A&A 676, A140 (2023)
https://doi.org/10.1051/0004-6361/202346325

Astronomy & Astrophysics

Galactic ArchaeoLogIcaL ExcavatiOns (GALILEO)

II. t-SNE portrait of local fossil relics and structures[⋆]

Mario Ortigoza-Urdaneta[1], Katherine Vieira[1], José G. Fernández-Trincado[2], Anna B. A. Queiroz[3,4,5], Beatriz Barbuy[6], Timothy C. Beers[7], Cristina Chiappini[3,5], Friedrich Anders[8,9,10], Dante Minniti[11,12,13], and Baitian Tang[14]

[1] Instituto de Astronomía y Ciencias Planetarias, Universidad de Atacama, Copayapu 485, Copiapó, Chile
e-mail: `mario.ortigoza@postgrados.uda.cl`
[2] Instituto de Astronomía, Universidad Católica del Norte, Av. Angamos 0610, Antofagasta, Chile
e-mail: `jose.fernandez@ucn.cl`
[3] Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, 14482 Potsdam, Germany
[4] Institut für Physik und Astronomie, Universität Potsdam, Haus 28 Karl-Liebknecht-Str. 24/25, 14476 Golm, Germany
[5] Laboratório Interinstitucional de e-Astronomia - LIneA, 20921-400 Rio de Janeiro, Brazil
[6] Universidade de São Paulo, IAG, Rua do Matão 1226, Cidade Universitária, São Paulo 05508-900, Brazil
[7] Department of Physics and Astronomy and JINA Center for the Evolution of the Elements, University of Notre Dame, Notre Dame, IN 46556, USA
[8] Departament de Física Quántica i Astrofísica (FQA), Universitat de Barcelona (UB), c. Martí i Franquès, 1, 08028 Barcelona, Spain
[9] Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (UB), c. Martí i Franquès, 1, 08028 Barcelona, Spain
[10] Institut d'Estudis Espacials de Catalunya (IEEC), c. Gran Capità, 2-4, 08034 Barcelona, Spain
[11] Depto. de Cs. Físicas, Facultad de Ciencias Exactas, Universidad Andrés Bello, Av. Fernández Concha 700, Las Condes, Santiago, Chile
[12] Vatican Observatory, 00120 Vatican City State, Italy
[13] Departamento de Fisica, Universidade Federal de Santa Catarina, Trinidade 88040-900, Florianopolis, Brazil
[14] School of Physics and Astronomy, Sun Yat-sen University, Zhuhai 519082, PR China

Received 4 March 2023 / Accepted 8 June 2023

ABSTRACT

Based on high-quality Apache Point Observatory Galactic Evolution Experiment (APOGEE) DR17 and *Gaia* DR3 data for 1742 red giants stars within 5 kpc of the Sun and not rotating with the Galactic disk ($V_\phi < 100\,\mathrm{km\,s^{-1}}$), we used the nonlinear technique of unsupervised analysis t-Distributed Stochastic Neighbor Embedding (t-SNE) to detect coherent structures in the space of ten chemical-abundance ratios: [Fe/H], [O/Fe], [Mg/Fe], [Si/Fe], [Ca/Fe], [C/Fe], [N/Fe], [Al/Fe], [Mn/Fe], and [Ni/Fe]. Additionally, we obtained orbital parameters for each star using the nonaxisymmetric gravitational potential `GravPot16`. Seven structures are detected, including Splash, Gaia-Sausage-Enceladus (GSE), the high-α heated-disk population, N-C-O peculiar stars, and inner disk-like stars, plus two other groups that did not match anything previously reported in the literature, here named Galileo 5 and Galileo 6 (G5 and G6). These two groups overlap with Splash in [Fe/H], with G5 having a lower metallicity than G6, and they are both between GSE and Splash in the [Mg/Mn] versus [Al/Fe] plane, with G5 being in the α-rich in situ locus and G6 on the border of the α-poor in situ one. Nonetheless, their low [Ni/Fe] hints at a possible ex situ origin. Their orbital energy distributions are between Splash and GSE, with G5 being slightly more energetic than G6. We verified the robustness of all the obtained groups by exploring a large range of t-SNE parameters, applying it to various subsets of data, and also measuring the effect of abundance errors through Monte Carlo tests.

Key words. stars: abundances – stars: chemically peculiar – solar neighborhood – Galaxy: halo – techniques: spectroscopic – methods: statistical

1. Introduction

Numerous studies have revealed that the various structures that make up the Milky Way (MW) have been affected by both intrinsic (e.g., secular evolution, Vera et al. 2016; Combes et al. 2017) and extrinsic (e.g., merger events, Kruijssen et al. 2020) processes. As a consequence, our galaxy has mixed populations comprising both in situ and ex situ star formation. The identification of these structures is fundamental to understanding the formation and evolutionary history of our galaxy.

In some cases, these structures still retain traces of their original kinematics, and can be identified by their peculiar orbits as compared to canonical populations of the MW (Belokurov et al. 2018; Koppelman et al. 2019a); however, in many other cases, their original motion is erased by their prolonged interaction with the MW. On the other hand, most chemical abundances in stellar atmospheres are unaffected by the dynamical history of stars. Thus, a chemical signature becomes a valuable piece of information to determine the presence and origin of a given population. Taken together, kinematics, dynamics, and chemical abundances are extremely useful to decode the history of such population (e.g., Freeman & Bland-Hawthorn 2002; Hasselquist et al. 2021; Horta et al. 2023).

[⋆] Full Table C.1 is only available at the CDS via anonymous ftp to `cdsarc.u-strasbg.fr` (`130.79.128.5`) or via `https://cdsarc.cds.unistra.fr/viz-bin/cat/J/A+A/676/A140`

Thanks to the data provided by *Gaia* (Gaia Collaboration 2021) and spectroscopic surveys (e.g., APOGEE, Galactic Archaeology with HERMES, *Gaia* Radial Velocity Spectrometer and General Stellar Parametriser-spectroscopy, Majewski et al. 2017; Buder et al. 2022; Recio-Blanco et al. 2023), and their unprecedented precision, great advances have been made in the detection and understanding of structures such as Gaia-Sausage Enceladus (GSE; Belokurov et al. 2018; Haywood et al. 2018; Helmi et al. 2018; Koppelman et al. 2018; Myeong et al. 2018) or the Helmi stream(s) (Helmi et al. 1999; Koppelman et al. 2019b; Chiba & Beers 2000). In both cases, an ex situ origin is attributed, coming from past mergers with the MW. On the other hand, there is “Splash”, a population thought to have likely arisen from the heating of the primordial disk by an early merger (Bonaca et al. 2017, 2020; Haywood et al. 2018; Di Matteo et al. 2019; Belokurov et al. 2020). Splash has also been interpreted as an outcome of clumpy star formation of the early disk (Amarante 2020; Fiteni et al. 2021).

While in the past century astronomical data were often scarce, particularly regarding chemical abundances, parallaxes, and proper motions, nowadays the situation is the opposite. This large influx of data allows for a more detailed analysis of known structures and the discovery of new populations, although the increased density of data points hampers a clean separation of the populations, which usually overlap in regards to the many different measurements available. More data – traditionally analyzed – means it is more difficult to separate the (probable) members individually.

The significant influx of astronomical data has driven the application of statistical techniques that allow for automatic searches to identify potential associations of stars by specific properties. It is necessary to understand how these methods work and to what extent they can help us analyze the data, as there are differences in data quality due to both instrumental and natural factors. Among the widely used techniques is principal component analysis (PCA), which has been successfully applied in many kinds of data with excellent results (Rebonato & Jäckel 2000; Ting et al. 2012).

In this paper, we present the use of t-SNE (Hinton & Roweis 2003; van der Maaten & Hinton 2008) dimensionality reduction technique for a study of local low-velocity structures ($V_\phi < 100\,\mathrm{km\,s^{-1}}$, less than 5 kpc from the Sun), from input data provided by APOGEE+*Gaia*. t-SNE has been used in several areas of astronomy, such as premain sequence identification (e.g., Rim et al. 2022), spectral classification (e.g., Verma et al. 2021), and abundance-space dissection (e.g., Anders et al. 2018), where we note the powerful applicability for automated searches of populations. We provide an analysis based on t-SNE detections obtained from ten chemical abundances of our sample and the dynamics provided by GravPot16[1].

In the last few years, several sub-populations have been identified in the MW halo, for example, GSE, Splash, Sequoia (Barbá et al. 2019), Kraken (Kruijssen et al. 2020), among others, due to their peculiar kinematics, dynamics, and/or chemical abundances. It has been proposed by Naidu et al. (2020) and other authors that the MW halo is built entirely from accreted dwarfs and heating of the disk. Those substructures have been generally identified at rather large vertical distances from the MW plane, but in this study, we look for evidence of their presence closer to the Sun where they would be colocated with the disk system.

This paper is organized as follows: data sources and selection are described in Sects. 2–4, the MW dynamical model in Sect. 5, the t-SNE method in Sects. 6 and 7, and our results in Sect. 8. Section 9 discusses the accretion origin of some of the structures detected, Sect. 10 is dedicated to a quick analysis of the effects of the bar pattern speed in the obtained structures, and our conclusions are in Sect. 11. Additional figures and tables are located in the appendix.

2. APOGEE-2 DR17

The data employed in this study were obtained as part of the second phase of the Apache Point Observatory Galactic Evolution Experiment (APOGEE-2; Majewski et al. 2017), which was one of the four Sloan Digital Sky Survey IV surveys (SDSS-IV; Blanton et al. 2017). APOGEE-2 was a high-resolution ($R \sim 22500$) near-infrared (NIR) spectroscopic survey containing observations of 657135 stars, whose spectra were obtained using the cryogenic, multi-fiber (300 fibers) APOGEE spectrograph (Wilson et al. 2019) mounted on the 2.5 m SDSS telescope (Gunn et al. 2006) at Apache Point Observatory to observe the Northern Hemisphere (APOGEE-2N), and expanded to include a second APOGEE spectrograph on the 2.5 m Irénée du Pont telescope (Bowen & Vaughan 1973) at Las Campanas Observatory to observe the Southern Hemisphere (APOGEE-2S). Each instrument records most of the H band (1.51 µm–1.70 µm) on three detectors, with coverage gaps between ~1.58–1.59 µm and ~1.64–1.65 µm, and with each fiber subtending a ~2″ diameter on-sky field of view in the northern instrument and 1.3″ in the southern one. We refer the interested reader to Zasowski et al. (2013, 2017), Beaton et al. (2021), and Santana et al. (2021) for further details regarding the targeting strategy and design of the APOGEE-2 survey.

The final version of the APOGEE-2 catalog was published in December 2021 as part of the 17th data release (DR17) of SDSS (Abdurro’uf 2022) and is available publicly online through the SDSS Science Archive Server and Catalog Archive Server[2]. The APOGEE-2 data reduction pipeline is described in Nidever et al. (2015), while stellar parameters and chemical abundances in APOGEE-2 have been obtained within the APOGEE Stellar Parameters and Chemical Abundances Pipeline (`ASPCAP`; García Pérez et al. 2016). `ASPCAP` derives stellar atmospheric parameters, radial velocities, and as many as 26 individual elemental abundances for each APOGEE-2 spectrum by comparing each to a multidimensional grid of theoretical `MARCS` model atmosphere grid (Zamora et al. 2015), employing a χ^2 minimization routine with the code `FERRE` (Allende Prieto et al. 2006) to derive the best-fit parameters for each spectrum. We used the abundances computed from the synspec_fix spectral synthesis code.

The accuracy and precision of the atmospheric parameters and chemical abundances are extensively analyzed in Holtzman et al. (2018), and Jönsson et al. (2018, 2020). Details regarding the customized H band line list are fully described in Shetrone et al. (2015), Hasselquist et al. (2016), Cunha et al. (2017), and Smith et al. (2021).

---

[1] `https://gravpot.utinam.cnrs.fr`

[2] SDSS DR17 data: `https://www.sdss4.org/dr17/irspec/spectro_data/`

3. Input parameters

3.1. Elemental abundances and radial velocities

We made use of high-quality elemental abundances for ten chemical species, including the light- (C, N), α- (O, Mg, Si, Ca), odd-Z (Al), and iron-peak (Mn, Fe, Ni) elements, as well as precise ($<1\,\mathrm{km\,s^{-1}}$) APOGEE-2 spectroscopic radial velocity measurements when calculating kinematics and orbits for giant stars in the APOGEE DR17 catalog (Abdurro'uf 2022). The other 16 abundances available from ASPCAP were not used because they are affected by telluric bands (e.g., Na), are blended with other atomic and/or molecular lines, or have a lower signal-to-noise ratio (S/N). Also, for some elements a reduced number of stars have measurements.

3.2. `StarHorse` distances

We also made use of the precise spectro-photo-astrometric distances ($d_{\texttt{StarHorse}}$) estimated with the Bayesian `StarHorse` code (Santiago et al. 2016; Queiroz et al. 2018, 2020, 2023; Anders et al. 2019), which have been published in the form of a value-added catalog (VAC)[3]. `StarHorse` combines high-resolution spectroscopic data from APOGEE-2 DR17 with broadband photometric data from several sources (`Pan-STARSS1`, `2MASS`, and `AllWISE`), as well as parallaxes from *Gaia* DR3 when available, along with their associated uncertainties, in order to derive distances, extinctions, and astrophysical parameters for APOGEE-2 stars through a Bayesian isochrone-fitting procedure. These parameters exhibit resilience against variations in the assumed Galactic priors and the *Gaia* parallax zero-point offset.

3.3. Proper motions

Astrometric data that are taken from *Gaia* DR3 (Gaia Collaboration 2021) are also available for the APOGEE-2 sample. For this study, when calculating orbital parameters, the renormalized unit weight error (`RUWE`) astrometric quality indicator was imposed to be `RUWE` < 1.40, in order to have astrometrically well-behaved sources (see, e.g., Lindegren et al. 2018).

4. Sample selection

The APOGEE-2 DR17 catalog contains more than six hundred thousand entries. Several cuts were applied to refine our sample. For quality control of the stellar parameters and elemental abundances, we first cleaned the sample for sources with unreliable parameters by keeping only sources with `STARFLAG` and `ASPCAPFLAG` equal to zero (see e.g., Holtzman et al. 2015, 2018), and stars with good spectra ($S/N > 70\,\mathrm{pixel^{-1}}$) and good stellar abundances, which are flagged as `X_FE_FLAG` = 0 (`X` = C, N, O, Mg, Al, Si, Ca, Mn, Fe, and Ni). These cuts ensure that there are not major flagged issues, such as a low S/N, poor synthetic spectral fit, stellar parameters near grid boundaries, and potential problematic object spectra.

We also cut in distance precision, by selecting only stars with less than 20% error distances from $d_{\texttt{StarHorse}}$. The application of these cuts yielded an initial sample of 144 476 stars. Then, we made a cut to select red giants by selecting, from the initial sample only those stars with surface gravity values $\log g < 3.6$, and stellar effective temperatures in the range $T_{\mathrm{eff}} < 5500\,\mathrm{K}$, yielding a sample of 142 305 stars. The upper effective temperature limit was imposed to remove stars at the hot edge of the main stellar-atmospheric model grids used in the APOGEE-2 analysis. The following cut was in kinematics, by choosing only stars with Galactocentric azimuthal velocity $V_{\phi} < 100\,\mathrm{km\,s^{-1}}$, which minimized the presence of the dominant Galactic disk population, yielding a sample of 9544 stars.

We also checked that this sample did not have repeated *Gaia* DR3 source IDs, because the main APOGEE catalog does have, in a few cases, pairs of different APOGEE IDs associated with the same *Gaia* source ID. This is relevant in the next cut for `RUWE` < 1.4, for which a true *Gaia* match is obviously needed, since the match is done via the *Gaia* DR3 source ID listed in the APOGEE catalog. The cut in `RUWE` yielded a sample of 9107 stars. Subsequently, we applied a selection criterion to limit the sample to stars confined within a 5 kpc radius from the Sun ($d_{\texttt{StarHorse}} < 5\,\mathrm{kpc}$). This filtering process resulted in a dataset comprising 1905 stars, which adequately samples a significant volume enabling exploration of the inner-halo region while effectively avoiding the influence of the Bulge region. In this paper, we refer to the inner halo as the part of the MW halo that is dominated by debris from past events of accretion, involving massive objects such as dwarf galaxies or globular clusters (GCs; e.g., Helmi et al. 2018).

Finally, we discarded known GC member stars as listed by Mészáros et al. (2020), Baumgardt & Vasiliev (2021), and Vasiliev & Baumgardt (2021). Our final sample for the analysis below contains 1742 red giant stars. It is also important to notice that we made no correction for the selection biases within the APOGEE-2 DR17 survey. Thus, stars close to the solar neighborhood are over-represented in our sample.

5. Dynamical model

In order to construct a comprehensive orbital study of the local structures across of the Solar vicinity, we used a state-of-the-art orbital integration model in a nonaxisymmetric gravitational potential that fits the structural and dynamical parameters based on recent knowledge of our Galaxy. For the computations in this work, we have employed the rotating "boxy/peanut" bar of the Galactic potential model called `GravPot16`, along with other composite stellar components. The considered structural parameters of our bar model, for example, mass, present-day orientation, and pattern speeds, are within observational estimations that lie in the range of $1.1 \times 10^{10}\,M_{\odot}$ and 20°, which is in line with that from Tang et al. (2018) and Fernández-Trincado et al. (2020d), and $31–51\,\mathrm{km\,s^{-1}\,kpc^{-1}}$ (see e.g., Bovy et al. 2019; Sanders et al. 2019) in increments of $10\,\mathrm{km\,s^{-1}\,kpc^{-1}}$, respectively. The bar scale lengths are $x_0 = 1.46\,\mathrm{kpc}$, $y_0 = 0.49\,\mathrm{kpc}$, $z_0 = 0.39\,\mathrm{kpc}$, and the middle region ends at the effective major semiaxis of the bar $\mathrm{Rc} = 3.28\,\mathrm{kpc}$ (Robin et al. 2012). The density profile of the adopted boxy/peanut bar is the same as in Robin et al. (2012).

`GravPot16` considers, on a global scale, a 3D steady-state gravitational potential for the MW, modeled as the superposition of axisymmetric and nonaxisymmetric components. The axisymmetric potential comprises the superposition of many composite stellar populations belonging to seven thin disks. For each i^{th} component of the thin disk, we implemented an Einasto density-profile law (Einasto et al. 1979), as described in Robin et al. (2003), superposed with two thick-disk components, with each one following a simple hyperbolic secant squared decreasing vertically from the Galactic plane plus an exponential

---

[3] https://data.sdss.org/datamodel/files/APOGEE_STARHORSE/

profile decreasing with Galactocentric radius, as described in Robin et al. (2014). We also implemented the density profile of the interstellar matter (ISM) component with a density mass as presented in Robin et al. (2003).

The model also correctly accounts for the underlying stellar halo, modeled by a Hernquist profile as already described in Robin et al. (2014), and surrounded by a single spherical dark matter halo component (Robin et al. 2003); no time dependence of the density profiles is assumed. The most important limitations of our model are as follows: (i) We ignored secular changes in the MW potential over time, which are expected although our Galaxy has had a quiet recent accretion history; (ii) We did not consider perturbations due to spiral arms, as an in-depth analysis is beyond the scope of this paper.

For reference, the Galactic convention adopted for this work is as follows: the X-axis is oriented toward $l = 0°$ and $b = 0°$, the Y-axis is oriented toward $l = 90°$ and $b = 0°$, and the disk rotates toward $l = 90°$; the velocities are also oriented in these directions. In this convention, the Sun's orbital velocity vector is $[U_{\odot}, V_{\odot}, W_{\odot}] = [11.1, 248.5, 7.25]\,\mathrm{km\,s^{-1}}$, which is in line with that from Brunthaler et al. (2011) and Reid & Brunthaler (2020). The model has been rescaled to the Sun's Galactocentric distance, $R_{\odot} = 8.178\,\mathrm{kpc}$ (GRAVITY Collaboration 2019) and $Z_{\odot} = 25\,\mathrm{pc}$ (Jurić et al. 2008).

The orbital trajectories for our entire sample were integrated by adopting a simple Monte Carlo scheme and the Runge-Kutta algorithm of seventh-eight order elaborated by Fehlberg (1968), in order to construct the initial conditions for each star, taking into account the uncertainties in the radial velocities provided by the APOGEE-2 survey, the absolute proper motions provided by the *Gaia* DR3 catalog (Gaia Collaboration 2023), and the spectro-photo-astrometric distances provided by `StarHorse` (Queiroz et al. 2023). The uncertainties in the input data (e.g., α, δ, distance, proper motions, and line-of-sight velocity errors) were randomly propagated as 1σ variations in a Gaussian Monte Carlo resampling. For each star, we computed one thousand orbits, computed backward in time over 3 Gyr. The average value of the orbital elements was found for these thousand realizations, with uncertainty ranges given by the 16th and 84th percentile values.

For this study, we used the average value of the amplitude of the vertical oscillation $|Z_{\mathrm{max}}|$, the perigalactic distance R_{peri}, apogalactic distance R_{apo}, the eccentricity e, the orbital Jacobi constant E_j computed in the reference frame of the bar and the "characteristic" orbital energy E_{char} as envisioned by Moreno et al. (2015). The E_{char} versus E_J plane was used to broadly discriminate orbit populations, the lower the E_{char} the more radially confined to the Galaxy, the lower the E_J the more vertically confined it is.

We also considered the minimum and maximum of the z component of the angular momentum in the inertial frame, which defines their orbital behavior. Under a nonaxisymmetric potential, the angular momentum is not conserved, and thus during the 3 Gyr integration time, a star can be always prograde (P), always retrograde (R), or change from P to R or vice versa (P-R). That is, a P orbit has $L_{\mathrm{min}} < 0$ and $L_{\mathrm{max}} < 0$, an R orbit has $L_{\mathrm{min}} > 0$ and $L_{\mathrm{max}} > 0$, and a P-R orbit has L_{min} and L_{max} with opposite signs. In general, P-R orbits are extremely eccentric, indicating that their transition from a P to R orbit, or vice versa, is achieved by following a very radial orbit, with their periastron rather close to the Galactic center. Their dynamics can be subject to chaotic effects with very sensitive orbits, and/or could be trapped in a resonance. Also, when orbits are isotropically heated, not just vertically, we suspect that the mechanism keeping them that way, internal or external, is different from the one that formed the MW thick disk.

6. t-SNE

The t-SNE algorithm is a nonlinear technique of unsupervised analysis for the study of high-dimensional ensembles through a dimension reduction that allows its visualization (2D or 3D), as presented by van der Maaten & Hinton (2008). Many areas of study make use of this technique. due to its utility in finding possibly related structures automatically and transferring the information to spaces that allow us to visualize them. In other words, t-SNE is able to detect similarity among data points in high-dimensional space, by transforming them onto 2D values (the t-SNE plane), so that nearby points are considered as likely members of the same population, or at least somehow related.

Astronomy presents an ideal scenario for the use of this technique since we now have large catalogs with a variety of information for each star. In the context of our work, t-SNE has proven useful, such as in Anders et al. (2018), where the interested reader can find a summarized yet comprehensive description of the operation of t-SNE. In the following paragraphs, we provide an overall description of how the method works.

The t-SNE algorithm comprises two main stages. First, it constructs a probability distribution over pairs of high-dimensional objects so that the closer and farther objects are the higher and lower probability assigned for being related, respectively. Second, starting from a Gaussian low-dimensional distribution, t-SNE iteratively constructs a probability distribution of points by minimizing the Kullback-Leibler divergence between the high- and low-dimensional distributions using gradient descent. As a result, the final low-dimensional distribution resembles the probability of the high-dimension one, so that nearby objects in one are also close in the other.

The method has one main parameter, called the `perplexity` p, which can be thought of as a guess about the number of close neighbors each point has; the ideal value for p depends on the sample size. Since a change in perplexity has –in many cases– a complex effect on the resulting map, it is recommended to try different values for p (Wattenberg et al. 2016). According to Linderman & Steinerberger (2017), two other hyper-parameters of t-SNE can be optimally chosen: the learning rate (~1) and the early exaggeration parameter (10% of the sample size). We used these recommended values as a starting point, but eventually other values were set, guided by the best separation of the detected groups in the t-SNE plane.

We used the Python implementation of t-SNE included in the `scikit-learn` package (Pedregosa et al. 2011). We would like to provide a word of caution: with the same data and parameter inputs, we recovered the exact same t-SNE planes up to version 1.1.3 of `scikit-learn`, but when using newer updates, that is version 1.2.0 onward, the look of the t-SNE plane changes. We checked that our detected groups remained virtually the same, and therefore our conclusions are not altered.

The distribution of points in the t-SNE plane not only depends on the multidimensional data input, but also on the order each of these vectors is entered. This means that the same data input sorted in different ways yields different-looking t-SNE planes, but in all cases the same points are associated quite similarly. In our case, the stellar data were inserted in order of increasing [Fe/H], as an additional check on the obtained results.

The t-SNE algorithm does have two major weaknesses: 1) no missing individual data are allowed, meaning each star must have measured values in all the dimensions considered; and 2) it does not consider the effect of individual uncertainties. We dealt with these issues in the following ways: 1) by using only stars with all ten chosen abundances measured and 2) by performing a Monte Carlo experiment to show that our results are resilient to fluctuations in abundance.

7. Strategy adopted with t-SNE

The input entries fed into t-SNE are the abundance ratios [Fe/H], [O/Fe], [Mg/Fe], [Si/Fe], [Ca/Fe], [C/Fe], [N/Fe], [Al/Fe], [Mn/Fe], and [Ni/Fe]. We also included [Mg/Mn], which has proven to be useful in separating accreted from in situ structures (Das et al. 2020; Naidu et al. 2020), considering also that from previous work we know of structures that overlap in [Mg/Mn] but are separate in [Mg/Fe] and [Mn/Fe], or vice versa. Our final results were obtained under `perplexity = 25`, `early_exageration = 130`, `init = pca`, `learning_rate = auto`, and `n_iter = 7000`, in which the detected associations of stars were more visibly separated (see Fig. 1, central panel). A total of seven structures were identified, and are presented for separate analysis in the coming sections. To choose this result as the final one, we thoroughly tested the effect of the input parameters and the abundance errors in the obtained results, as described below. Also, we checked their dynamical parameters to help confirm if the detected populations also exhibited additional differences in their orbits.

Stars in all the structures identified were persistently associated despite changing the t-SNE parameters over a range of values: 20 < `perplexity` < 40 and 80 < `early_exageration` < 130, although a few of the smaller groups were more unstable in their location in the t-SNE plane, moving from one location to another and getting closer or farther from the other groups. The diagrams generated by t-SNE can strongly vary the shape of the distribution in the plane depending on the initial parameters imposed (see Fig. A.1), but indeed some of the identified structures predominate and simply relocate, always staying as a single structure (Groups G1, G2, G5, and G7 mainly). Therefore, the final selection shown took into consideration how their location in the t-SNE plane changed with respect to the other detected associations. Another important test was to verify the separation of adjacent or nearby structures by running t-SNE on smaller input subsets. Moreover, we used the obtained t-SNE values as additional inputs for a second t-SNE run, to check that the selected groups were "squeezed" further into the new t-SNE plane. These tests helped us to trace the boundaries between nearby groups. The results of this test are shown in Appendix A and Fig. A.2. We also found that the most extended structures (G1 and G2) kept each one cohesively in these runs, without further separation.

To understand the effect of abundance errors, we generated random samples based on the uncertainty of each abundance and then fed those as input to t-SNE. For each star, 100 random measurements were drawn from a Gaussian centered on the measured abundance and dispersion corresponding to its uncertainty. Samples were then created by randomly taking one out of 100 simultaneously in all input abundances and applying the t-SNE to each generated sample with the same parameters as the one chosen for our final results. We found that, despite a significant variation in the distribution shape of the points in the t-SNE plane (see Fig. A.3), the same structures stayed together consistently and were detected in the different t-SNE realizations. Therefore, we conclude that abundance uncertainties do not lead to significant changes in the group identifications. Moreover, this reinforces the fact that the structures detected do not correspond to random associations.

8. Structures detected in the t-SNE projection

The large panel shown in Fig. 1 is our reference t-SNE projection, on which we have identified and named seven substructures that clearly emerge from it. The figure also shows the distribution of these substructures in several elemental-abundance diagrams (small panels), which have been color-coded according to the substructures' names. This t-SNE plane is also shown in Fig. 2 and color-coded by the indicated abundances and dynamical parameters e, $|Z_{\rm max}|$ in kiloparsecs, and $R_{\rm apo}$ in kiloparsecs.

From inspection of the t-SNE map in Fig. 1, we can immediately appreciate two dominant and well-populated substructures, which fit very well with the descriptions of the Splash-like population (G1, blue points, Fig. 1) and a significant past merger event (G2, beige points, Fig. 1) established in the literature. Five other smaller substructures were identified as well. Each substructure is described individually below, considering not only their chemical abundances but also their dynamical properties, which were not fed into the t-SNE but exhibit distinctive ranges. In all corresponding figures, V_ϕ is in km s^{-1}, distances are given in kiloparsecs, and energies are in 10^5 km^2 s^{-2}. Eccentricity e and t-SNE X and t-SNE Y values are dimensionless. Figure 3 shows the distribution of the chemical abundances for some of the detected groups, to prove they were all indeed different from each other in at least one of the elements. Table B.1 contains the median and median absolute deviations for all the chemical abundances and dynamical parameters of each detected group.

8.1. Group 1 (G1): Splash

This is the largest and most consistent structure detected by the t-SNE method, showing higher α-element abundances $\gtrsim$+0.20 and slightly enriched levels of Al ([Al/Fe] > +0.1) at metallicities [Fe/H] > −1 (Fig. 1, blue points) to solar metallicity, with chemistry consistent with the thick disk. Dynamically (see Fig. 4), it exhibits a wide range of $|Z_{\rm max}|$, spanning from in-plane to 14 kpc from the Galactic plane, clearly overlapping with the inner-halo region of the Galaxy.

We note that $R_{\rm apo}$ spans mostly from 3 to 12 kpc, but a few stars reach up to 25 kpc, while $R_{\rm peri}$ is mostly below 2.5 kpc. As expected from such values, these stars follow eccentric orbits, typically with $e \gtrsim 0.6$ (see also the lower panels of Fig. 2, color-coded by e, $|Z_{\rm max}|$, and $R_{\rm apo}$). The orbital energies of these stars span a large range, from low to mid values. Much of this group fits very well with the descriptions established for the so-called Splash by different authors (Bonaca et al. 2017, 2020; Haywood et al. 2018; Di Matteo et al. 2019; Belokurov et al. 2020).

Group 1 is mostly (76%) comprised of P orbits stars (see Table B.2), which supports the idea of its origin as heated-disk stars. Nonetheless, a nonnegligible portion of G1 stars (see Table B.2) are P-R and have R orbits, which are not only vertically but also radially heated. Therefore, these stars have the chemical signature of the MW thick disk, but are dynamically hotter, pointing at some additional mechanism (internal or external) giving these stars a higher "kick".

8.2. Group 2 (G2): Gaia-Enceladus-Sausage-like

This structure matches the group dubbed GSE very well (Helmi et al. 2018; Belokurov et al. 2018; Haywood et al. 2018; Mackereth et al. 2019), showing its characteristic low concentrations of aluminum and iron: [Al/Fe] $\lesssim$ 0 and [Fe/H] < −0.6, in its component stars (see Fig. 3). Our Group 2 (G2) is deficient

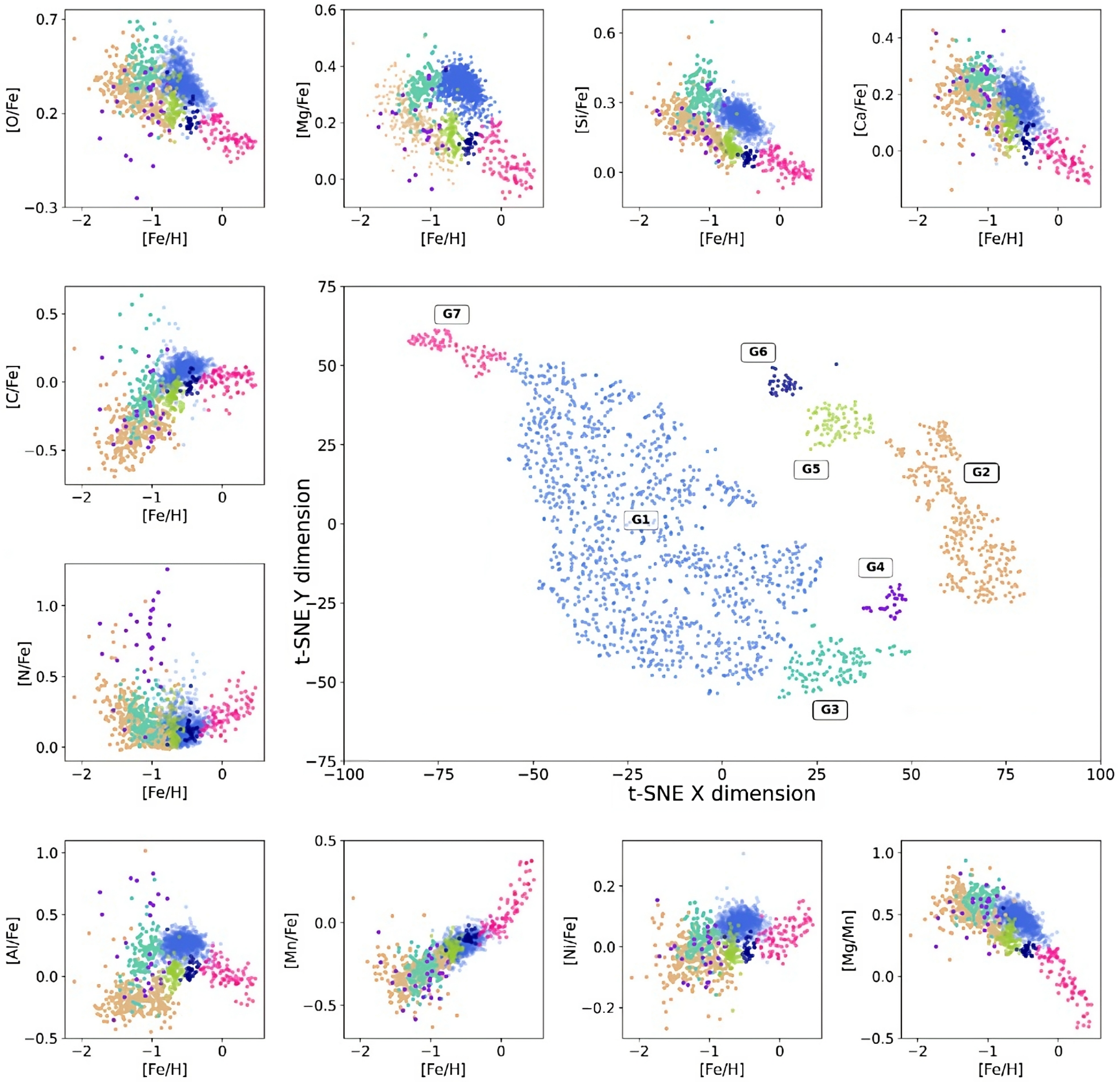

Fig. 1. t-SNE plane, shown in the large main panel, with the selected groups identified by colors and labels. The surrounding panels correspond to the ten APOGEE abundances used as input for t-SNE.

in carbon, with [C/Fe] $\lesssim 0$, although some outliers are seen with higher values. For the ratios [Si/Fe], [Mg/Fe], and [Ca/Fe], G2 follows the known sequence vs [Fe/H] for GSE, including the so-called *knee* at [Fe/H] ~ -1. The elements [N/Fe], [O/Fe], and [Ni/Fe] are more spread out.

The G2 stars (see Fig. 5) exhibit rather eccentric orbits, spanning mainly between $0.6 \leq e \leq 1$ with a visible concentration at $e > 0.9$, and heights exceeding 2 kpc, reaching 20 kpc. The energy plane E_{char} versus E_J exhibits high values (mainly $E_J > -2$ and $E_{\mathrm{char}} > -2$), spread in a way consistent with the dynamics of the so-called inner-halo component, first identified by Carollo et al. (2007, 2010) and Beers et al. (2012). Most of the stars in G2 are similarly distributed between P and R orbits with a smaller fraction in the P-R classification (see Table B.2), which means G2 stars are evenly distributed in L_Z, not showing a preferential rotation direction. We noticed that G2 stars with the lowest V_ϕ values – potentially members of Sequoia – are all also R, but they spread all over the G2 locus, meaning t-SNE did not detect anything closely in common among them. Therefore, we kept those as G2 members. The periastron distances for many G2 stars are within 1 kpc of the Galactic center, with apoastron distances between 8 and 25 kpc. The highest vertical distance from the Galactic plane for these stars spans between 2 and 25 kpc. One star (APOGEE ID: 2M13393889+1836032) reaches $R_{\mathrm{apo}} \sim 80$ kpc and $|Z_{\mathrm{max}}| \sim 80$ kpc as well. All these features are in agreement with previous results (Mackereth & Bovy 2020; Naidu et al. 2020; Feuillet et al. 2021; Buder et al. 2022).

8.3. Group 3 (G3): High-α heated-disk population

From the t-SNE plane in Fig. 1, Group 3 (G3) is next and seemingly connected to G1, yet we diagnosed G3 as a separate group because in many tests (see Figs. A.1 and A.3) it always appeared at one extreme of G1, looking similar to an appendage connected by only a few points, and being always more metal-poor than G1 (see Fig. 2, row 2, Col. 5). We also noticed that, in the t-SNE double iteration (see Fig. A.2), G3 is detached from G1 while G1 itself, despite exhibiting some internal clustering, remains

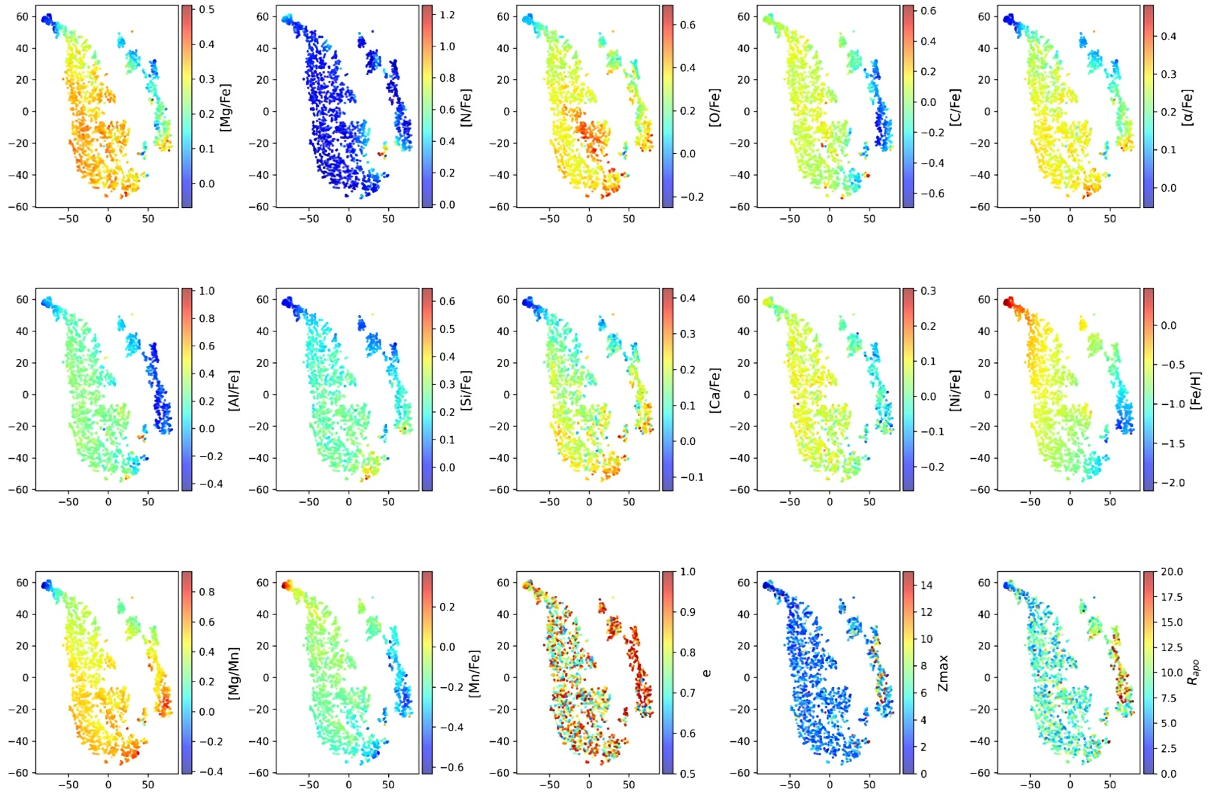

Fig. 2. t-SNE plane color-coded by the indicated abundances and dynamical parameters e, $|Z_{\max}|$ in kiloparsecs, and R_{apo} in kiloparsecs.

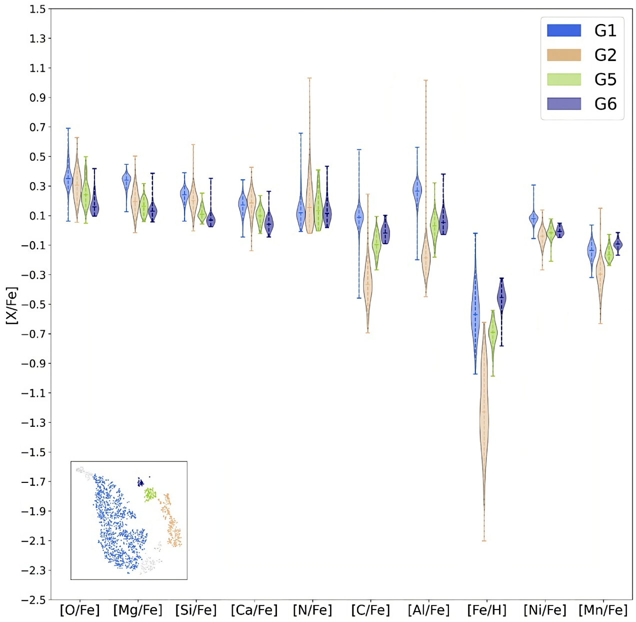

Fig. 3. Comparison of Groups G1 (Splash), G2 (GSE), G5 (Galileo 5), and G6 (Galileo 6), in each of the chemical abundances used in this work, displayed in the form of a violin plot. A common trend in the compared abundances for the α elements can be seen, as well as in the iron-peak elements. The observed distributions demonstrate that each group is different from the others in at least one chemical abundance.

mostly together. Additionally, when applying t-SNE only to the stars we identified in G1 and G3 alone, under various conditions, G3 is often separated from G1, while the latter keeps its integrity.

Although our Group 3 (G3) stellar abundances overlap with either G1 and/or G2 in various elements, it is clear from Fig. 1 that this structure sits between them as a separate population in the [Si/Fe] versus [Fe/H] and [Al/Fe] versus [Fe/H] planes, with it being the structure that is the richest in silicon out of our whole dataset. As Hawkins et al. (2015) describe in their Fig. 1, the abundance locus of canonical halo MW stars in the [α/Fe] versus [Fe/H] plane is also fulfilled well by G3 stars. Dynamically though, G3 stars' orbital parameters are spread over ranges similar to those of G1 (Splash; see Fig. 6), although they are more dispersed. A population of such features has been found by Nissen & Schuster (2010), for their high-α heated population, and also by Hayes et al. (2018) for their high-Mg thick disk-like population.

Nissen & Schuster (2010) suggested that these stars may be ancient disk or bulge stars "heated" to halo kinematics by merging satellite galaxies, or they could have formed as the first stars during the collapse of a proto-Galactic gas cloud, while Hayes et al. (2018) claim that the HMg population is likely associated with in situ formation. In Fig. 7, G3 sits on the high-α in situ locus. In the chemical-evolution model for a MW galaxy proposed by Horta et al. (2021, their Fig. 2), this location roughly corresponds to an in situ early disk population. We conclude that G3 corresponds to this structure, which we tag as the high-α population.

8.4. Group 4 (G4): N-C-O peculiar stars

This is a small substructure in t-SNE space, whose abundances relative to Fe are shown in small panels in Fig. 1. A large fraction of the stars in this substructure are typified by near-solar and anomalously high levels of [N/Fe]

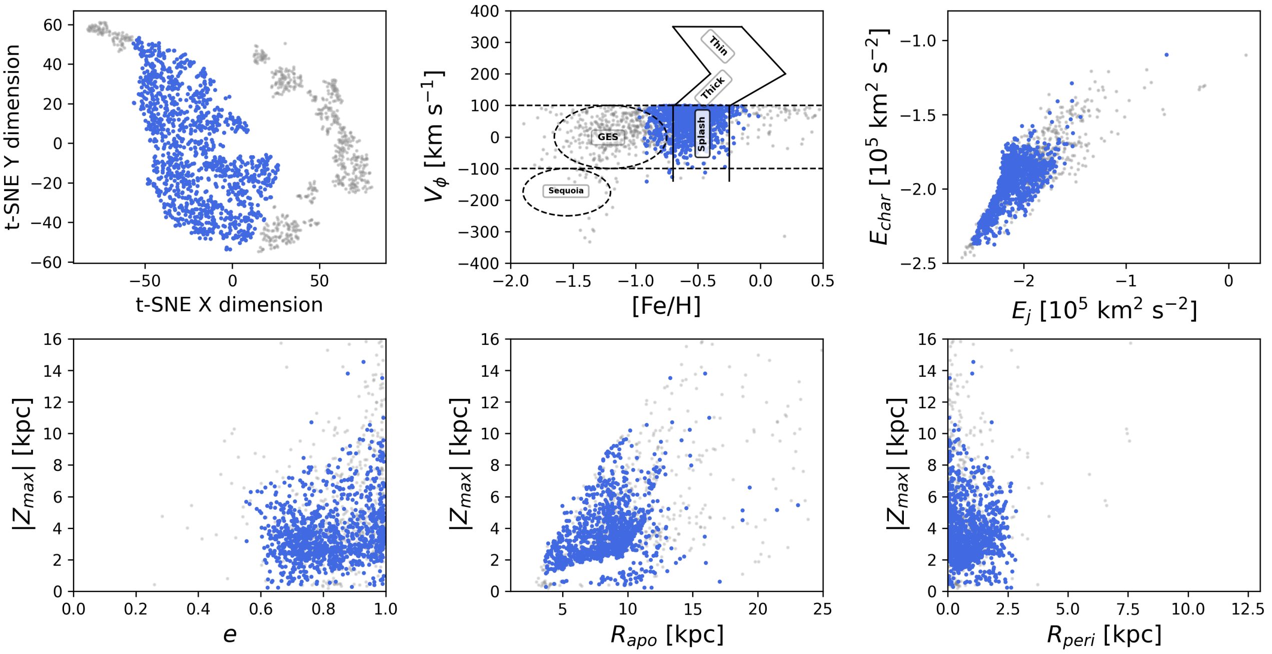

Fig. 4. G1 identified as Splash. Energies units are 10^5 km^2 s^{-2}. Details are explained in Sect. 8.1.

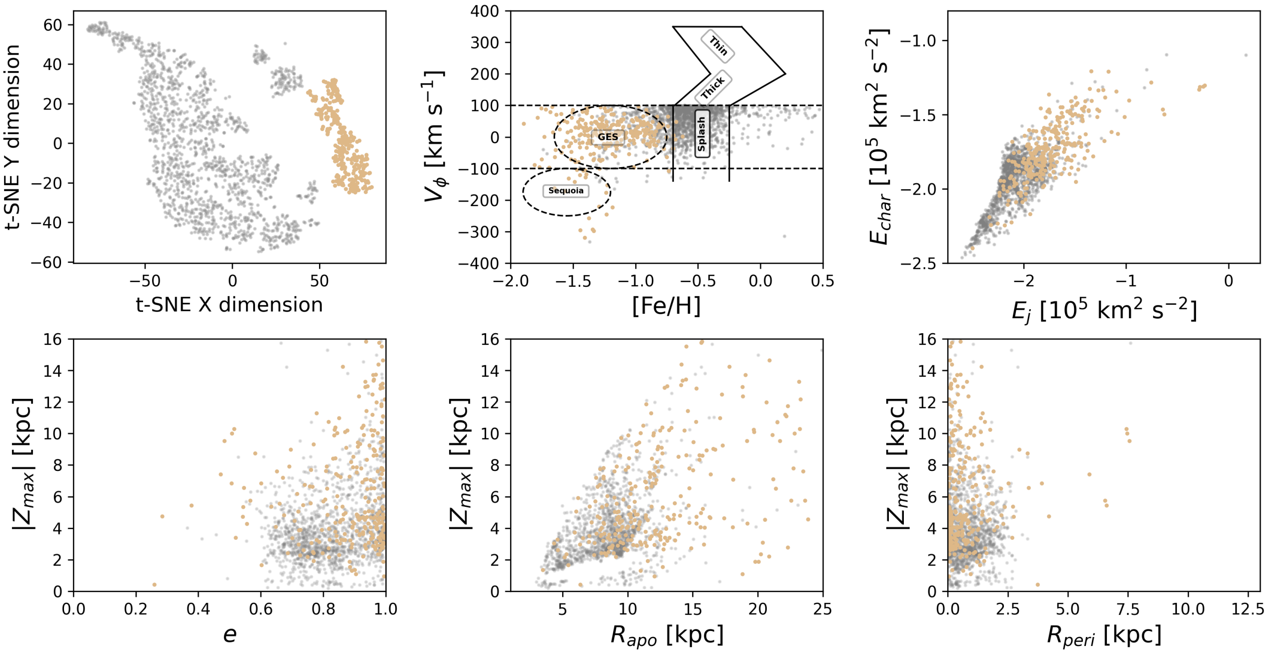

Fig. 5. G2 identified as GSE. Details are explained in Sect. 8.2. Panels and units are the same a those in Fig. 4.

that are well above ($\gtrsim$+0.5) typical Galactic levels over a range of metallicities, accompanied by decreased abundances of [C/Fe] < +0.15, as seen in Figs. 1 and 2. A portion of its stars also exhibit high [Al/Fe]. This substructure also matches very well with the atypical nitrogen-enriched population, as envisioned by a series of works (see, e.g., Schiavon et al. 2017; Fernández-Trincado et al. 2016, 2017, 2019a,b, 2020a–c, 2021a–c, 2022) that have attributed these stars to exhibit GC second-generation-like chemical patterns, and are likely GC debris (Fernández-Trincado et al. 2021a,b) that have now become part of the Galactic field population. The dynamics of these stars have inner-halo-like orbits, with values spreading with no particular concentration in any parameter (see Fig. 8).

A closer examination of their abundances reveals that G4 contains three subgroups of stars:

– N-rich and C-poor stars: [N/Fe] > +0.5 and [C/Fe] < +0.15, amounting to 20 stars falling into this category, including 17 already reported by Fernández-Trincado et al. (2022), plus three others (APOGEE IDs: 2M06572697+5543115, 2M21372238+1244305 and 2M16254515-2620462), located within the abundance limits defined by Fernández-Trincado et al. (2022) for these kinds of stars, and proposed by these authors to GC second-generation or chemically anomalous debris. These three stars are genuine newly identified N-rich stars, with the last one (2M16254515-2620462) being a potential extra-tidal star of M 4.

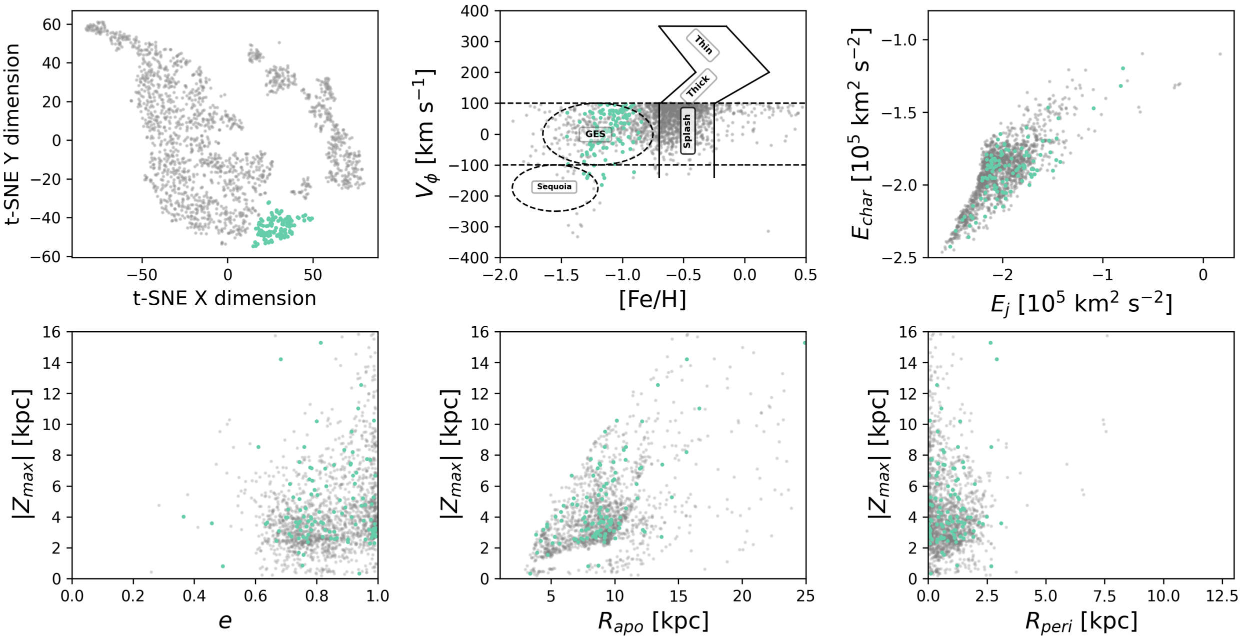

Fig. 6. G3 identified as the high-α heated-disk population. Details are explained in Sect. 8.3. Panels and units are the same as those in Fig. 4.

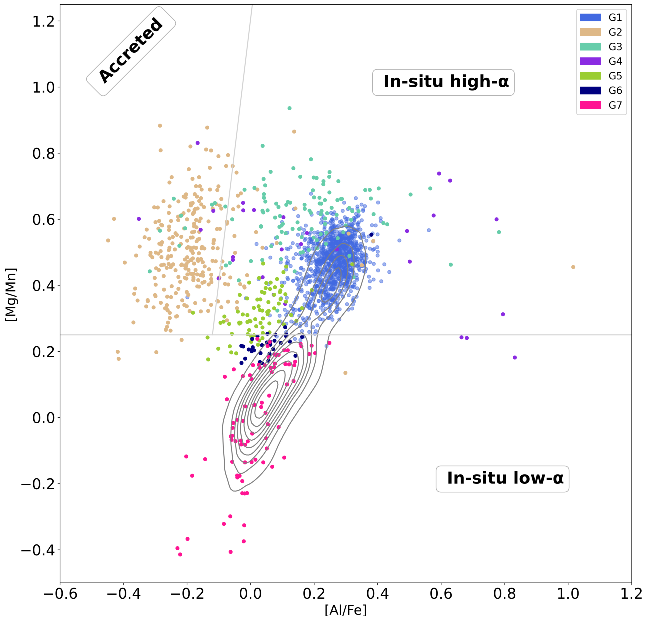

Fig. 7. [Mg/Mn] versus [Al/Fe] plot, color-coded for each detected group. A kernel density estimation (KDE) plot of MW disk stars is shown with gray curves. These were selected from APOGEE DR17 following the same quality criteria as our halo stars, and $V_\phi \geq 100\,\mathrm{km\,s^{-1}}$. Straight lines mark the boundaries of the accretion and in situ high- and low-α loci used in Das et al. (2020). The G2 (GSE) locus is mostly limited in this plot by $[\mathrm{Al/Fe}] < 0$ and $[\mathrm{Mg/Mn}] > +0.2$.

– N-poor and C-poor stars: $[\mathrm{N/Fe}] < +0.55$ and $[\mathrm{C/Fe}] < -0.1$, which amount to five stars that also exhibit subsolar [Al/Fe] abundance ratios. When examined in the [Mg/Mn]–[Al/Fe] plane these stars clearly fall in the accreted halo region (see Fig. 7). We found that three of the five stars (APOGEE IDs 2M12211605-2310262, 2M23364836-1135404 and 2M21420647-3019385) are within GSE's t-SNE footprint in all the tests shown in Fig. A.3. We postulate that they are part of merger remnants (and/or likely associated with GSE, or to a number of unknown events contributing to this sub-population) present in the MW.

– N-rich, C-rich, and O-rich stars: $[\mathrm{N/Fe}] > +0.4$, $[\mathrm{C/Fe}] > +0.15$ and $[\mathrm{O/Fe}] > +0.2$, which amount to three stars. These stars also have super-solar values in Mg and Si. Stars with such enhancements have been reported by Beers & Christlieb (2005), for example. It is likely that the atmosphere of these mildly carbon-enhanced stars has been contaminated either by an intrinsic process, such as self-enrichment (likely a thermally pulsing asymptotic giant branch), or by a past extrinsic event, that is, the mass-transfer hypothesis (binary mass-transfer systems). However, with the available radial velocity scatter (`VSCATTER` $< 1\,\mathrm{km\,s^{-1}}$) from three APOGEE-2 visits, it is not possible to support or reject either possibility.
In our entire sample of 1742 stars, we have 24 matches with the 412 N-rich stars listed by Fernández-Trincado et al. (2022), of which the seven that are not in G4 were assigned by t-SNE to G2 (GSE). Three of these latter stars have $[\mathrm{Al/Fe}] > 0$, which is not consistent with the main feature that distinguishes GSE, but the other four have $[\mathrm{Al/Fe}] < 0$, and we postulate these as second-generation or chemically anomalous debris from GSE's own GCs. This is corroborated by their locus in Fig. 7.

Some of the differences between the results of our work and that of Fernández-Trincado et al. (2022) could also be explained by two factors: (i) these authors based their analysis on a proprietary catalog internally distributed to the collaboration, which contained the resulting `ASPCAP` solutions from the "…133…" runs, while our present study is based on the published `ASPCAP` run "…synspec_fix…"; and (ii) t-SNE does an *un*-supervised search for stars associated with each other in the used abundances space, while Fernández-Trincado et al. (2022) did a supervised one.

8.5. Group 5 (G5): New structure

Stars in Group 5 (G5) are located next to the metal-rich border of G2 (GSE) in many abundances versus [Fe/H] in the small panels of Fig. 1. This small structure appears frequently

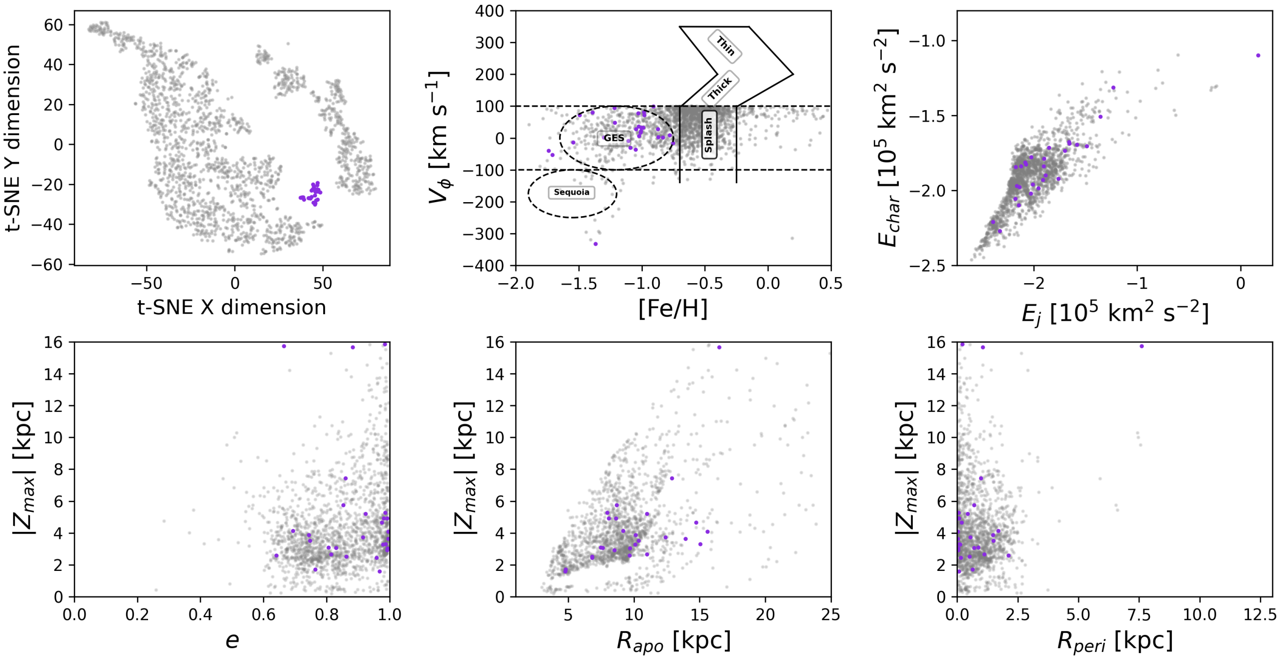

Fig. 8. G4 identified as N-C-O peculiar stars. Details are explained in Sect. 8.4. Panels and units are the same as those in Fig. 4.

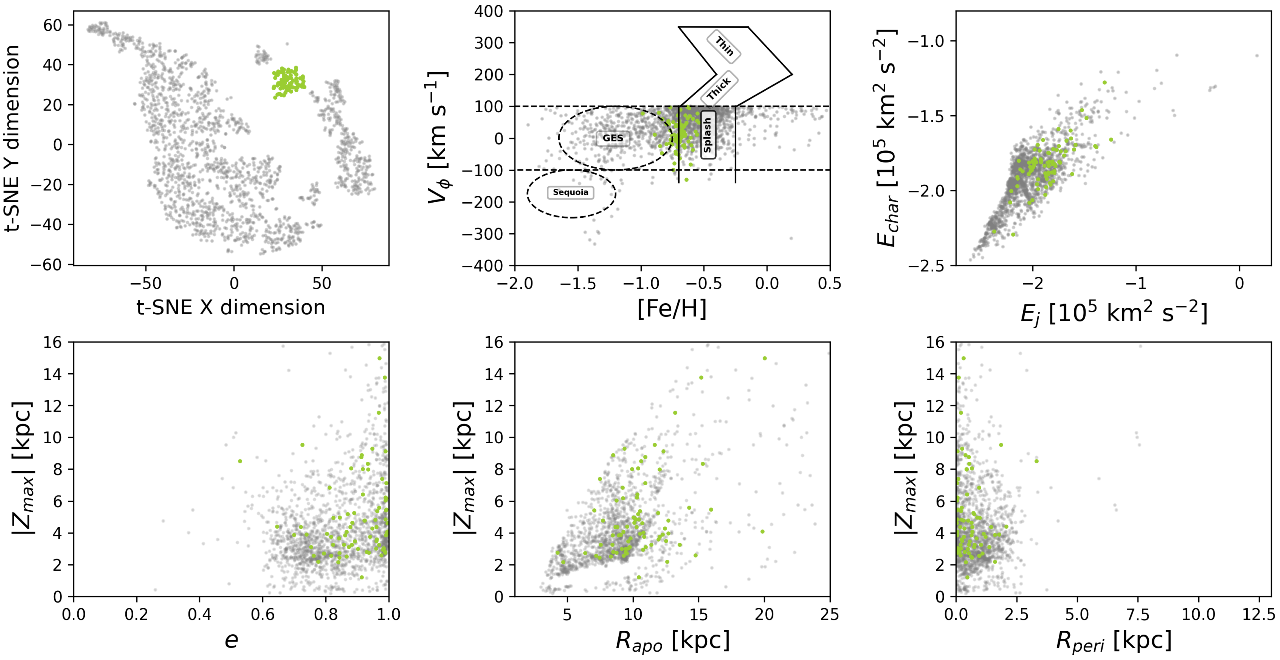

Fig. 9. G5, a new structure identified in this work and named Galileo 5 (G5). Details are explained in Sect. 8.5. Panels and units are the same as those in Fig. 4.

between G2 and G7 in the many t-SNE tests performed in this work. The G5 structure sits in a distinctive separate locus in the [Mg/Fe], [Si/Fe], and [Al/Fe] versus [Fe/H] planes. Compared to G1 (Splash), G5 is lower in oxygen, magnesium, silicon, carbon, and aluminum, although it has a G1-like metallicity. Dynamically, the values of $e, R_{\rm peri}, R_{\rm apo}$, and $|Z_{\rm max}|$ in G5 stars spread over the values spanned by the full dataset, but its energy and eccentricity value are slightly higher than G1 (see Fig. 2, lowest row, fifth panel). The G5 structure has mostly $-2.2 \lesssim E_J \lesssim -1.5$ and $-2.1 \lesssim E_{\rm char} \lesssim -1.6$ in the energy plot of Fig. 9. We notice one star, APOGEE ID 2M15113246+4813218, which is an outlier of G5 in [Mg/Fe, [Si/Fe], [Ca/Fe], and [Al/Fe], despite its location on the t-SNE plane being well within the footprint of this group.

Chemically, there is an overlap between G5 and the metal-poor side of the structure Eos described by Myeong et al. (2022). Unlike these authors' Eos, our G5 group does not overlap much with either Splash (G1) or GSE (G2). Our detection also extends the eccentricity range toward slightly lower values, as we did not cut on this parameter. Similar to the metal-poor half of Eos, G5 is lower in nickel than the MW disk, and low values of [Ni/Fe] have been linked to accretion from present-day dwarf

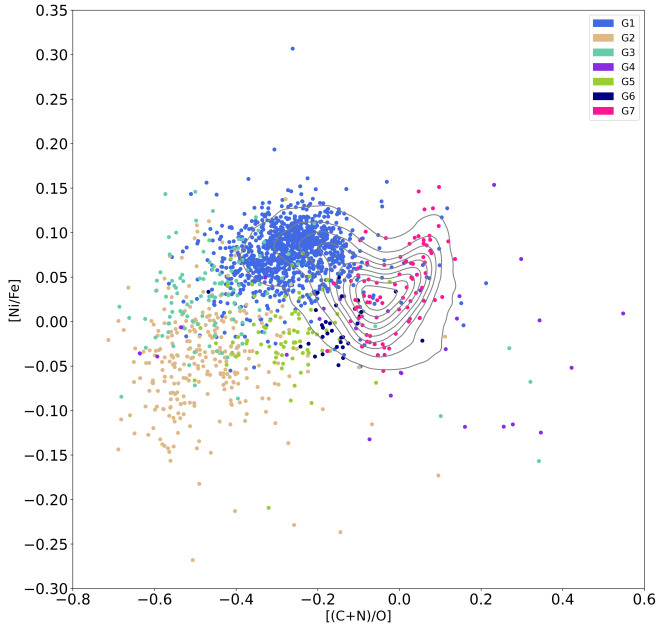

Fig. 10. [Ni/Fe] versus [(C+N)/O] plot, color-coded for each detected group. A kernel density estimation (KDE) plot of MW disk stars is shown with gray curves. These were selected from APOGEE DR17 following the same quality criteria as our halo stars, and $V_\phi \geq 100\,\mathrm{km\,s^{-1}}$. We note that $[\mathrm{Ni/Fe}] < 0$ and $[(\mathrm{C+N})/0] < -0.2$ have been linked to ex situ stars. In this work, G2 (Splash) and G5 (Galileo 5) fall in the locus of accreted stars. G6 (Galileo 6) overlaps with the thick-disk population.

spheroidal galaxies (Nissen & Schuster 2010), although these are more metal poor ($[\mathrm{Fe/H}] < -0.8$) than G5. The investigation of Montalbán et al. (2021) carefully dated ex situ stars with $e > 0.7$ that fall below the line $[\mathrm{Mg/Fe}] = 0.2\,[\mathrm{Fe/H}] + 0.05$, and found that they also exhibit sub-solar values of [Ni/Fe] and are slightly richer than GSE in [(C+N)/O], as seen in Fig. 10. These stars are found to be slightly younger than the high-Mg in situ halo stars. Most of our G5 stars fall below the abovementioned line and occupy the same locus in the plot [Ni/Fe] versus [(C+N)/O] as the confirmed ex situ low-Mg stars.

We did a cross-match between the 177 stars labeled as members of Eos by Myeong et al. (2022, priv. comm.) with our full sample. Only 49 were in common between both studies: 11 (22%) in G1, seven (14%) in G2, one (2%) in G3, zero (0%) in G4, 19 (39%) in G5, eigth (16%) in G6, and three (6%) in G7. Given their chemical similarity, it is not unexpected that a good portion of the common stars (55%) belongs to G5 and G6, yet a nonnegligible portion of them (22%) were assigned by t-SNE to G1 (Splash). Only 169 listed members of Eos have a $d_{\mathtt{StarHorse}}$, of which 65 are beyond 5 kpc, and therefore they are outside the volume sampled by our study.

While Myeong et al. (2022) propose that Eos originated from the gas polluted by the GSE and evolved to resemble the (outer) thin disk of the MW, we instead speculate that what they detected as the single structure Eos, was separated by t-SNE in our work into two portions, with the more metal-poor one being our G5 and of ex situ origin, and with the more metal-rich one which may correspond to our detected group G6 being described next. We postulate that G5 is a new structure, which we name Galileo 5, after the name of the project Galactic ArchaeoLogIcaL ExcavatiOns (GALILEO), thus retaining the G5 acronym.

8.6. Group 6 (G6): New structure or Eos

Stars in Group 6 (G6) are more metal-rich than G5 but more metal-poor than G7 (see Fig. 11), and they are slightly lower than G5 in all the α elements, as seen in Fig. 1. In several t-SNE tests, G6 stars were very close to or attached to G7. Similar to the previous group, G6 also occupies a distinctive separate locus in the [Mg/Fe], [Si/Fe], and [Al/Fe] versus [Fe/H] planes, and is more α-poor than G1 (Splash), overlapping partially with the abundance locus of thin-disk stars. Dynamically, similar to G5, G6 also exhibits high eccentricities, but is slightly less energetic than G5.

Figure 3 shows an element-by-element comparison between groups G1 (Splash), G2 (GSE), G5 (Galileo 5), and G6. For all the α elements, G1, G2, G5, and G6 go from richest to poorest in their mean values; for the iron-peak elements, aluminum, and carbon, G2, G5, and G6 follow an ascending trend, with G1 being either the richest or the second richest. G1 stands out as the richest structure in aluminum; G2 is easily separated from the other structures by its extremely low values in carbon, aluminum, and iron; G5 and G6 separate from each other in iron. We notice one star in G6, APOGEE ID 2M16045805-2418515, which sits a bit far from the locus of G6 in the t-SNE plane, and also it looks similar to an outlier in several abundances. This star was associated with G6 when running t-SNE only on the G5 and G6 data.

As commented on the previous subsection, G6 overlaps with the more metal-rich part of the reported structure Eos. Our data selection is more stringent than that of Myeong et al. (2022), and this may have helped t-SNE in separating G6 from G5 in [Fe/H]. In our data, G5 and G6 have distinctive values and a distribution of metallicity, as seen in Fig. 3, and therefore we trust they are separate structures. G6 abundances overlap with Aleph, as reported by Horta et al. (2023), but Aleph clearly follows a disk-like orbit, which is not the case for G6. When studying various halo substructures in the MW halo, Horta et al. (2023) explain that halo substructures with low nickel due to a low star formation rate can be identified by them having disk-like values of [Mn/Fe], which is the case for both G5 and G6. On the other hand, unlike G5, G6 overlaps in many abundances with the thick-disk population, although G6 stars' orbits are clearly more energetic. If we assume G6 stars were formed in the disk, a mechanism different or additional to the one that formed the thick disk is needed to keep G6 stars in those more energetic orbits. With G6 it is harder to ascertain if it has an ex situ origin or if it matches Eos, which is proposed to be formed in situ. We postulate that G6 is a likely new ex situ structure that formed by debris from a past merger event and we name it Galileo 6, keeping the G6 acronym. Nonetheless, we recognize there is also a chance of it being the Eos structure.

8.7. Group 7 (G7): Inner disk-like stars

Our Group 7 (G7) contains the richest stars of the whole dataset in the iron-peak elements, $[\mathrm{Fe/H}] \gtrsim -0.2$ and $[\mathrm{Mn/Fe}] \gtrsim -0.2$, and it is also poorer in all the α elements than all the other groups (see Fig. 12). G7 has the chemical signature of inner disk-like stars. About half of the G7 stars have low energy values, and this subset is also concentrated at $V_\phi \sim 80\,\mathrm{km\,s^{-1}}$, eccentricity at $e \sim 0.82, Z_{\max} \lesssim 2\,\mathrm{kpc}$, $R_{\mathrm{apo}} < 5\,\mathrm{kpc}$, and $R_{\mathrm{peri}} < 0.5\,\mathrm{kpc}$. Nonetheless, this subset does not cluster in a particular locus within the G7 footprint in the t-SNE plane, and not all the stars around $V_\phi \sim 80\,\mathrm{km\,s^{-1}}$ belong to this subset. A nonnegligible portion of the $V_\phi \sim 80\,\mathrm{km\,s^{-1}}$ stars have higher energies and

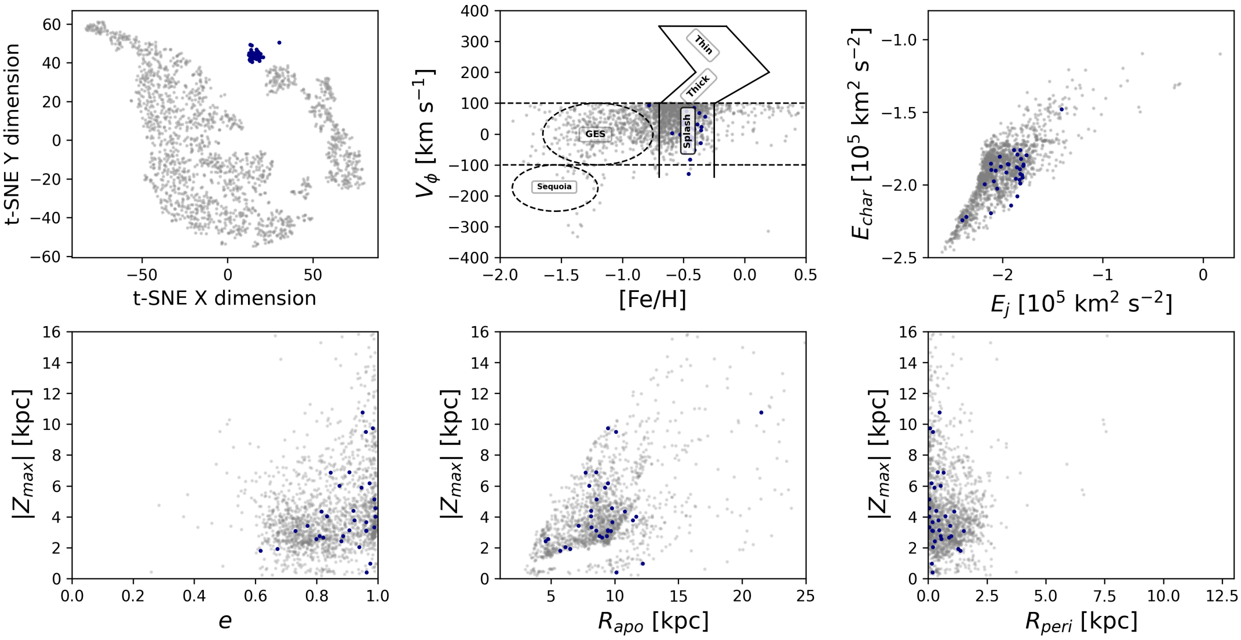

Fig. 11. G6, a new structure identified in this work and named Galileo 6 (G6). Details are explained in Sect. 8.6. Panels and units are the same as those in Fig. 4.

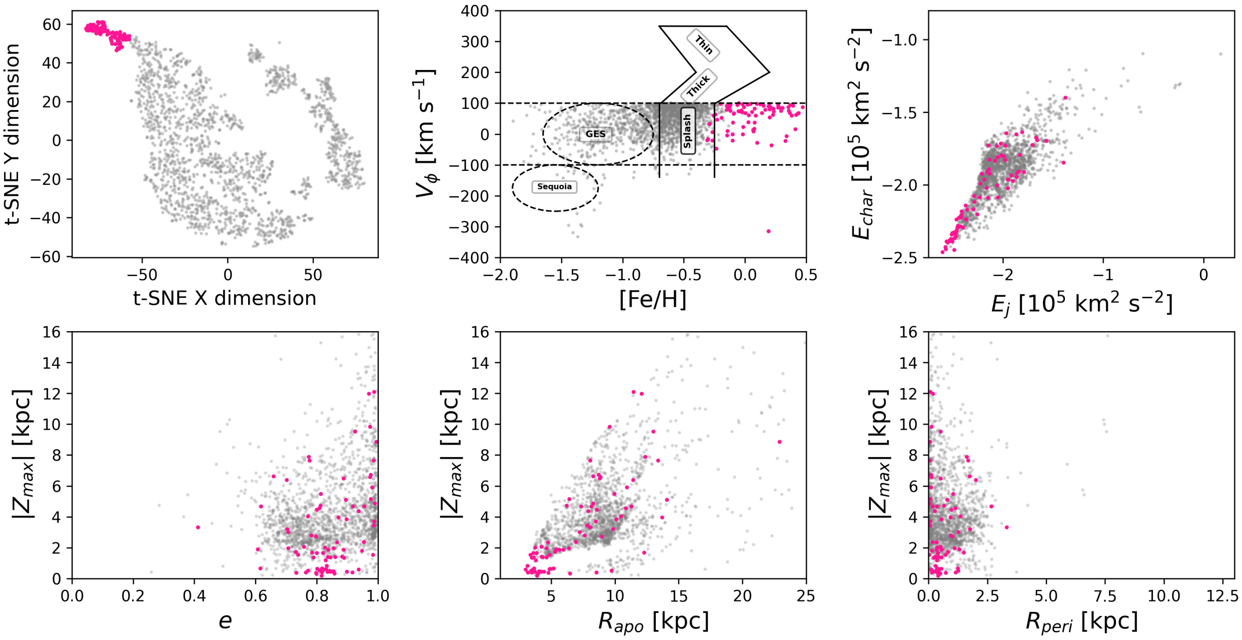

Fig. 12. G7 identified as inner disk-like stars. Details are explained in Sect. 8.7. Panels and units are the same as those in Fig. 4.

halo-like orbits. The energies of G7 stars in the $E_{\rm char}$ versus E_J plane fall within the footprint of MW Bulge GCs shown in Fernández-Trincado et al. (2021d) in their Fig. 4 panel d. In the t-SNE runs, we noticed that G7 kept close to G1 in several tests but it did separate sometimes, while G1 – despite experiencing some internal clustering – always retained its integrity. This was a hint of G7 being different from G1, which was independently confirmed by its dynamics.

9. Accretion origin for some structures

The plot [Mg/Mn] versus [Al/Fe], shown in Fig. 7, has been used by several authors (e.g., Das et al. 2020; Naidu et al. 2020) as an empirical diagnostic for the origin of structures as either in situ or accretion. Moreover, Horta et al. (2021) prove the validity of this plot with their chemical-evolution model. From the structures detected in this work, only G2 (which we identify as GSE) falls right on the locus of accreted structures in this plot, confirming its nature. Also, there are five G4 N-poor and C-poor stars that fall on this locus, which we postulate as GSE's GC first-generation debris. We also notice that, among the G2 stars, there are four that are N-rich and probable GSE's GC second-generation debris.

As discussed in Sects. 8.5 and 8.6, the plot of [Ni/Fe] versus [(C+N)/O] has been used by Montalbán et al. (2021) to pinpoint the loci of identified in situ and ex situ stars, for which also precise seismic ages were calculated. The ex situ stars are slightly younger than the in situ halo stars. Low values of [Ni/Fe] with disk-like values of [Mn/Fe] have been linked to satellite galaxies of the MW with low star formation rates, and such an abundance pattern has been observed in both G5 and G6 (see Fig. 10). Both reasons motivate us to postulate that G5 and G6 are newly identified accreted structures in the MW inner halo.

10. The Galactic bar pattern speed $\Omega_{\rm bar}$ role in the detected inner halo groups

In previous works studying peculiar chemical and/or kinematical structures in the solar vicinity (e.g., Naidu et al. 2020, 2022; Horta et al. 2023), the dynamical analysis was performed using axisymmetric gravitational potentials. This work is the first to analyze the role of a nonaxisymmetric potential, in this case, the one induced by the bar on the dynamics of the inner-halo groups detected – a relevant issue as they spend a good portion of their orbits in the volume occupied by the MW disk, and therefore some effect is expected. Our analysis considered $\Omega_{\rm bar} = 31, 41$, and 51 km s^{-1} kpc^{-1}. The results obtained for each of these values were very similar; we did not detect a visible or significant difference in the distribution of each of the dynamical parameters examined, $|Z_{\rm max}|, R_{\rm peri}, R_{\rm apo}, e, E_j$, and $E_{\rm char}$. Therefore, the bar pattern velocity does not have an effect on the dynamical behavior of the structures detected, and we conclude that its role seems to be absent or undetectable at the error level of the data.

11. Conclusions

From this investigation, we conclude that:

– We have used high-quality APOGEE abundances and *Gaia* astrometric data for 1742 red giant stars located within 5 kpc of the Sun, in order to detect coherent structures in the chemical-kinematic-dynamical space. We limited our study to stars not rotating with the disk, that is $V_\phi < 100$ km s^{-1}.

– We determined orbital parameters using the nonaxisymmetric galactic potential model `GravPot16`, which, together with the stellar components for the disk and halo of the MW, also includes a rotating boxy/peanut bar. The bar pattern velocity adopted was 41 km s^{-1} kpc. They were also determined for velocities 31 and 51 km s^{-1} kpc in order to measure the impact on the orbital parameters, since these stars may spend a significant portion of their orbits close to the Galactic plane. We did not detect significant variations in this regard.

– The search for structures was performed with the nonlinear algorithm of dimensional reduction known as t-SNE, using ten chemical-abundance ratios as input data: [Fe/H], [O/Fe], [Mg/Fe], [Si/Fe], [Ca/Fe], [C/Fe], [N/Fe], [Al/Fe], [Mn/Fe], and [Ni/Fe].

– Seven structures were detected, of which the following were traced to previously identified populations in the literature: G1 ∼ Splash, G2 ∼ GSE, G3 ∼ high-α heated-disk population, G4 ∼ N-C-O peculiar stars, and G7 ∼ inner disk-like stars. In the many tests performed, t-SNE occasionally separated the metal-poor high-α plateau portion of GSE from the metal-rich lower-α declining section at the so-called *knee*.

– No known structures were detected to be unambiguously similar to groups G5 and G6, and therefore we posit that they are new structures found in this work. These groups have low values of [Ni/Fe] but [Mn/Fe] with thin disk-like chemical abundances, which hints at an ex situ origin, that is to say an accreted satellite galaxy. Both G5 and G6 have Splash-like dynamics. G6 may also correspond to Eos, which was proposed by Myeong et al. (2022) to be formed in situ from gas contaminated by GSE.

– All groups except G4 (N-C-O peculiar stars) as detected by t-SNE in this work occupy distinctive loci in the [Mg/Fe], [Si/Fe], and [Al/Fe] versus [Fe/H] planes, with little overlap between the groups. We recommend using these abundances in future searches for these structures.

– We report three new N-rich stars and probable GC debris: APOGEE IDs 2M06572697+5543115, 2M21372238+1244305, and 2M16254515-2620462.

– The t-SNE algorithm was thoroughly tested on several issues. 1) A wide range of parameters were tested for the parameters `perplexity, early exaggeration`, and `number of iterations` to find those that produced the greatest visible separation in the t-SNE plane. 2) In those cases where the groups separation was not completely clean in the t-SNE plane, several procedures were applied: (i) t-SNE was applied iteratively, that is the outputs' t-SNE X and t-SNE Y were introduced as two additional inputs for a second run of the method; (ii) t-SNE was applied to subsamples in which two or more groups were suspected to coexist, to check if t-SNE separated them or not; and (iii) t-SNE was applied to the larger individual groups to determine whether or not they remained as a single entity. 3) Through Monte Carlo realizations considering the uncertainties in the abundances used, it was determined that the errors in these do not appreciably affect the groups detected by t-SNE.

Finally, the t-SNE algorithm proved to be very useful for exploring high-dimensional datasets, objectively separating real structures at various scales present in a large number of data, for example, small hidden groups with a very dominant feature, or large ensembles with more dispersed properties that overlap in some variables with other structures and whose boundaries are less evident. In any case, it is recommended to cross-check t-SNE results with independent data (e.g., dynamics) or previous findings, and also gauge the stability of the structures detected, as described above.

A table with all the relevant data for the full dataset, including a tag for each of the seven groups detected, is available at the CDS. A sample is shown in Table C.1.

Acknowledgements. This work has received funding from the grant support provided by Agencia Nacional de Investigación y Desarrollo de Chile (ANID) under the Proyecto Fondecyt Iniciación 2022 Agreement No. 11220340 (PI: José G. Fernández-Trincado) and from ANID under the Concurso de Fomento a la Vinculación Internacional para Instituciones de Investigación Regionales (Modalidad corta duración) Agreement No. FOVI210020 (PI: José G. Fernández-Trincado) and from the Joint Committee ESO-Government of Chile 2021 under the Agreement No. ORP 023/2021 (PI: José G. Fernández-Trincado) and from Becas Santander Movilidad Internacional Profesores 2022, Banco Santander Chile (PI: José G. Fernández-Trincado). T.C.B. acknowledges partial support for this work from grant PHY 14-30152; Physics Frontier Center/JINA Center for the Evolution of the Elements (JINA-CEE), and from OISE-1927130: The International Research Network for Nuclear Astrophysics (IReNA), awarded by the US National Science Foundation. D.M. also acknowledges support from CNPq/Brazil through project350104/2022-0. B.T. gratefully acknowledges support from the Natural Science Foundation of Guangdong Province under grant No. 2022A1515010732, the National Natural Science Foundation of China through grants No. 12233013, and the China Manned Space Project No. CMS-CSST-2021-B03. Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the US Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS website is `www.sdss.org`. SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofisica de Canarias, The Johns Hopkins

University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatory of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (`http://www.cosmos.esa.int/gaia`), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, `http://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.

References

Abdurro'uf, Accetta, K., Aerts, C., , et al. 2022, ApJS, 259, 35
Allende Prieto, C., Beers, T. C., Wilhelm, R., et al. 2006, ApJ, 636, 804
Amarante, J. A. S., Beraldo e Silva, L., Debattista, V. P., & Smith, M. C. 2020, ApJ, 891, L30
Anders, F., Chiappini, C., Santiago, B. X., et al. 2018, A&A, 619, A125
Anders, F., Khalatyan, A., Chiappini, C., et al. 2019, A&A, 628, A94
Barbá, R. H., Minniti, D., Geisler, D., et al. 2019, ApJ, 870, L24
Baumgardt, H., & Vasiliev, E. 2021, MNRAS, 505, 5957
Beaton, R. L., Oelkers, R. J., Hayes, C. R., et al. 2021, AJ, 162, 302
Beers, T. C., & Christlieb, N. 2005, ARA&A, 43, 531
Beers, T. C., Carollo, D., Ivezić, Ž., et al. 2012, ApJ, 746, 34
Belokurov, V., Erkal, D., Evans, N. W., Koposov, S. E., & Deason, A. J. 2018, MNRAS, 478, 611
Belokurov, V., Sanders, J. L., Fattahi, A., et al. 2020, MNRAS, 494, 3880
Blanton, M. R., Bershady, M. A., Abolfathi, B., et al. 2017, AJ, 154, 28
Bonaca, A., Conroy, C., Wetzel, A., Hopkins, P. F., & Kereš, D. 2017, ApJ, 845, 101
Bonaca, A., Conroy, C., Cargile, P. A., et al. 2020, ApJ, 897, L18
Bovy, J., Leung, H. W., Hunt, J. A. S., et al. 2019, MNRAS, 490, 4740
Bowen, I. S., Vaughan, A. H., & J., 1973, Appl. Opt., 12, 1430
Brunthaler, A., Reid, M. J., Menten, K. M., et al. 2011, Astron. Nachr., 332, 461
Buder, S., Lind, K., Ness, M. K., et al. 2022, MNRAS, 510, 2407
Carollo, D., Beers, T. C., Lee, Y. S., et al. 2007, Nature, 450, 1020
Carollo, D., Beers, T. C., Chiba, M., et al. 2010, ApJ, 712, 692
Chiba, M., & Beers, T. C. 2000, AJ, 119, 2843
Combes, F. 2017, in SF2A-2017: Proc. Annual Meeting of the French Society of Astronomy and Astrophysics, eds. C. Reylé, P. Di Matteo, & F. Herpin (Paris), 223
Cunha, K., Smith, V. V., Hasselquist, S., et al. 2017, ApJ, 844, 145
Das, P., Hawkins, K., & Jofré, P. 2020, MNRAS, 493, 5195
Di Matteo, P., Haywood, M., Lehnert, M. D., et al. 2019, A&A, 632, A4
Einasto, J. 1979, IAU Symp., 84, 451
Fehlberg, E. 1968, NASA Technical Report, 315
Fernández-Trincado, J. G., Robin, A. C., Moreno, E., et al. 2016, ApJ, 833, 132
Fernández-Trincado, J. G., Zamora, O., García-Hernández, D. A., et al. 2017, ApJ, 846, L2
Fernández-Trincado, J. G., Beers, T. C., Tang, B., et al. 2019a, MNRAS, 488, 2864
Fernández-Trincado, J. G., Mennickent, R., Cabezas, M., et al. 2019b, A&A, 631, A97
Fernández-Trincado, J. G., Beers, T. C., Minniti, D., et al. 2020a, ApJ, 903, L17
Fernández-Trincado, J. G., Beers, T. C., Minniti, D., et al. 2020b, A&A, 643, L4
Fernández-Trincado, J. G., Chaves-Velasquez, L., Pérez-Villegas, A., et al. 2020c, MNRAS, 495, 4113
Fernández-Trincado, J. G., Minniti, D., Beers, T. C., et al. 2020d, A&A, 643, A145
Fernández-Trincado, J. G., Beers, T. C., Minniti, D., et al. 2021a, A&A, 647, A64
Fernández-Trincado, J. G., Beers, T. C., Minniti, D., et al. 2021b, A&A, 648, A70
Fernández-Trincado, J. G., Beers, T. C., Queiroz, A. B. A., et al. 2021c, ApJ, 918, L37
Fernández-Trincado, J. G., Minniti, D., Souza, S. O., et al. 2021d, ApJ, 908, L42
Fernández-Trincado, J. G., Beers, T. C., Barbuy, B., et al. 2022, A&A, 663, A126
Feuillet, D. K., Sahlholdt, C. L., Feltzing, S., & Casagrande, L. 2021, MNRAS, 508, 1489
Fiteni, K., Caruana, J., Amarante, J. A. S., Debattista, V. P., & Beraldo e Silva, L., 2021, MNRAS, 503, 1418
Freeman, K., & Bland-Hawthorn, J. 2002, ARA&A, 40, 487
Gaia Collaboration (Brown, A. G. A., et al.) 2021, A&A, 649, A1
Gaia Collaboration (Vallenari, A., et al.) 2023, A&A, 674, A1
García Pérez, A. E., Allende Prieto, C., Holtzman, J. A., et al. 2016, AJ, 151, 144
GRAVITY Collaboration (Abuter, R., et al.) 2019, A&A, 625, L10
Gunn, J. E., Siegmund, W. A., Mannery, E. J., et al. 2006, AJ, 131, 2332
Hasselquist, S., Shetrone, M., Cunha, K., et al. 2016, ApJ, 833, 81
Hasselquist, S., Hayes, C. R., Lian, J., et al. 2021, ApJ, 923, 172
Hawkins, K., Jofré, P., Masseron, T., & Gilmore, G. 2015, MNRAS, 453, 758
Hayes, C. R., Majewski, S. R., Shetrone, M., et al. 2018, ApJ, 852, 49
Haywood, M., Di Matteo, P., Lehnert, M. D., et al. 2018, ApJ, 863, 113
Helmi, A., White, S. D. M., de Zeeuw, P. T., & Zhao, H. 1999, Nature, 402, 53
Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, Nature, 563, 85
Hinton, G., & Roweis, S. 2003, Adv. Neural Inf. Process. Syst., 15, 857
Holtzman, J. A., Shetrone, M., Johnson, J. A., et al. 2015, AJ, 150, 148
Holtzman, J. A., Hasselquist, S., Shetrone, M., et al. 2018, AJ, 156, 125
Horta, D., Schiavon, R. P., Mackereth, J. T., et al. 2021, MNRAS, 500, 1385
Horta, D., Schiavon, R. P., Mackereth, J. T., et al. 2023, MNRAS, 520, 5671
Jönsson, H., Allende Prieto, C., Holtzman, J. A., et al. 2018, AJ, 156, 126
Jönsson, H., Holtzman, J. A., Allende Prieto, C., et al. 2020, AJ, 160, 120
Jurić, M., Ivezić, Ž., Brooks, A., et al. 2008, ApJ, 673, 864
Koppelman, H., Helmi, A., & Veljanoski, J. 2018, ApJ, 860, L11
Koppelman, H. H., Helmi, A., Massari, D., Price-Whelan, A. M., & Starkenburg, T. K. 2019a, A&A, 631, L9
Koppelman, H. H., Helmi, A., Massari, D., Roelenga, S., & Bastian, U. 2019b, A&A, 625, A5
Kruijssen, J. M. D., Pfeffer, J. L., Chevance, M., et al. 2020, MNRAS, 498, 2472
Lindegren, L., Hernández, J., Bombrun, A., et al. 2018, A&A, 616, A2
Linderman, G. C., & Steinerberger, S. 2017, ArXiv e-prints [arXiv:`1706.02582`]
Mackereth, J. T., & Bovy, J. 2020, MNRAS, 492, 3631
Mackereth, J. T., Schiavon, R. P., Pfeffer, J., et al. 2019, MNRAS, 482, 3426
Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, AJ, 154, 94
Mészáros, S., Masseron, T., García-Hernández, D. A., et al. 2020, MNRAS, 492, 1641
Montalbán, J., Mackereth, J. T., Miglio, A., et al. 2021, Nat. Astron., 5, 640
Moreno, E., Pichardo, B., & Schuster, W. J. 2015, MNRAS, 451, 705
Myeong, G. C., Evans, N. W., Belokurov, V., Sanders, J. L., & Koposov, S. E. 2018, ApJ, 863, L28
Myeong, G. C., Belokurov, V., Aguado, D. S., et al. 2022, ApJ, 938, 21
Naidu, R. P., Conroy, C., Bonaca, A., et al. 2020, ApJ, 901, 48
Naidu, R. P., Ji, A. P., Conroy, C., et al. 2022, ApJ, 926, L36
Nidever, D. L., Holtzman, J. A., Allende Prieto, C., et al. 2015, AJ, 150, 173
Nissen, P. E., & Schuster, W. J. 2010, A&A, 511, L10
Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, J. Mach. Learn. Res., 12, 2825
Queiroz, A. B. A., Anders, F., Santiago, B. X., et al. 2018, MNRAS, 476, 2556
Queiroz, A. B. A., Anders, F., Chiappini, C., et al. 2020, A&A, 638, A76
Queiroz, A. B. A., Anders, F., Chiappini, C., et al. 2023, A&A, 673, A155
Rebonato, R., & Jäckel, P. 2000, Journal of Risk, 2, 17
Recio-Blanco, A., de Laverny, P., Palicio, P. A., et al. 2023, A&A, 674, A29
Reid, M. J., & Brunthaler, A. 2020, ApJ, 892, 39
Rim, P., Steinhardt, C., & Clark, T. 2022, in AAS Meet. Abstr., 54, 241.38
Robin, A. C., Reylé, C., Derrière, S., & Picaud, S. 2003, A&A, 409, 523
Robin, A. C., Marshall, D. J., Schultheis, M., & Reylé, C. 2012, A&A, 538, A106
Robin, A. C., Reylé, C., Fliri, J., et al. 2014, A&A, 569, A13
Sanders, J. L., Smith, L., & Evans, N. W. 2019, MNRAS, 488, 4552
Santana, F. A., Beaton, R. L., Covey, K., et al. 2021, AJ, 162, 303
Santiago, B. X., Brauer, D. E., Anders, F., et al. 2016, A&A, 585, A42
Schiavon, R. P., Zamora, O., Carrera, R., et al. 2017, MNRAS, 465, 501
Shetrone, M., Bizyaev, D., Lawler, J. E., et al. 2015, ApJS, 221, 24
Smith, V. V., Bizyaev, D., Cunha, K., et al. 2021, AJ, 161, 254
Tang, B., Fernández-Trincado, J. G., Geisler, D., et al. 2018, ApJ, 855, 38
Ting, Y.-S., Freeman, K. C., Kobayashi, C., De Silva, G. M., & Bland-Hawthorn, J. 2012, MNRAS, 421, 1231
van der Maaten, L., & Hinton, G. 2008, J. Mach. Learn. Res., 9, 2579
Vasiliev, E., & Baumgardt, H. 2021, MNRAS, 505, 5978
Vera, M., Alonso, S., & Coldwell, G. 2016, A&A, 595, A63
Verma, M., Matijevič, G., Denker, C., et al. 2021, ApJ, 907, 54
Wattenberg, M., Viégas, F., & Johnson, I. 2016, Distill, `http://doi.org/10.23915/distill.00002`
Wilson, J. C., Hearty, F. R., Skrutskie, M. F., et al. 2019, PASP, 131, 055001
Zamora, O., García-Hernández, D. A., Allende Prieto, C., et al. 2015, AJ, 149, 181
Zasowski, G., Johnson, J. A., Frinchaboy, P. M., et al. 2013, AJ, 146, 81
Zasowski, G., Cohen, R. E., Chojnowski, S. D., et al. 2017, AJ, 154, 198

Appendix A: t-SNE tests

The following figures show the various tests performed with t-SNE to assess its performance, and to make the final selection of adjacent groups that were not clearly separated in the t-SNE plane chosen, as described in Section 7. Figure A.1 shows the effect of varying the input parameters: 25 <`perplexity`< 40, and 80 <`early_exageration`< 130. Figure A.2 shows the *iterated* t-SNE plane, and Figure A.3 shows the effect of abundance errors. These tests helped to identify truly stable groups that t-SNE kept finding regardless of the changes introduced.

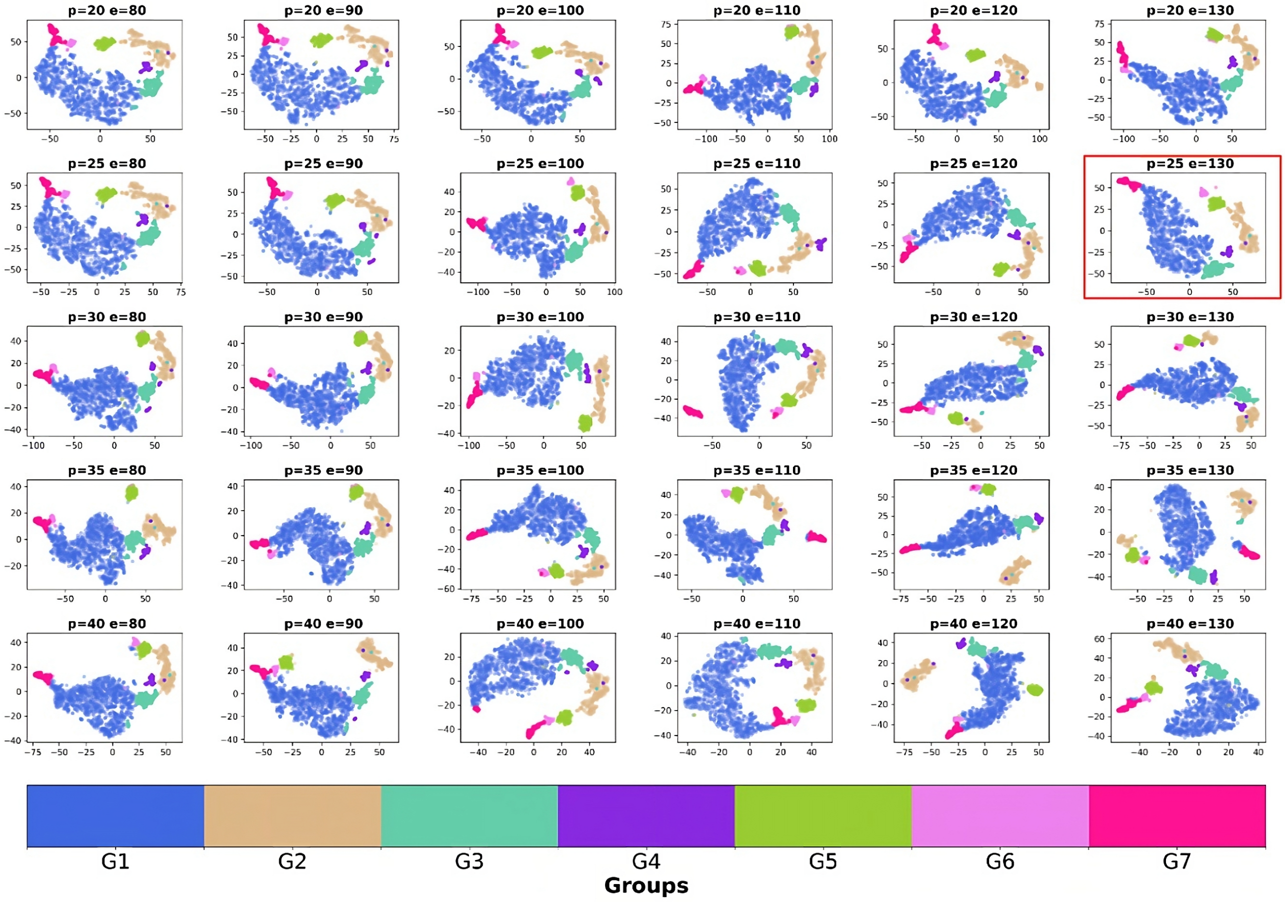

Fig. A.1. t-SNE plane for different input parameter values, as indicated in each panel. Horizontal and vertical axes correspond to the t-SNE X and Y dimensions, respectively. The panel marked in red is the final one chosen — i.e. Figure 1 — and best to select the structures found in this investigation.

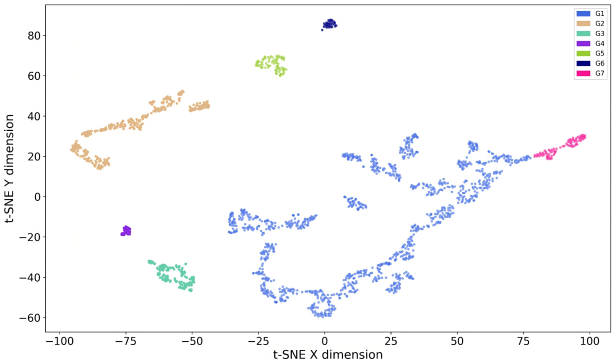

Fig. A.2. Iterated t-SNE plane. The values of t-SNE X and t-SNE Y from Figure 1 in the central panel were fed as additional data to another run of t-SNE, so as to check if adjacent groups in the original t-SNE plane were separated or not, and to determine where the borders occur. It is important to notice how the overall shape is kept, but in fact in this case t-SNE flipped the orientation of the points in the plane.

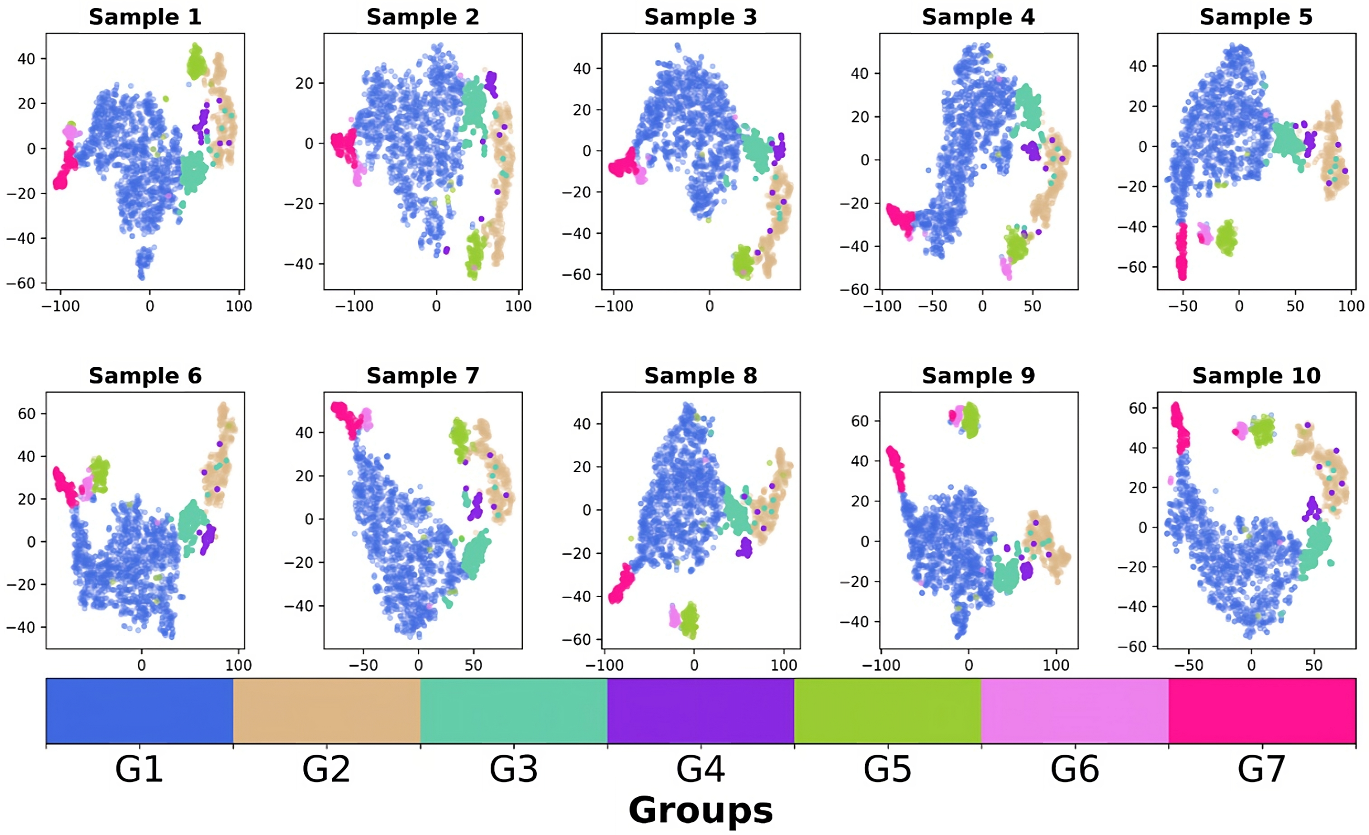

Fig. A.3. Effect of abundance errors on the t-SNE plane. Horizontal and vertical axes correspond to the t-SNE X and Y dimensions, respectively. The Monte-Carlo realizations of the abundances were used as inputs into the t-SNE algorithm to check if the groups selected in Figure 1 were still associated and kept as coherent structures.

Appendix B: Summary of chemical abundances and dynamical parameters of each detected group

Table B.1 contains the median value and median absolute deviation median($X-\mathrm{median}(X)$) of each the abundances and dynamical parameters studied in this work. Table B.2 shows the percentage of P, P-R, and R orbits per group. Readers can refer to Section 8 and the subsections within for details.

Table B.1. Chemodynamical properties of the detected groups.

	G1	G2	G3	G4	G5	G6	G7
[C/Fe]	+0.087 ±0.035	-0.366 ±0.077	-0.124 ±0.095	-0.239 ±0.115	-0.097 ±0.037	-0.019 ±0.037	+0.034 ±0.034
[N/Fe]	+0.120 ±0.044	+0.155 ±0.099	+0.195 ±0.082	+0.772 ±0.153	+0.134 ±0.066	+0.114 ±0.027	+0.235 ±0.063
[O/Fe]	+0.349 ±0.040	+0.305 ±0.067	+0.429 ±0.071	+0.212 ±0.122	+0.241 ±0.046	+0.159 ±0.028	+0.068 ±0.043
[Mg/Fe]	+0.339 ±0.026	+0.195 ±0.053	+0.327 ±0.028	+0.176 ±0.065	+0.162 ±0.040	+0.129 ±0.018	+0.052 ±0.049
[Al/Fe]	+0.265 ±0.039	-0.187 ±0.062	+0.170 ±0.088	+0.106 ±0.215	+0.031 ±0.041	+0.052 ±0.044	+0.007 ±0.056
[Si/Fe]	+0.240 ±0.027	+0.201 ±0.041	+0.334 ±0.045	+0.178 ±0.056	+0.109 ±0.020	+0.069 ±0.014	+0.031 ±0.031
[Ca/Fe]	+0.172 ±0.034	+0.186 ±0.049	+0.258 ±0.029	+0.229 ±0.058	+0.098 ±0.033	+0.041 ±0.031	-0.025 ±0.035
[Mn/Fe]	-0.136 ±0.034	-0.297 ±0.055	-0.279 ±0.054	-0.321 ±0.081	-0.166 ±0.025	-0.093 ±0.015	+0.026 ±0.079
[Fe/H]	-0.570 ±0.103	-1.230 ±0.197	-1.105 ±0.105	-1.027 ±0.136	-0.692 ±0.042	-0.455 ±0.031	+0.058 ±0.170
[Ni/Fe]	+0.079 ±0.017	-0.040 ±0.033	+0.028 ±0.030	-0.007 ±0.039	-0.016 ±0.020	-0.009 ±0.019	+0.039 ±0.034
[Mg/Mn]	+0.479 ±0.050	+0.497 ±0.087	+0.619 ±0.055	+0.541 ±0.085	+0.311 ±0.048	+0.218 ±0.022	-0.007 ±0.157
$R_{peri}{}^{(a)}$	0.808 ±0.529	0.457 ±0.360	0.688 ±0.542	0.471 ±0.406	0.414 ±0.315	0.372 ±0.211	0.391 ±0.228
$R_{apo}{}^{(a)}$	8.673 ±1.268	11.759 ±2.724	8.890 ±1.153	9.889 ±2.340	10.212 ±1.354	9.116 ±0.974	5.955 ±2.256
e	0.811 ±0.093	0.933 ±0.053	0.830 ±0.089	0.920 ±0.070	0.923 ±0.061	0.922 ±0.052	0.835 ±0.059
$\|Z_{max}\|^{(a)}$	3.303 ±0.914	5.065 ±2.075	3.857 ±1.202	3.681 ±1.046	4.183 ±1.183	3.534 ±0.998	1.894 ±1.431
$E_j{}^{(b)}$	-2.115 ±0.100	-1.789 ±0.166	-2.003 ±0.154	-1.934 ±0.207	-1.867 ±0.116	-1.877 ±0.077	-2.284 ±0.218
$E_{char}{}^{(b)}$	-1.927 ±0.097	-1.747 ±0.125	-1.887 ±0.092	-1.838 ±0.132	-1.813 ±0.077	-1.898 ±0.071	-2.117 ±0.211

Notes. Median and median absolute deviation of each of the chemical abundances and dynamical parameters investigated in this work, and plotted in Figs. 1 to 10, for each detected group. (a)Measured in kpc. (b)Measured in 10^5 km^2 s^{-2}.

Table B.2. Orbital orientation of each group.

G(a)	Name	# of P(b)	# of P-R(c)	# of R(d)	Total
1	Splash	851 (76%)	175 (16%)	99 (8%)	1125
2	GSE	136 (49%)	55 (20%)	89 (31%)	280
3	High-α	61 (54%)	19 (17%)	34 (29%)	114
4	N-rich	13 (46%)	9 (32%)	6 (22%)	28
5	Galileo 5	39 (49%)	17 (21%)	24 (30%)	80
6	Galileo 6	14 (44%)	6 (19%)	12 (37%)	32
7	Inner Disk	64 (77%)	10 (12%)	9 (11%)	83

Notes. Number and percentage of stars in each group detected according to their orbit orientation. (a)G = group number assigned in this work. (b)P = prograde. (c)P-R = prograde-retrograde. (d)R = retrograde.

Appendix C: Supplementary online data

A catalog with all the relevant information, including a tag for each detected group, is available online. A sample table showing the column labels is shown in Table C.1.

Table C.1. Relevant information for each star used in this investigation, including a tag for each detected group.

ID #	Column name	Units	Column Description
1	APOGEE_ID		APOGEE id
2	RA	deg	$\alpha(J2000)$
3	DEC	deg	$\delta(J2000)$
4	SNR	pixel^{-1}	Spectral signal-to-noise
5	RV	km s^{-1}	APOGEE-2 radial velocity
6	VSCATTER	km s^{-1}	APOGEE-2 radial velocity scatter
7	LOGG	[cgs]	Surface gravity
8	ERROR_LOGG	[cgs]	Uncertainty in LOGG
9	C_FE		[C/Fe] from ASPCAP
10	ERROR_C_FE		Uncertainty in [C/Fe]
11	N_FE		[N/Fe] from ASPCAP
12	ERROR_N_FE		Uncertainty in [N/Fe]
13	O_FE		[O/Fe] from ASPCAP
14	ERROR_O_FE		Uncertainty in [O/Fe]
15	MG_FE		[Mg/Fe] from ASPCAP
16	ERROR_MG_FE		Uncertainty in [Mg/Fe]
17	AL_FE		[Al/Fe] from ASPCAP
18	ERROR_AL_FE		Uncertainty in [Al/Fe]
19	SI_FE		[Si/Fe] from ASPCAP
20	ERROR_SI_FE		Uncertainty in [Si/Fe]
21	CA_FE		[Ca/Fe] from ASPCAP
22	ERROR_CA_FE		Uncertainty in [Ca/Fe]
23	MN_FE		[Mn/Fe] from ASPCAP
24	ERROR_MN_FE		Uncertainty in [Mn/Fe]
25	FE_H		[Fe/H] from ASPCAP
26	ERROR_FE_H		Uncertainty in [Fe/H]
27	NI_FE		[Ni/Fe] from ASPCAP
28	ERROR_NI_FE		Uncertainty in [Ni/Fe]
29	VR	km s^{-1}	Galactocentric radial velocity
30	VPHI	km s^{-1}	Galactocentric azimuthal velocity
31	PERIGALACTICON	kpc	Perigalactocentric distance
32	ERROR_PERIGALACTICON	kpc	Uncertainty in PERIGALACTICON
33	APOGALACTICON	kpc	Apogalactocentric distance
34	ERROR_APOGALACTICON	kpc	Uncertainty in APOGALACTICON
35	Eccentricity		Orbital eccentricity
36	ERROR_Eccentricity		Uncertainty in eccentricity
37	$\lvert Zmax \rvert$	kpc	Maximum vertical excursion from the Galactic plane
38	ERROR_Zmax	kpc	Uncertainty in Zmax
39	E_j	10^2 km^2 s^{-2}	Jacobi energy
40	ERROR_E_j	10^2 km^2 s^{-2}	Uncertainty in Jacobi energy
41	E_{char}	10^2 km^2 s^{-2}	Characteristic orbital energy
42	ERROR_E_{char}	10^2 km^2 s^{-2}	Uncertainty in characteristic orbital energy
43	L_{zmin}	10^1 km^2 s^{-2} kpc	Minimum z-component of angular momentum
44	ERROR_L_{zmin}	10^1 km^2 s^{-2} kpc	Uncertainty in minimum z-component of angular momentum
45	L_{zmax}	10^1 km^2 s^{-2} kpc	Maximum z-component of angular momentum
46	ERROR_L_{zmax}	10^1 km^2 s^{-2} kpc	Uncertainty in maximum z-component of angular momentum
47	GAIADR3_SOURCE_ID		*Gaia* DR3 source id
48	ruwe		*Gaia* DR3 renormalized unit weight error
49	d_STARHORSE	kpc	Bayesian StarHorse distance, 50th percentile
50	ERROR_d_STARHORSE	kpc	Uncertainty in d_STARHORSE, ERROR = (84th-16th)/2 percentile
51	pmRA_GaiaDR3	mas yr^{-1}	μ_α cos (δ) from *Gaia* DR3
52	ERROR_pmRA_GaiaDR3	mas yr^{-1}	Uncertainty in μ_α cos (δ) from *Gaia* EDR3
53	pmDEC_GaiaDR3	mas yr^{-1}	μ_α from *Gaia* EDR3
54	ERROR_pmDEC_GaiaDR3	mas yr^{-1}	Uncertainty in μ_α from *Gaia* EDR3
55	Group		Group as determined by t-SNE