $\alpha$ -selection described in this work (equations 1 and 2). In all the panels, the 16<sup>th</sup>, 50<sup>th</sup>, and 84<sup>th</sup> percentile tracks are shown as black lines, mean tracks are shown as black dashed lines and mode tracks are shown as gray dashed lines. It is important to note that the mode tracks trace the underlying distribution the best when the column normalisation is calculated by the amplitude of the distribution. In Figure B.1 bottom panels, we compare the column-normalised by sum 2D histograms of orbital circularity versus metallicity and their corresponding  $v_\phi$  1D histograms for Chandra et al. (2023)  $\alpha$ -selection and the  $\alpha$ -selection described in this work (equations 1 and 2). This is equivalent to what is presented in the rest of this paper (tracing the PDF of each distribution, equivalent to normalising such that the area under the curve is equal to 1). In all the panels, the 16<sup>th</sup>, 50<sup>th</sup>, and 84<sup>th</sup> percentile tracks are shown as black lines.

There are two main inferences that can be made from this comparison, as discussed below:

- The  $\alpha$ -selection described in this work is more efficient in removing the last major merger (accreted GES) from the high- $\alpha$  selection (*in situ* equivalent), better than the  $\alpha$ -selection described by Chandra et al. (2023). Therefore, our high- $\alpha$  stars are cleaner than the high- $\alpha$  stars from Chandra et al. (2023).
- Column normalisation of 2D histograms can be done in many ways. From our comparison, we conclude that the column normalisation is scaled by the amplitude of the distribution in Chandra et al. (2023), whereas in this work, we column normalise the histograms by the sum of the distribution. Column normalisation by amplitude scales the peak of each 1D histogram to 1.0 whereas column normalisation by sum scales the sum of the histogram (equivalent to integral under the histogram curve) to be equal to 1.0 thereby tracing the probability distribution function of each 1D histogram. Column normalisation by amplitude traces the peak



**Fig. A.1.** Column-normalised (by sum) 2D histogram of stars in the  $[M/H]$ - $v_\phi$  plane for all the stars (right), high- $\alpha$  selection (left), and low- $\alpha$  selection (center) towards the Galactic center (top panels) and away from the Galactic center (anticenter, bottom panels). The running median track is shown as dashed black line and the 16<sup>th</sup> and 84<sup>th</sup> percentile tracks are shown as black lines in all panels. We can see that the low- $\alpha$  stars are not fully two separate populations towards the inner Galaxy due to contamination from the proto-Galaxy in the low- $\alpha$  end, that creates a small connection between the thin disc and halo populations in the overall  $[M/H]$ - $v_\phi$  plane for low- $\alpha$  plane as seen in Figure 3. Low- $\alpha$  stars towards the galactic anti-center show a cleaner, two separate population of thin disc and halo stars as expected. This effects are also reflected in the *all stars* panels. This effect is minimal in the APOGEE  $\alpha$ -selection most likely due to the high resolution  $\alpha$ -measurements.

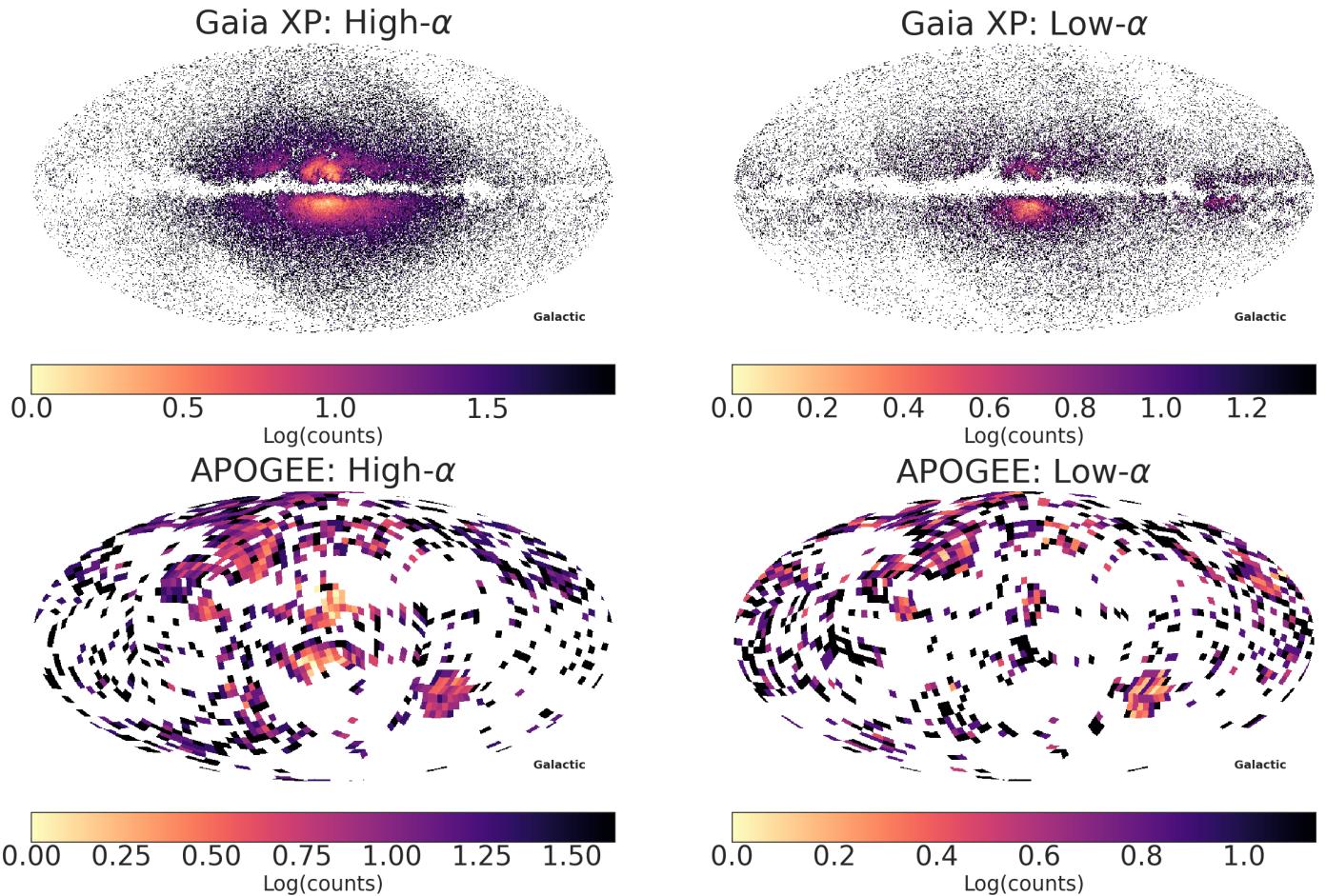
of each 1D histogram with no easily interpretable connection between each 1D histograms, and therefore produces a noisy view of the underlying data. This can be seen by the noisy streaks in the 2D histograms and the noisy mode tracks in the top panels of Figure B.1. Column normalisation by amplitude traces the mode of the distribution, which is also tricky in case of bimodal distributions. The metal-poor end of our underlying azimuthal velocity is bimodal due to a slowly-rotating proto-Galactic population and the high- $\alpha$  remnants of accreted stellar systems (mostly isotropic). The mode simply traces the peak of the distribution and therefore looks like a step function when one Gaussian dominates over the other. This affects the high- $\alpha$  population the most as the low- $\alpha$  stars are already two almost disjoint distributions in metallicities (thin disc at higher metallicities and accreted halo at lower metallicities). Therefore, we emphasize that column normalisation by sum is more appropriate to understand the underlying distribution of orbital circularity or azimuthal velocity as it traces the PDF of the distribution and not just the peak/the mode.

These two reasons together explain why one would interpret the spin up to be rapid and drastic, whereas the underlying dis-

tribution is slowly gaining rotation over a wide range of metallicities.

In order to better understand the difference in the  $\alpha$ -separation between our work and Chandra et al. (2023), we also show row normalised 2D histograms of  $[M/H]$ - $\eta$  space for high- $\alpha$  stars in this work and high- $\alpha$  stars from Chandra et al. (2023) in the top and bottom panels of Figure B.2. Because of row-normalisation, higher metallicity stars are more highlighted in these figures. In both these panels, we can see the highly rotating high- $\alpha$  disc dominating higher metallicities ( $[M/H] > -0.4$ ). Below these metallicities, we see high- $\alpha$  stars with a broad range of circularities, even down to retrograde orbits, but have metallicities that are representative of high- $\alpha$  disc. The most plausible origin for these stars are that they were born in the old high- $\alpha$  disc and got kicked-up into halo-like orbits by the last major merger, GES (Bonaca et al. 2017; Helmi et al. 2018; Belokurov et al. 2020). The most interesting difference between our high- $\alpha$  stars and that of Chandra et al. (2023) is that their high- $\alpha$  stars have larger number of stars in isotropic orbits around  $[M/H] \sim -1.2$ , reminiscent of the GES merger. This is almost absent in our high- $\alpha$  stars. This leads us to believe that our high- $\alpha$  star selection is purer than the one in Chandra et al. (2023). It is easier for the eye to trace the excess of retrograde low-metallicity stars in both the pan-

$$-1.3 < [\text{M}/\text{H}] < -0.7$$



**Fig. A.2.** Logarithmic Galactic map our sample of stars with high- $\alpha$  (left) and low- $\alpha$  (right) selections for the all-sky Gaia XP sample (top) and APOGEE DR17 sample (bottom), in the metallicity regime where the gradual spin-up in high- $\alpha$  versus the step function due to two different stellar populations in higher and lowe metallicities in low- $\alpha$  is reflected the most:  $-1.3 < [\text{M}/\text{H}] < -0.7$ . We can see that the poor old heart/the proto-Galaxy is still present (contamination) in low- $\alpha$  towards the inner Galaxy, however, to a smaller extent than in high- $\alpha$ . This is not the case for APOGEE stars which have a much cleaner low- $\alpha$  population towards the inner Galaxy with the bulk of proto-Galaxy towards the inner Galaxy in the high- $\alpha$  Galactic map.

els, but these are only highlighted due to the row-normalisation (because retrograde stars are almost only present due to halo accretion events in the lower-metallicity end) and in reality, these stars are much less in number. However, we still see a population of accreted stars in our high- $\alpha$  selection (to a much lower extent than using Chandra et al. (2023)  $\alpha$ -separation), which can be attributed to the evolution of any stellar system that has a high- $\alpha$  low-metallicity tail which cannot simply be removed using a simple  $\alpha$ -separation. We model this population along with the evolution of high- $\alpha$  disc in subsection 4.1 and subsection 5.1.

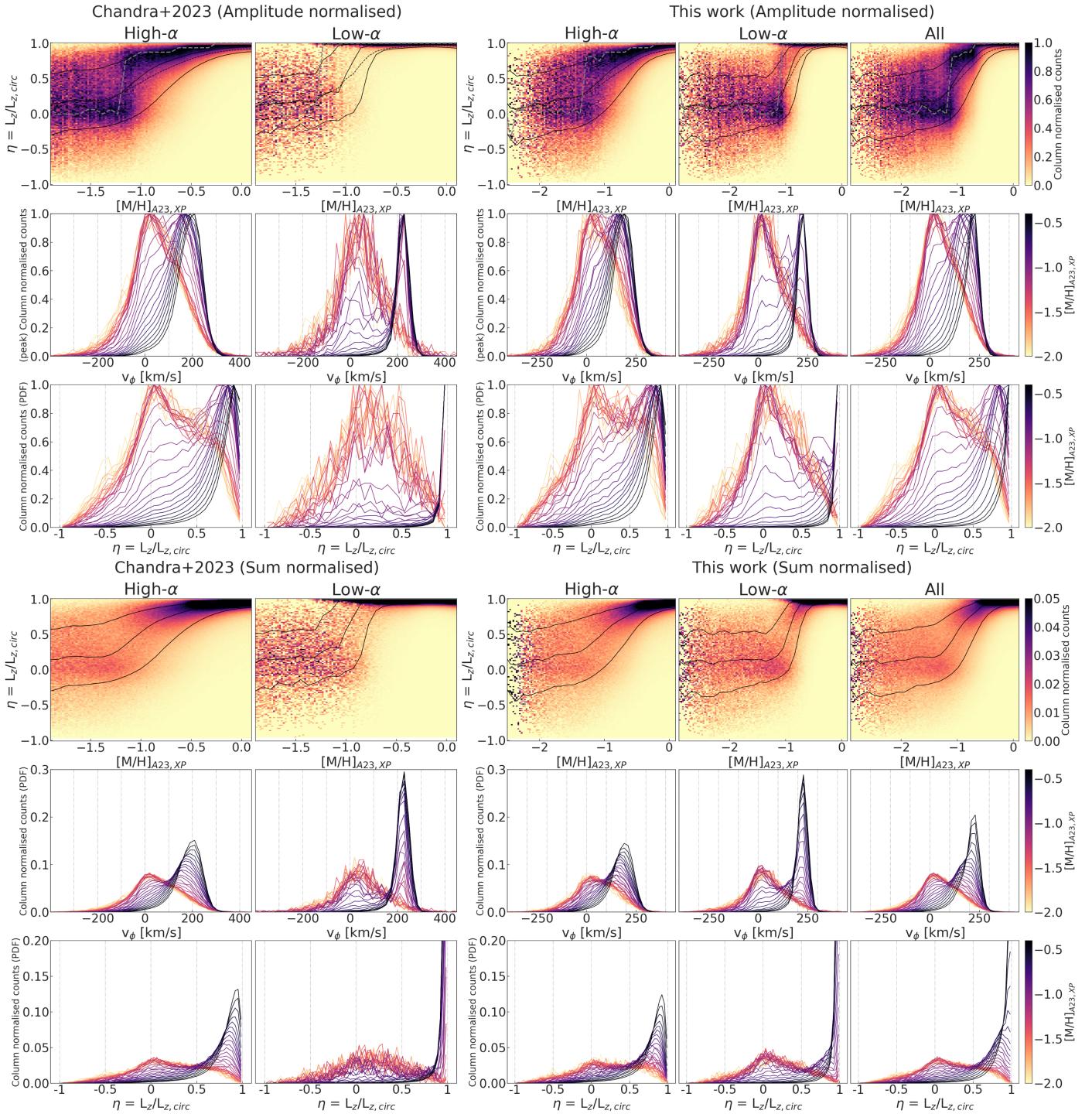
### Appendix C: Velocity evolution versus metallicity for high- $\alpha$ stars used as a cosmic clock

In this section, we show the evolution of the most prograde GMM component (based on the GMM runs explained in subsection 6.3) in its different velocity components for high- $\alpha$  stars. In Figure C.1, we see the evolution of  $(V_r, V_\phi, V_z)$  in the top panel, the evolution in their velocity dispersion  $(\sigma_r, \sigma_\phi, \sigma_z)$  in the middle panel and their rotational support  $(V_\phi/\sigma_z$  or  $V_\phi/\sigma_{tot}$ ) in the bottom panel across different [M/H] bins with the respective

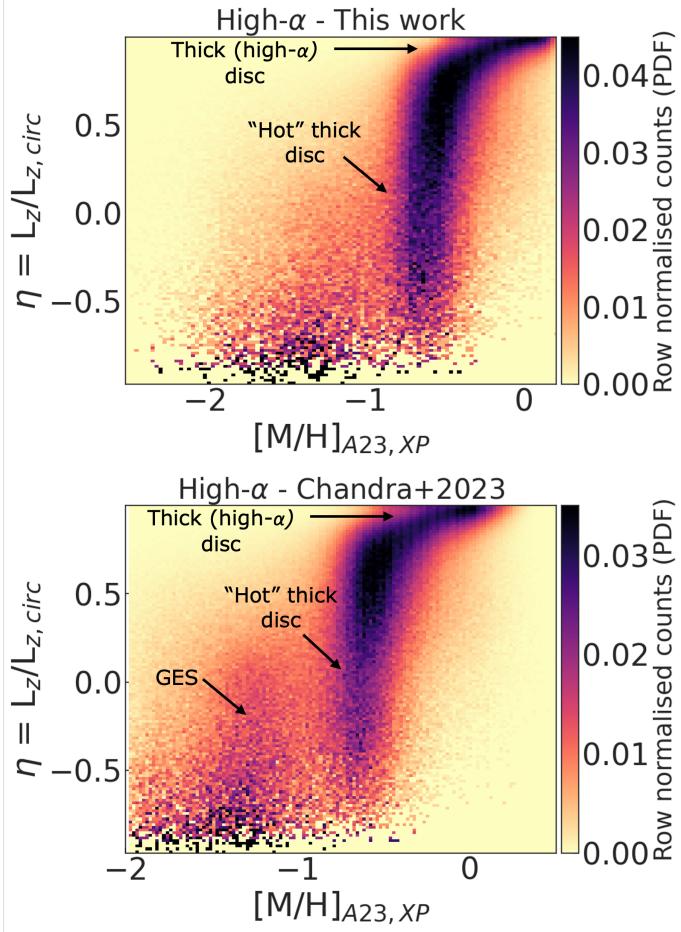
component names next to each curve. It is important to note that these velocities and their dispersions trace the present dynamical evolution and not those at formation. Using FIRE-2 simulations, McCluskey et al. (2024) shows that the rotational velocities can increase now compared to at formation in the pre-disc era due to stars being torqued into rotational orbits as the disc settles, and the rotational velocities can decrease now compared to at formation in the late-disc era due to dynamical heating. In case of rotational velocity dispersion, it also does not directly reflect the formation history, as it monotonically increases due to post-formation dynamical heating, adding to the velocity dispersion at formation. Therefore, even though metal-poor stars trace old stellar populations, the velocities do not simply trace the formation velocities. However, this simple analysis of velocity evolution can give us an idea of the evolution of the high- $\alpha$  disc as we see it now.

In the top panel of Figure C.1, we see that the radial velocity and vertical velocity is almost close to zero, with the rotational velocity (and the total velocity) increasing slowly with increasing metallicities. This is reminiscent of a slowly rotating proto-Galactic population that settles into a high- $\alpha$  disc at higher metallicities. We see that this spinning-up phase is more gradual

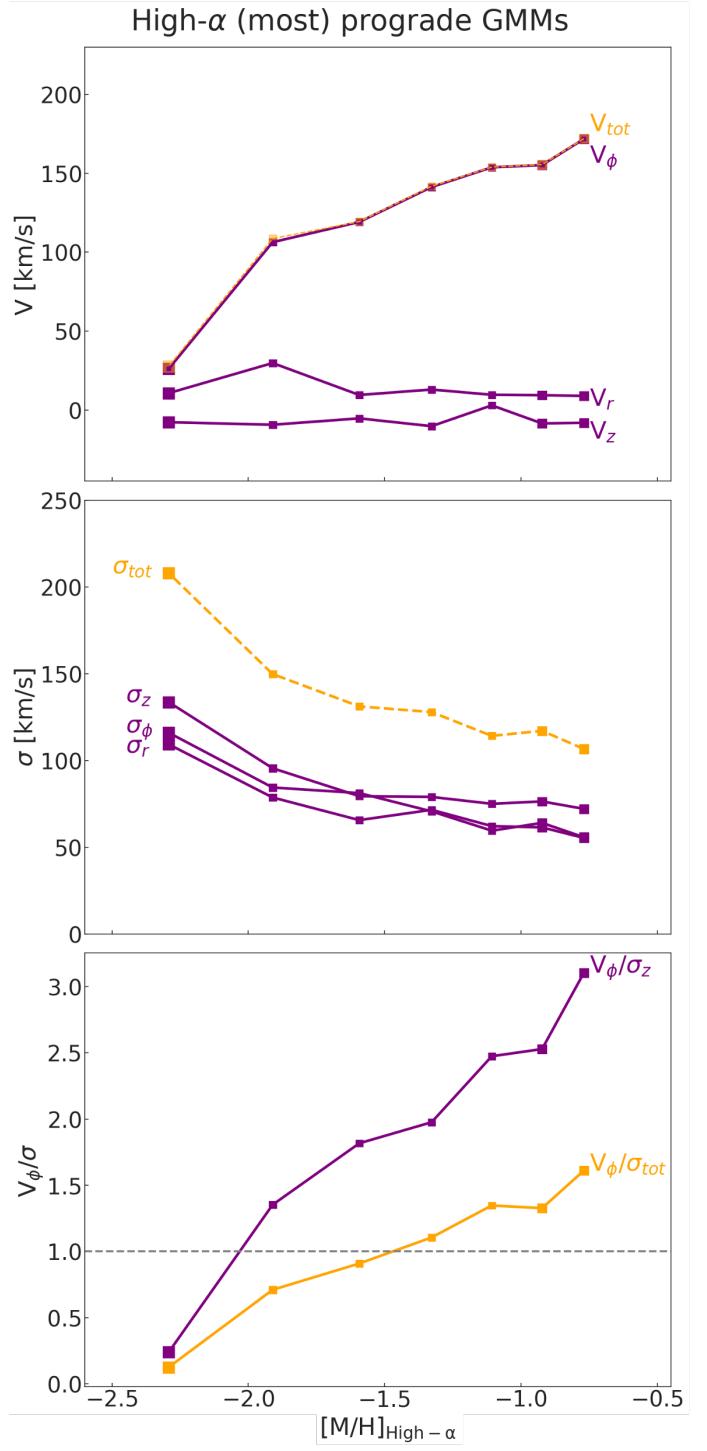
than previously reported. However, if this gradual spin-up comes from the time of its formation or due to post-formation heating is an open question. In the middle panel of Figure C.1, we see the trends for velocity dispersions in all three directions decreasing with increasing metallicities, as the high- $\alpha$  disc begin to settle. In the bottom panel of Figure C.1, we see the evolution of rotational support as a function of metallicity. McCluskey et al. (2024) show that in late- and early-disc era, both  $v_\phi/\sigma_z$  and  $v_\phi/\sigma_{tot}$  decrease by a factor of 2 between formation and present state, due to post-formation dynamical heating that increases their velocity dispersions. In the bottom panel of Figure C.1, we see the evolution of rotational support gradually increasing with increasing metallicity, also reminiscent of a proto-Galactic population slowly gaining rotation and settling into the high- $\alpha$  disc. The metallicity at which this population becomes more rotation supported ('discy') is difficult to infer, as it differ between  $\sigma_z$  and  $\sigma_{tot}$ , and also that the overall  $V_\phi/\sigma_{tot}$  reduces at present when compared to what it was at formation. Therefore, our main conclusion from these velocity evolution curves is that the spin-up phase of a slowly-rotating proto-Galaxy is gradual across a large range of metallicities and not as rapid as previously reported.



**Fig. B.1.** Orbital circularity versus metallicity for high- $\alpha$ , low- $\alpha$  and *all* stars, and 1D histogram of azimuthal velocity and orbital circularity at different metallicity bins for column-normalisation by amplitude (top panels) and by sum (bottom panels) using Chandra et al. (2023)  $\alpha$ -selection (left panels) versus the  $\alpha$ -selection implemented in this work (right panels). Running median (black line), mean (black dashed line), mode (gray dashed line), 16<sup>th</sup> and 84<sup>th</sup> percentile (black lines) tracks are overlaid for the 2D histograms with column normalisation by amplitude. We can see that mean and median tracks follow each other closely except at higher metallicities where the non-Gaussianity of thin and thick discs dominates, whereas the mode tracks follow the peaks of the background and is very noisy and resembles the step function behaviour seen by Chandra et al. (2023), even for high- $\alpha$  stars where we see a gradual spin-up using column normalisation by sum. Running median, 16<sup>th</sup> and 84<sup>th</sup> percentile tracks are overlaid as black lines for the 2D histograms with column normalisation by sum. We also see that our  $\alpha$ -selection has a cleaner high- $\alpha$  sample and isolates the bulk of GES into the low- $\alpha$  sample.



**Fig. B.2.** Row-normalized conditional metallicity distribution for high- $\alpha$  stars across orbital circularities for our  $\alpha$ -selection (top) and [Chandra et al. \(2023\)](#)  $\alpha$ -selection (bottom). The high- $\alpha$  disc, in-situ halo (hot high- $\alpha$  disc), and accreted components can be seen in both panels. However, contamination from accreted GES merger is stronger in [Chandra et al. \(2023\)](#)  $\alpha$ -selection than the  $\alpha$ -selection proposed in this work.



**Fig. C.1.** Mean velocities (top), velocity dispersions (middle) and  $V_\phi/\sigma_z$  ratio (bottom) showing how rotationally supported the stars are, for the most prograde Gaussian component in each metallicity bin from our GMM runs for high- $\alpha$  stars.