Astronomy
&
Astrophysics

The stellar halo in Local Group Hestia simulations

III. Chemical abundance relations for accreted and in situ stars

Sergey Khoperskov[1] , Ivan Minchev[1], Noam Libeskind[1,2], Vasily Belokurov[3,4], Matthias Steinmetz[1],
Facundo A. Gomez[5,6], Robert J. J. Grand[7,8,9], Yehuda Hoffman[10], Alexander Knebe[11,12,13], Jenny G. Sorce[14,15,1],
Martin Spaare[16,1], Elmo Tempel[17,18], and Mark Vogelsberger[19]

[1] Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, 14482 Potsdam, Germany
   e-mail: sergey.khoperskov@gmail.com
[2] Univ. Lyon, Univ. Claude Bernard Lyon 1, CNRS, IP2I Lyon/IN2P3, IMR 5822, 69622 Villeurbanne, France
[3] Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK
[4] Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA
[5] Instituto de Investigación Multidisciplinar en Ciencia y Tecnología, Universidad de La Serena, Raúl Bitrán 1305, La Serena, Chile
[6] Departamento de Astronomía, Universidad de La Serena, Av. Juan Cisternas 1200 Norte, La Serena, Chile
[7] Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str 1, 85748 Garching, Germany
[8] Instituto de Astrofísica de Canarias, Calle Váa Láctea s/n, 38205 La Laguna, Tenerife, Spain
[9] Departamento de Astrofísica, Universidad de La Laguna, Av. del Astrofísico Francisco Sánchez s/n, 38206 La Laguna, Tenerife,
   Spain
[10] Racah Institute of Physics, Hebrew University, Jerusalem 91904, Israel
[11] Departamento de Física Teórica, Módulo 15, Facultad de Ciencias, Universidad Autónoma de Madrid, 28049 Madrid, Spain
[12] Centro de Investigación Avanzada en Física Fundamental (CIAFF), Facultad de Ciencias, Universidad Autónoma de Madrid,
   28049 Madrid, Spain
[13] International Centre for Radio Astronomy Research, University of Western Australia, 35 Stirling Highway, Crawley,
   Western Australia 6009, Australia
[14] Université Paris-Saclay, CNRS, Institut d'Astrophysique Spatiale, 91405 Orsay, France
[15] Univ. Lille, CNRS, Centrale Lille, UMR 9189 CRIStAL, 59000 Lille, France
[16] Institut für Physik und Astronomie, Universität Potsdam, Campus Golm, Haus 28, Karl-Liebknecht Straße 24-25, 14476 Potsdam,
   Germany
[17] Tartu Observatory, University of Tartu, Observatooriumi 1, 61602 Tõravere, Estonia
[18] Estonian Academy of Sciences, Kohtu 6, 10130 Tallinn, Estonia
[19] Department of Physics, Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge,
   MA 02139, USA

ABSTRACT

Stellar chemical abundances and kinematics provide key information for recovering the assembly history of galaxies. In this work we explore the chemo-chrono-kinematics of accreted and in situ stellar populations, by analyzing six M31/Milky Way (MW) analogues from the HESTIA suite of cosmological hydrodynamics zoom-in simulations of the Local Group. We show that elemental abundances ([Fe/H], [Mg/Fe]) of merger debris in the stellar haloes are chemically distinct from the survived dwarf galaxies, in that they are [α/Fe]-enhanced and have lower metallicity in the same stellar mass range. Therefore, mergers debris have abundances expected for stars originating from dwarfs that had their star formation activity quenched at early times. Accreted stellar haloes, including individual debris, reveal [Fe/H] and [Mg/Fe] gradients in the $E - L_z$ plane, with the most metal-rich, [α/Fe]-poor stars, which have formed in the inner parts of the disrupted systems before the merger, contributing mainly to the central regions of the host galaxies. This results in negative metallicity gradients in the accreted components of stellar haloes at $z = 0$, seen also for the individual merger debris. We suggest, therefore, that abundance measurements of halo stars in the inner MW will allow constraining better the parameters, such as the mass and merger time, of MW's most massive merger Gaia-Sausage-Enceladus. The metallicity distribution functions (MDFs) of the individual debris show several peaks and the majority of debris have lower metallicity than the in situ stars in the prograde part of the $E - L_z$ space. At the same time, non-rotating and retrograde accreted populations are very similar to the in situ stars in terms of [Fe/H] abundance. Prograde accreted stars show a prominent knee in the [Fe/H]–[Mg/Fe] plane, reaching up to solar [Mg/Fe], while retrograde stars typically contribute to the high-[Mg/Fe] sequence only. We find that the most metal-poor stars ([Fe/H] $\lesssim -1$) of the HESTIA galaxies exhibit net rotation up to $80\,\mathrm{km\,s^{-1}}$, which is consistent with the Aurora population recently identified in the MW. At higher metallicities ([Fe/H] $\approx -0.5 \pm 0.1$) we detect a sharp transition (spin-up) from the turbulent phase to a regular disk-like rotation. Different merger debris appear similar in the [Fe/H]–[Mg/Fe] plane, thus making it difficult to identify individual events. However, combining a set of abundances, and especially stellar age, makes it possible to distinguish between different debris.

Key words. galaxies: evolution – galaxies: halos – galaxies: kinematics and dynamics – galaxies: structure

1. Introduction

Although a substantial amount of stellar mass is formed in situ (Marchesini et al. 2009, 2010; Brammer et al. 2011), galaxies also grow via hierarchical clustering of smaller systems. Their stellar content at $z = 0$ represents all the systems accreted at different epochs (White & Frenk 1991; Lacey & Cole 1993; Springel et al. 2005; Fakhouri et al. 2010; Frenk & White 2012). Earlier mergers, happening at high redshifts, appear to contribute to the diffuse inner stellar halo, while more recent ones can be seen as coherent streams, shells, and other unmixed structures (Ibata et al. 1994; Johnston et al. 1995; Bullock & Johnston 2005; Abadi et al. 2006; Belokurov et al. 2006; Fardal et al. 2007; Cooper et al. 2010; Martínez-Delgado et al. 2012; Martin et al. 2014; Shipp et al. 2018; van Dokkum et al. 2019; Carlin et al. 2019). In this scenario, stellar haloes are made of different stellar populations that spatially overlap with each other, thus making it hard to uncover their parental galaxy remnants. Using kinematic information does allow for the capture of some individual stellar substructures (Helmi et al. 1999; Johnston et al. 2001; Helmi 2004; Smith et al. 2009; Belokurov et al. 2018), however, since merging galaxies are affected by the dynamical friction and evolving gravitational field. As a result, different merger debris can be fully phase-mixed with each other (Knebe et al. 2005; Vogelsberger et al. 2008; Li & Helmi 2008; Gómez et al. 2013; Amorisco 2017; Tissera et al. 2018). In this situation, the chemical abundances of stars are the only fossil records of the physical conditions of the star formation in the galaxies at different epochs. Therefore, an analysis of the chemical abundance patterns of resolved stellar populations in the halo with the phase-space data makes it possible to extract more precise information on galactic building blocks (Font et al. 2006; Deason et al. 2016; Roederer et al. 2018; Helmi 2020; Cunningham et al. 2022).

Recent studies of the Milky Way (MW) have undoubtedly shown that the inner stellar halo is dominated by accreted stars from a single dwarf galaxy (Gaia-Sausage-Enceladus, GSE; Belokurov et al. 2018; Haywood et al. 2018; Helmi et al. 2018) accreted at $10-11$ Gyr ago along (Gallart et al. 2019), with dynamically heated in situ stars (Di Matteo et al. 2019; Belokurov et al. 2020). However, these two main building blocks are not the only components that contribute to the diffuse stellar halo. In particular, one more component was discovered soon after the GSE remnants, namely: Sequoia (Myeong et al. 2019, see also Matsuno et al. 2019 and Monty et al. 2020), which is characterised by a weak retrograde rotation while the GSE's remnant stars are on radial orbits. The sequoia remnant was discovered by using the globular clusters (GCs) chemical abundances and kinematics and the same data sets were used to predict the presence of several other merger debris (see, e.g., Kruijssen et al. 2019; Forbes 2020). Although the presence of many other potential merger debris has been identified, the number of stars that can be associated with them is quite limited. In this context, it is worth mentioning Wukong (Naidu et al. 2020), Thamnos (Helmi et al. 2017, but see also Koppelman et al. 2019), I'itoi, and Arjuna (Naidu et al. 2020); however, the estimated masses of their parental galaxies are by a factor of ≈ 10 smaller compared to the GSE progenitor (see Table 1 in Naidu et al. 2022, and references therein).

The above-mentioned substructures have mainly been identified by using combinations of the integrals of motions, orbital parameters and some chemical abundances. For instance, Wukong appears as a pair of overdensities in the $E - L_z$ coordi-

nates with a cutoff at low [Fe/H] (Naidu et al. 2020). At the same time, Wukong is not a prominent group of stars in the orbital parameters and [α/Fe]–[Fe/H] spaces. Thamnos's stars were selected using several circularity cuts in the retrograde region of the $E - L_z$ space where the debris is presumably made of two components with different V_ϕ (Koppelman et al. 2019). We note that, in contrast, for instance, to Wukong, there are no GCs found in the Thamnos selection regions. In the [α/Fe]–[Fe/H] space, while having slightly larger [Mg/Fe], Thamnos substantially overlaps with the GSE and other substructures. Arjuna and I'itoi are a pair of substructures with the net retrograde rotation, again selected with sharp boundaries in $E - L_z$ coordinates, together with the Sequoia result in three peaks in the metallicity distribution function (MDF; Naidu et al. 2020). Similarly to the previously reported structures, Arjuna and I'itoi, overlap in the [α/Fe]–[Fe/H] space and exhibit a broad range of orbital eccentricities $\approx 0.1-0.7$.

The identification of prograde structures in the halo is even more difficult compared to the retrograde ones because of a strong contamination from the heated MW disc. Nevertheless, the list of potential mergers debris includes Helmi stream (Helmi et al. 1999), Aleph (Naidu et al. 2020), Nyx (Necib et al. 2020), Cetus (Newberg et al. 2009), and Orphan/Chenab (Koposov et al. 2019; Li et al. 2021). For instance, the Aleph substructure was identified as a sequence below the high-[α/Fe] sequence, or as a continuation of the low-[α/Fe] sequence ([Fe/H] ≈ -0.5), with a very low-eccentricity, typical for the disk stars. In the $E - L_z$ space, Aleph also looks like a continuation of the prograde disk populations. For a more comprehensive complete compilation of the halo substructures, we refer to recent works by Naidu et al. (2020), Malhan et al. (2022), and Mateu (2023).

In practice, a group of stars selected in the kinematical (chemical) space results in a broad distribution in chemical (kinematical) coordinates. This problem was recently investigated, for instance, by Buder et al. (2022), who showed how different chemo-kinematic selections affect the resulting sample of potentially accreted stars in other coordinates (see also Feuillet et al. 2021; Horta et al. 2023). Another concern regarding the origin of the substructures identified in the MW halo, is that often a new group of stars is being 'discovered' because some of GCs (sometimes one or two) appearing in the same region of the $E - L_z$. However, it is not evident that the GCs and the field stars of a dwarf galaxy, once it is undergoing accretion, experience the same orbital evolution (Sanders & Binney 2013; Pagnini et al. 2023), at least because of the dynamical friction affecting the motions of the globular clusters but with no impact on a low-density field stars (Pagnini et al. 2023).

As mentioned above, a large number of substructures have been discovered in the halo of the MW and more features in the halo phase-space will be identified thanks to the publication of forthcoming *Gaia* data releases. Nowadays, there is a tendency to associate each halo substructure with an accretion of a dwarf galaxy. However, the link between a given substructure and the accreted dwarf galaxy is not so trivial. In particular, a set of difficulties of the halo substructures identification and their link to ancient merger events raises from the results of various types of simulations (cosmological and tailored mergers of galaxies) which clearly show that: i) a single merger debris span over the large range in $E - L_z$ space (Gómez et al. 2013; Jean-Baptiste et al. 2017; Grand et al. 2019; Simpson et al. 2019); ii) a single merger could result in a number of prominent overdensities in $E - L_z$ coordinates and other kinematic spaces (Jean-Baptiste et al. 2017;

Grand et al. 2019; Koppelman et al. 2020; Khoperskov et al. 2023a); iii) finally, both E and L_z are not conserved quantities in case of evolving and non-axisymmetric galaxy, thus the stars accreted >8−10 Gyr can change their orbital characteristics (including actions) due to substantial mass growth of the galaxy (Panithanpaisal et al. 2021; Khoperskov et al. 2023a) and/or close passages of massive satellites – for instance, the LMC/SMC system (Erkal et al. 2021; Conroy et al. 2021) or Sgr dwarf galaxy (Laporte et al. 2018) in the case of the MW.

In a series of works based on a new set of the HESTIA[1] high-resolution cosmological simulations of Local Group (LG) galaxies, we investigate the impact of the ancient mergers on the in situ stellar populations (Khoperskov et al. 2023b, hereafter Paper I) and phase-space evolution as well as the present-day structure of the merger debris (Khoperskov et al. 2023a, hereafter Paper II). In this paper, we study the chemical abundance patterns as a function of the stellar ages and kinematics of both accreted and in situ stellar populations formed in six M31/MW analogues from the HESTIA simulations. The paper is structured as follows. The HESTIA simulations are briefly described in Sect. 2. In Sect. 3 we analyse the relations between the mean stellar metallicity of the stellar debris and the mass and accretion time. We also compare the chemical composition of the merger debris and surviving satellites. In Sect. 4, we focus on the metallicity variations of the individual debris in the $E - L_z$ space. In Sect. 5, we analyse the properties of accreted and in situ stellar populations in the [Mg/Fe]–[Fe/H] space as a function of stellar kinematics. In Sect. 6, we present the chemical abundances variations for a number of elements. Finally, in Sect. 7, we summarise our main results.

2. Model

2.1. HESTIA simulations

In this study, we analyse the three highest resolution HESTIA simulations of the LG. Each simulation reproduces many LG properties, which have previously been described by Libeskind et al. (2020). These include two massive disc galaxies resembling the MW and M31 analogues, as well as the population of orbiting smaller satellites at $z = 0$.

The HESTIA simulations use the AREPO code (Springel 2005; Pakmor et al. 2016), with gravitational forces computed using a hybrid TreePM technique, described in Springel (2005). The galaxy formation model used is from Grand et al. (2017). HESTIA's cosmology is consistent with the best fit values (Planck Collaboration XVI 2014): σ_8 = 0.83 and H0 = 100 h km s^{-1}Mpc^{-1}, where h = 0.677. Moreover, Ω_Λ = 0.682, Ω_M = 0.270, and Ω_b = 0.048. We identified halos and subhaloes at each redshift by using the publicly available AHF[2] halo finder (Knollmann & Knebe 2009). We refer to the HESTIA simulations introductory paper (Libeskind et al. 2020) and Paper I for more details.

Throughout the paper, 'in situ' refers to stars formed in the most massive M31 or MW galaxy progenitor, while 'accreted' stars are formed in other galaxies and subsequently accreted onto the most massive one (see e.g., Fattahi et al. 2020; Agertz et al. 2021). We define a merger event as the accretion of a dwarf galaxy that becomes gravitationally unbound from its own halo and gets bound to the main progenitor, according to the MergerTree tool from AHF. Therefore, all particles associated with a sub-halo right before this event are marked as accreted from

a single merger and the last snapshot is used as the time of the merger. Smaller systems that are still bound at $z = 0$, while remaining inside the virial radius of the main progenitor, represent the population of surviving dwarf galaxies of M31/MW analogues.

2.2. Orbital parameters calculation

For each M31/MW simulated galaxy and every snapshot output, we defined a coordinate system (x, y, z) centred at the position of the 10% most bound in situ star particles and aligned with the principal axes of this in situ stellar component, such as the disk plane of the host galaxy located in the XY-plane and the z-axis along the direction of galactic rotation. Our study uses velocity in galactocentric cylindrical coordinates, tangential velocity, V_ϕ, radial velocity, V_r, and velocity in the z direction, V_z. We also made use of the integrals of motion, focusing on the angular momentum in the z direction L_z, and the total orbital energy per mass E. To calculate the orbital eccentricity of star particles, we did not trace the motion of the particles across the snapshots, but instead integrated the orbits of star particles in a smooth potential. This allowed us to obtain the instantaneous values of the orbital parameters for each snapshot independently, which is not possible directly from the output data.

To compute the instantaneous orbital parameters of star particles, we used AGAMA (Vasiliev 2019) to model the smooth potential. In order to avoid the perturbation of orbits from massive satellites, we interpolated the galaxy potential using the particles associated with the host galaxy only. The potential due to dark matter and halo gas was represented by a symmetric expansion in spherical harmonics up to $l = 4$, while the potential of stars and the gaseous disk was approximated by an azimuthal harmonic expansion up to $m = 4$. Next, we integrated the orbits of star particles in this potential for 20 Gyr. This long timescale was chosen to account for the many periods expected for halo particles with small orbital frequencies. The orbits were integrated using the eighth-order Runge-Kutta DOP853 integrator with an adaptive time step from AGAMA (Vasiliev 2019).

2.3. HESTIA: Merger histories

The merger history of the HESTIA galaxies is explored in detail in Khoperskov et al. (2023a,b; see also Dupuy et al. 2022 for an analysis of direction of accretion of satellites). The total number of mergers in M31/MW HESTIA galaxies varies within the range of ≈10−40 for different galaxies; however, the number of massive mergers with $\mu_* > 0.2$ (i.e., at least 1:5 mergers) is only 1−4 (see Fig. 4 in Khoperskov et al. 2023b). For the analysis of individual mergers in this work, we selected the five most significant ones per galaxy, namely the mergers with the largest stellar mass relative to the host galaxy at the time of the merger. We refer to these mergers as 'significant' because not all of them can be classified as 'major' mergers. For most of the galaxies, the last significant merger has taken place >8 Gyr ago. This expectation is similar to what is considered for the MW (see Helmi 2020, for review and references in the introduction), with the single exception of the M31 analogue in the 09−18 simulation. In the analysis, we marked the five most significant mergers in each M31/MW galaxy as M1−M5, that is, from the earliest to the latest one. Their stellar masses lie in the range $5 \times 10^8 - 2 \times 10^{10}\,M_\odot$, which corresponds to 0.05−0.96 stellar mass relative to the host at the time of the merger. We note that some of these mergers are rather massive relative to our expectations regarding the merger history of the MW (see e.g., Deason et al. 2019;
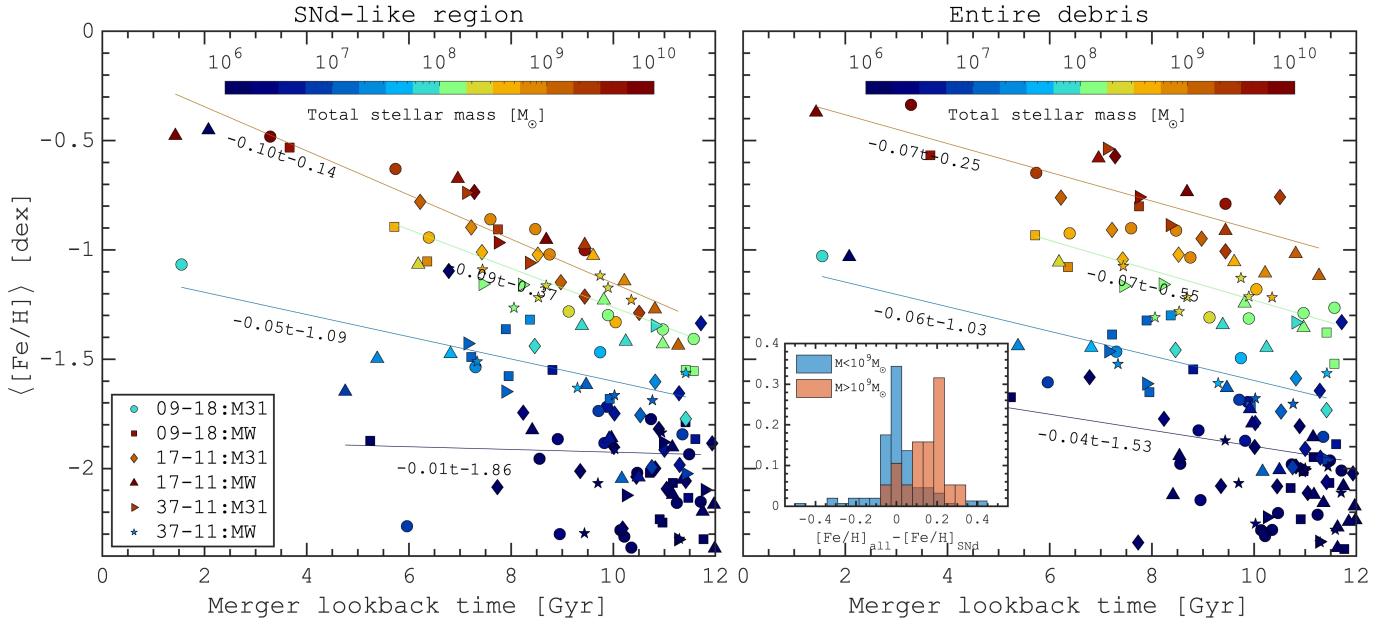
Fig. 1. Relation between mean metallicity and merger time for all the merger debris in the six HESTIA galaxies (indicated by different symbols in the bottom left of the left panel). The symbols are colour-coded according to the debris stellar mass. The left and right panels show the SNd-like and the entire population, respectively. In the right panel, the subplot depicts the difference between the mean metallicity of the entire debris and that of the SNd-like region $((0.5-2)R_d$ and $|z| < 10$ kpc, where R_d is the disk scale length from Libeskind et al. 2020). Two distributions are shown for the low-mass debris $(<10^9 \, M_\odot$, blue) and the high-mass debris $(>10^9 \, M_\odot$, red). The lines of different colours show the linear fits of the mean metallicity-and-merger time relation in several bins of the stellar mass.

Kruijssen et al. 2020); however, since the HESTIA simulations cover pairs of the M31/MW analogues, they are in agreement with the merger history of the M31 galaxy (D'Souza & Bell 2018).

3. Age-mass-metallicity relations for mergers debris and satellite galaxies

Chemical abundances of the merger debris are being used not only to identify substructures in the halo but also to recover the star formation history and the total mass of the parental dwarf galaxy (see e.g., Helmi et al. 2018; Vincenzo et al. 2019; Fernández-Alvar et al. 2019; Mackereth et al. 2019; Feuillet et al. 2021; Sanders et al. 2021). These properties can be compared to the population of surviving dwarf galaxies orbiting around the MW. In some sense, both surviving satellites and those that have merged at different epochs are expected to experience similar evolutionary paths and to exhibit similarities in their chemical composition. However, in the MW, the elemental abundance measurements of stars in local dwarf galaxies show that these systems are chemically distinct from halo stars and often from other low-mass systems (Shetrone et al. 2003; Tolstoy et al. 2009; Hasselquist et al. 2017; Fattahi et al. 2020; Nidever et al. 2020; Reichert et al. 2020). The observed differences suggest that the life paths of these systems are quite different; the evolution of merged galaxies stops once they merge and get buried in the host galaxy, while the surviving dwarfs experience a much longer evolution and enrichment history – unless they have quenched their star formation.

3.1. Metallicity-accretion time relation

Before comparing the merger debris with the sample of dwarf galaxies in the HESTIA simulations, we look at the relation between the mean stellar metallicity of the merger debris as

a function of the merger lookback time. In Fig. 1 we show the compilation of all the mergers in six M31/MW HESTIA galaxies colour-coded according to the total stellar mass of the debris. In the figure, we compare the relations based on the mean metallicity calculated for the solar neighbourhood (SNd)-like region (left)[3] and the entire debris (right). First we can see a wide range of metallicities of the debris, from [Fe/H] ≈ -2.5 to [Fe/H] ≈ -0.5, where, as expected, lower-mass systems have on average lower mean metallicities. Also we find a clear trend where more massive debris have higher metallicity at the same accretion time and merger debris of similar masses have higher metallicity if they accreted later. These relations are somewhat expected because in order to form a more massive dwarf at the same time the SFR should be higher and thus, the chemical evolution needs to be faster and the metallicity reaches higher values on a shorter timescale. An interesting implication of the relations we find is that we can fit the mean metallicity as a function of the accretion time with a simple linear function in different mass bins (see numbers in the figure). The fits we provide can be used to constrain the time of accretion of the merger debris. For instance, if we use the mass of the GSE of $10^{8.7} \, M_\odot$ and the mean metallicity of -1.18 (Naidu et al. 2022) then, following the HESTIA galaxies relation, we can estimate the accretion time of ≈ 11 Gyr, which agrees well with other estimates (Helmi et al. 2018; Di Matteo et al. 2019; Gallart et al. 2019). As shown in Fig. 1, the relation between the accretion time and mean metallicity of stellar debris, estimated in a SNd-like region, is slightly different from the one based on the entire debris. In this case, however, our estimate of ≈ 9 Gyr is still consistent with the lower limit of GSE accretion time estimates (Mackereth et al. 2019; Kruijssen et al. 2020; Lu et al. 2022).

---

[3] We define the SNd-like region as a circular annulus $(0.5-2)R_d$ within $|z| < 10$ kpc where R_d is the radial scale length of the disk from Libeskind et al. (2020).
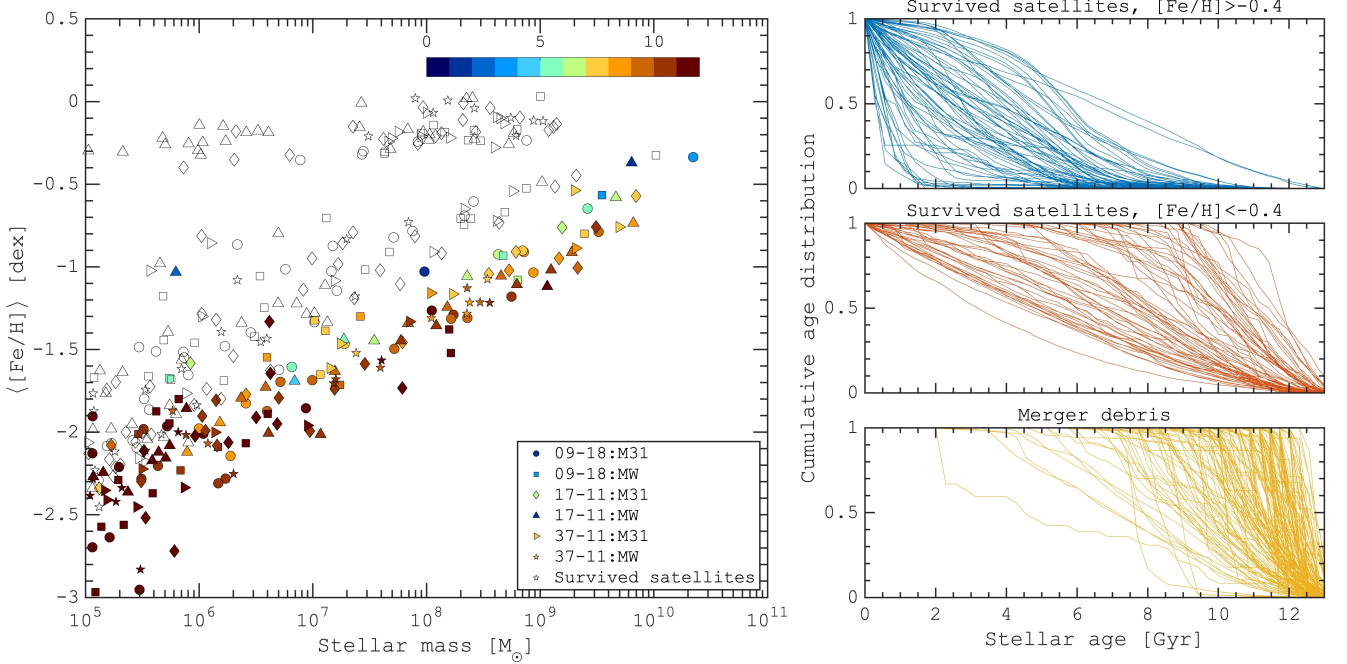
Fig. 2. Relation between stellar mass and mean stellar metallicity for all merged satellites (coloured symbols) and surviving satellites (empty symbols) for the six HESTIA M31/MW galaxies shown in the left panel. As in Fig. 1, different galaxies are shown by different symbols, yet here they are colour-coded according to the merger lookback time, as indicated in the colour bar. The right panel shows the mass-weighted cumulative stellar age distribution for metal-rich satellites (top), metal-poor satellites (middle), and merger debris (bottom).

Another issue we address in Fig. 1 is the difference between the mean metallicity measured across the entire debris and in a local SNd-like region. As we show in the subpanel on the right, the difference usually does not exceed 0.2 dex, however, it is slightly biased towards larger metallicities for the entire debris compared to the locally measured values. Although the difference we find is not significant for most of the mergers remnants, already at this point, we can see that the metallicity is not uniform across the debris, which we investigate in the subsequent section (Sect. 4). Nevertheless, the [Fe/H] difference found for the most massive debris ($>10^9\,M_\odot$) can reach up to ≈ 0.2 dex. This mass range is close to the estimated mass of the GSE, and since its mass is based on the data biased towards the SNd, one could expect that the mean metallicity of the GSE debris is higher and, thus, its total stellar mass can be underestimated. Consequently, the GSE could have been accreted at a smaller redshift. Therefore, we suggest that in order to constrain better the GSE merger parameters, it is vital to analyse the abundances of accreted stars in the central parts of the MW (Arentsen et al. 2020; Queiroz et al. 2021), which will be soon accessible for the MOONS (Gonzalez et al. 2020) and 4MOST (de Jong et al. 2019) spectroscopic surveys.

3.2. Mass-metallicity relation: Merger debris and surviving dwarf galaxies

Dependence of the stellar metallicity on the mass is a fundamental relation reflecting the efficiency of the star formation and the physical conditions in galaxies (Kauffmann et al. 2003; Tremonti et al. 2004; Tolstoy et al. 2009; Kirby et al. 2013). In Fig. 2, we show the relation between the mean metallicity of stars and the total stellar mass of merger debris (coloured symbols) and surviving at $z = 0$ dwarf galaxies (empty symbols), where the symbols correspond to the M31/MW HESTIA galaxies in three simulations (as noted in the legend). Once we com-

bined the information about all the merger debris from all the galaxies in our sample, we could see a very narrow relation between their stellar mass and the mean metallicity. However, all the merger debris set up a 'metallicity floor' at a given stellar mass. A striking feature that can be seen in Fig. 2 is a split of the relation for the surviving satellites, where the most metal-rich dwarfs seen as a tight sequence with the metallicities from ≈ -0.4 to 0. The split of the relation is explained by the star formation history, which we show in the right panels of the figure and where we calculate a cumulative mass fraction as a function of the stellar age. As we can see, the metal-rich satellites (top right) have extended star formation histories with a rather high intensity in a recent epoch (<5 Gyr). The star formation history for the low-metallicity satellites is qualitatively different because, in most cases, the active phase of the star formation stops at early times (>5 Gyr) or the star formation is quiescent on a very long time scale (middle right). The SFH for the merger debris is even a more extreme case, whereby the star formation shuts down early on at the time of the merger. Therefore, the differences in the star formation histories aptly explain the mass-metallicity relation for the dwarf galaxies and the merger debris. In other words, the merger debris have the SFHs similar to the dwarf galaxies that quenched the star formation early on, and, as a result, their metallicities are more similar for the same stellar mass systems.

Next, we analysed the age-metallicity relation by comparing the merger debris with surviving satellites. In Fig. 3, we show the mean stellar metallicity as a function of the stellar age, where surviving satellites are shown by the grey lines and the merger debris are colour-coded according to the lookback time of the merger. This figure clearly explains the difference between the metallicity behaviour of dwarf galaxies and the merger debris. The enrichment of the merger debris stops much earlier compared to the dwarf galaxies, especially the ones that have extended chemical evolution. Meanwhile, at early times,
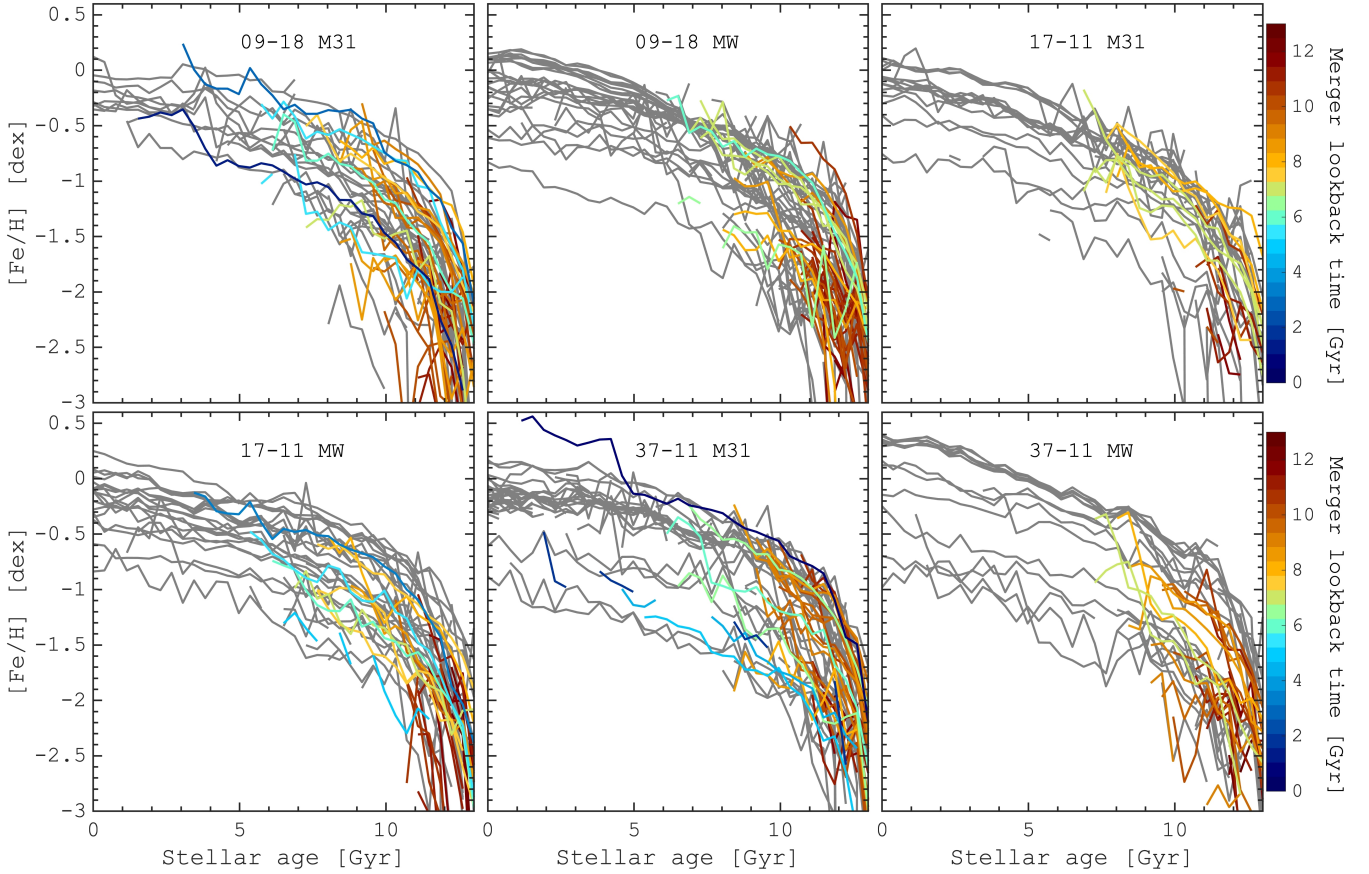
Fig. 3. Age-metallicity relation for the stellar merger debris (coloured lines) and surviving satellites (grey lines), for the six M31/MW HESTIA galaxies. The debris are colour-coded according to the merger lookback time.

the age-metallicity relations of the merger debris and surviving dwarfs are essentially the same; thus, they represent a similar type of high-redshift objects. If we turn this point around, the stellar population of dwarf galaxies with quenched star formation should represent the stellar populations of accreted systems buried in the halo. Therefore, we suggest that the analysis of old dwarf galaxies without recent star formation should offer a better idea about the chemical composition of the merger debris.

3.3. [Mg/Fe]–[Fe/H] relation for the merger debris and dwarf galaxies

It is widely known that the $[\alpha/\text{Fe}] - [\text{Fe/H}]$ relation provides a vital information about the star formation activity and the chemical enrichment variations (see Matteucci 2021, for the review). In our analysis, we used [Mg/Fe] as a proxy of the $[\alpha/\text{Fe}]$ elements. Similarly to Fig. 1, in Fig. 4, we show the [Mg/Fe]–[Fe/H] plane for the mean abundances of the merger debris in the SNd-like region (left) and the entire debris (right). In the right panel, we also add the mean abundances of the surviving dwarf satellites, shown by the empty symbols. In both panels, the colour of the symbols indicate the total stellar mass. Although there is a large scatter of the mean abundances and significant overlap between the merger debris, it is evident that the dwarf galaxies are more metal-rich and have lower $[\alpha/\text{Fe}]$ abundances. For the same stellar mass, disrupted galaxies, which are accreted early on, have little time to form their stellar content, which is likely not enough to enrich the ISM by the SNI and, thus, to reduce the $[\alpha/\text{Fe}]$-abundances of newly formed stars. On the other hand, dwarf galaxies have more time to evolve, and they

have the same masses as the merged ones if the star formation activity is quiescent on a longer times scale. The differences in the star formation activity that lead to the observed $[\alpha/\text{Fe}]$ relations are shown in Fig. 2. Our results regarding the $\alpha - [\text{Fe/H}]$ relation between the merger debris and the dwarf galaxies are in quantitative agreement with recent findings by Naidu et al. (2022), who compared the merger debris abundances delivered by the H3 Survey with the MW dwarfs.

4. Chemical abundance variations in $E - L_z$

4.1. Metallicity distribution in $E - L_z$ for accreted and in situ stars

In this section, we turn our attention from the mean chemical abundances of the merger debris to the metallicity and [Mg/Fe] variations of the accreted stars in several phase-space coordinates, which we also compare to the in situ stellar populations. As we already noted in the introduction, differences in metallicity in different regions of the $E - L_z$ space is often used to claim different dwarf galaxy progenitor for the MW halo substructures. In order to discuss the metallicity behaviour in the HESTIA galaxies similarly to the ones available in the MW, in Fig. 5 we show the $E - L_z$ space for the SNd-like selection (based on the circular annulus $(0.5-2)R_d$ within $|z| < 10$ kpc). In the top row, we show the stellar density distribution, while the following rows present the mean stellar metallicity for all stars, both the in situ and accreted stellar populations. In the top panels, we can see a number of substructures in the $E - L_z$ space; however, as we showed in Paper II, not all these features correspond
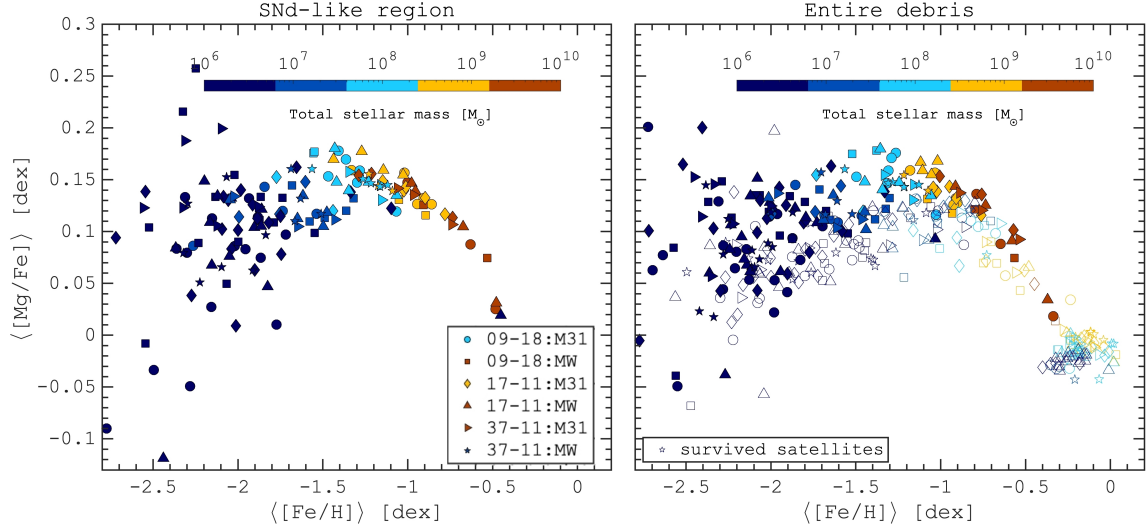
Fig. 4. Mean [Mg/Fe] abundance vs. mean metallicity, $\langle[Mg/Fe]\rangle - \langle[Fe/H]\rangle$, for the merger debris (filled symbols) and surviving satellites (empty symbols). In both panels, the symbols are colour-coded according to the total stellar mass. The left panel, similarly to Fig. 1, shows the relation for the SNd-like selection, while the right panel shows the values averaged over the entire system.
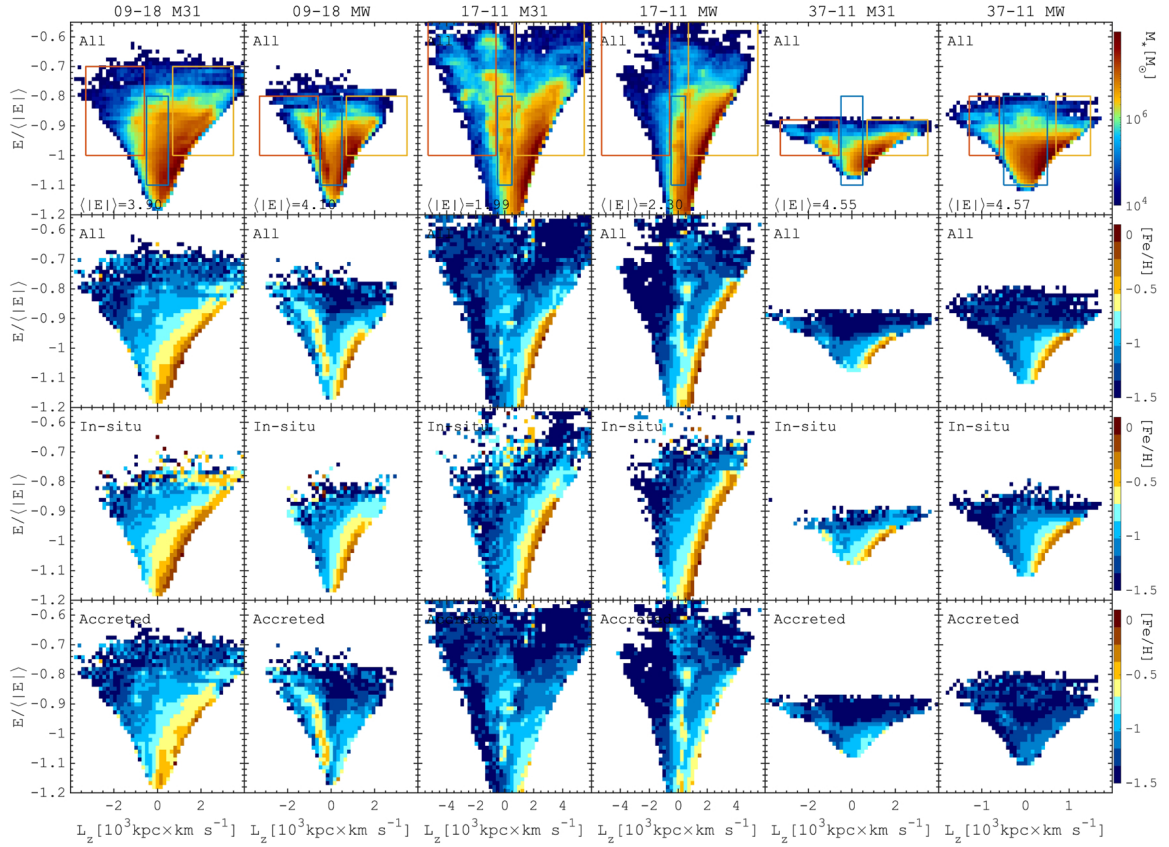

Fig. 5. Energy-angular momentum relation for the SNd-like region for all (top two rows), in situ (third row), and accreted (bottom row) populations. The first and second rows are coloured by the stellar number density and the mean stellar metallicity, respectively. In the bottom two rows, the colour corresponds to mean metallicity. The coloured boxes in the top row show the selection of retrograde (red), non-rotating (blue), and prograde (yellow) stellar populations that are featured in Fig. 6. The energy is normalised to the mean energy of the corresponding sample for a given galaxy, marked at the bottom of each top panel and in units of $10^6 \, kpc \times km^2 \, s^{-2}$.

to individual merger events (see also Jean-Baptiste et al. 2017; Grand et al. 2019). In particular, the internal structure of a dwarf galaxy, prior to the merger, can result in the formation of different substructures once the galaxy is accreted (see, e.g., Koppelman et al. 2020; Amarante & Debattista 2022). The

mean metallicity distribution is also not featureless. On top of the global [Fe/H] diagonal gradient where the metallicity increases from larger energies to the lower and from negative L_z to the positive; there are several relatively high-metallicity features. These local metallicity peaks are the most prominent near $L_z \approx 0$ in
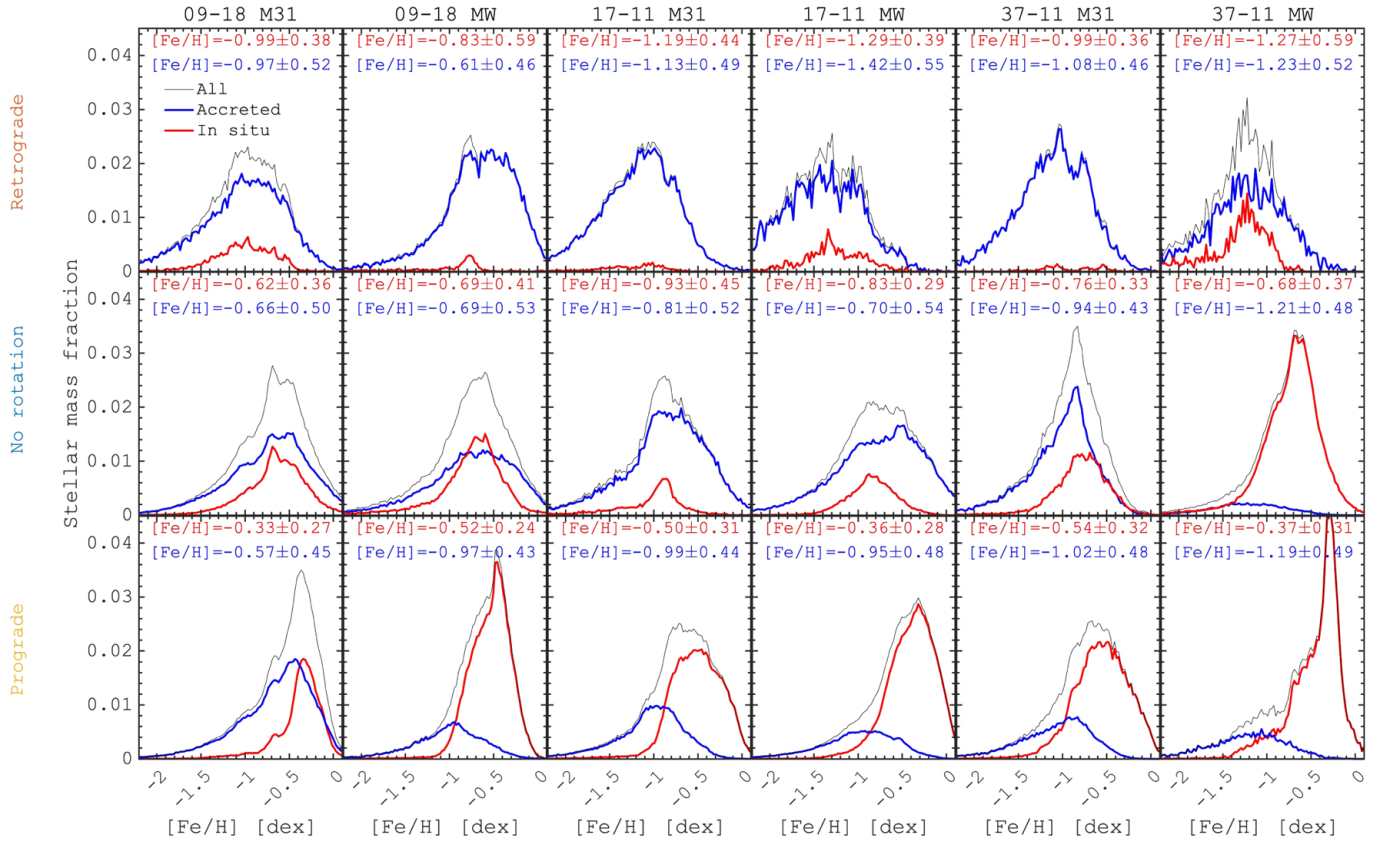
Fig. 6. Metallicty distribution functions (MDFs) for the retrograde (top), non-rotating (middle), and prograde (bottom) stars selected in the $E - L_z$ coordinates in Fig. 5. All accreted and in situ populations are shown by the blue and red lines, respectively, while the black lines correspond to the overall distribution. The MDFs are normalised by the total mass of stars, including both accreted and in situ populations. Prograde in situ MDFs always peak at higher [Fe/H] compared to prograde accreted ones, while the two are mostly degenerate for the retrograde and non-rotating stars.

both M31/MW 17-11 galaxies and in the retrograde tail of the MW analogue in the 09–18 simulation. The location of the high-[Fe/H] spots suggest that they are linked to the accretion events.

Two bottom rows in Fig. 5 show the metallicity distributions for the in situ and accreted stars separately. It is clear that in most of the galaxies, the larger-scale diagonal gradient in $E - L_z$ (see the second row in the figure) is mainly caused by the in situ stellar populations where the metallicity increases from the retrograde to prograde tails of the overall distribution. At the same time, the accreted stars display a diverse array of behaviours with regard to the metallicity distribution, where higher metallicity traces the location of the most massive, and, thus, more metal-rich debris. This debris is the most prominent on top of the overall distribution, especially if the accreted stars are not aligned with the host galaxy rotation. If the merger debris is prograde, then it is also difficult to distinguish the accreted stars from the in situ ones because the fraction of accreted stars in the prograde region is way smaller compared to the in situ component (see more details in Paper II).

For a deeper understanding of the [Fe/H] behaviour in $E - L_z$ space, we analysed the metallicity distribution function (MDF) of stars selected in three regions: retrograde, non-rotating, and prograde, which are marked by the red, blue and yellow rectangles, respectively, in the top row of Fig. 5. In Fig. 6, we present the [Fe/H] distributions separating accreted (blue lines) and in situ (red lines) stars where the mean [Fe/H] values are also marked in each panel. Interestingly, the MDFs are very similar in terms of the mean values in the retrograde part of the $E - L_z$; however, the accreted stars dominate in this

region. The mean [Fe/H] is also very similar for the non-rotating stars selection (second row), where (with a single exception: MW-37-11) the number of accreted stars is larger than the in situ ones. In this region, the accreted stars have a broader MDF. Therefore, the substantial overlap makes it very hard to disentangle between the accreted and in situ stars by using the metallicity information, only because here we show the stars with similar energy and angular momentum. The prograde region is the only one where we detected a prominent difference between the MDFs of accreted and in situ stars. Obviously, the in situ stars have a narrower distribution with a higher mean metallicity. Such behaviour is clear because at the time of accretion, all the merging dwarf galaxies are smaller than the main progenitor and, thus, they have lower metallicity (see below for more details). Therefore, the HESTIA simulations, similarly to the case of the MW (see e.g., Helmi et al. 2018; Di Matteo et al. 2019), suggest that the in situ stars which are heated up (that is why they are in the same $E - L_z$ region) by the mergers are metal-rich compared to the merger debris.

4.2. Origin of the [Fe/H] variations in the $E - L_z$ plane for accreted stars and individual mergers

In Fig. 6 we can notice several peaks in the MDFs of distribution of the accreted stars, seen mainly in the retrograde part of the $E - L_z$ space. We recall that the presence of peaks in the MDFs is often used as an argument in favour of multiple distinct substructures in the MW halo; in other words, it is assumed that different peaks of the MDFs are associated with different
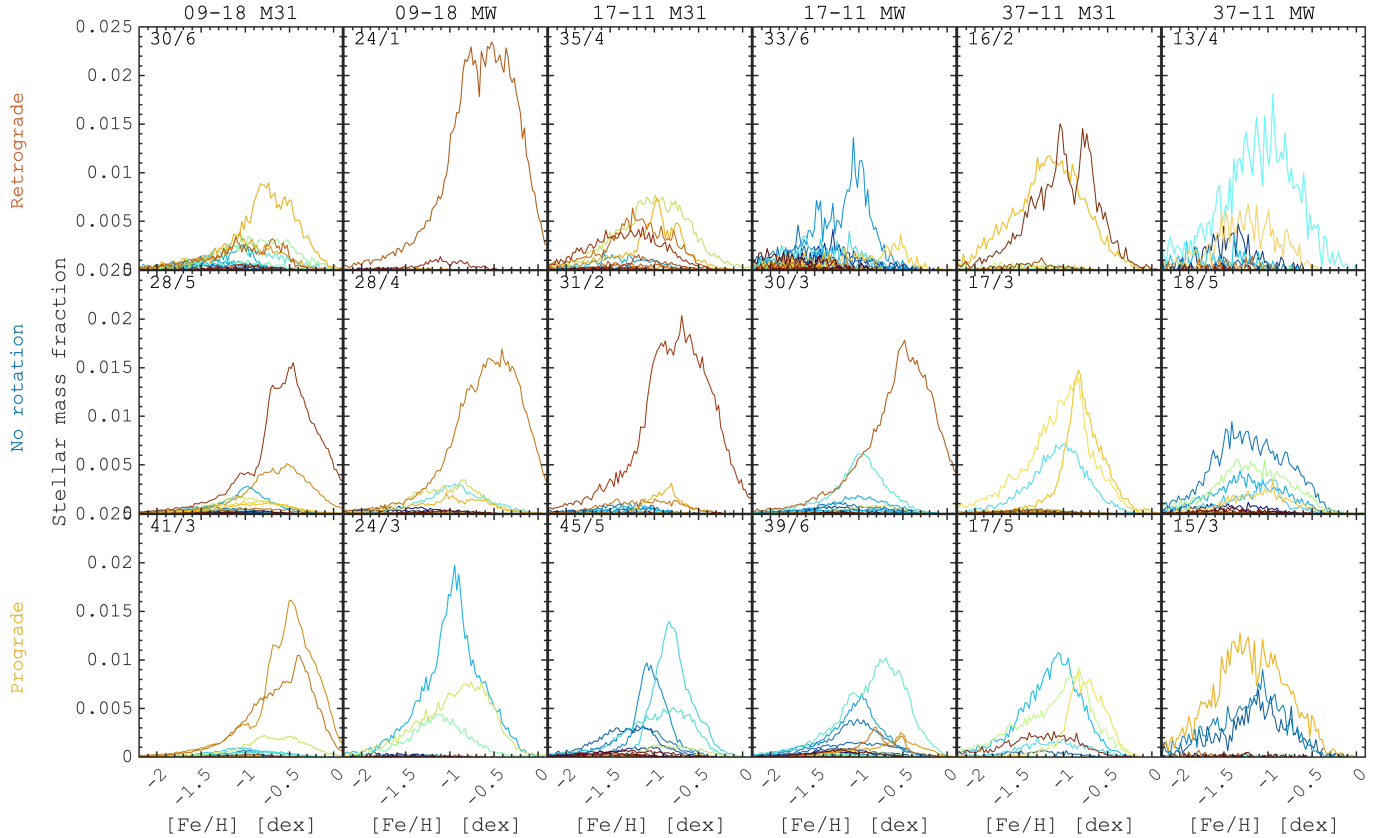
Fig. 7. MDFs for the retrograde (top), non-rotating (middle), and prograde (bottom) stars but showing individual mergers in lines of different colours. The numbers in the top-left corner of each panel indicate the number of all mergers and the number of mergers, which contribute at least 5% of the total mass of the accreted stellar populations in a given $E - L_z$ region. The MDFs are normalised by the total mass of accreted stellar populations. This figure shows that both retrograde and non-rotating accreted stars in the SNd-like region predominantly belong to just a few merger events.

progenitors (see e.g., Naidu et al. 2020). The relevance of such assumption is easy to check in the HESTIA simulations by comparing the MDFs of the individual merger debris in the same $E - L_z$ regions. In Fig. 7, we show the contribution of all the merger debris to the MDFs of the accreted stars. The individual MDFs show that both retrograde and non-rotating accreted stars of the SNd-like region predominantly belong to just a few merger events. Meanwhile, in the prograde region (bottom row), the number of significant mergers is way larger. This picture depends on the metallicity values, where at lower metallicities, the number of mergers increases, but their total mass fraction is rather low. Nevertheless, the MDFs are mainly shaped by just a few merger remnants, which, however, often show a complex behaviour. In some cases, for instance, M31 in 37-11, M31/MW in 09–18 simulations, the MDFs of the most significant mergers have a few prominent peaks, suggesting that some of the halo substructures in the MW may originate from a single progenitor, despite the fact that they have different metallicities and even linked to the peaks of the MDF.

The complex metallicity behaviour of the individual merger debris is also essential to explore in the $E - L_z$ coordinates. In Fig. 8, we show the mean [Fe/H] maps for five of the most significant merger debris for each M31/MW HESTIA galaxy (see Fig. 4 in Paper I) by selecting stars in the radial range $(0.5-2)R_d$ and within $|z| < 10$ kpc. The merger debris are sorted from the top to the bottom by decreasing the lookback time of accretion; in other words, the mergers accreted earlier are shown on the top. Before we discuss the distribution of the metallicity in the

individual merger debris, we discuss a few general trends we observe in Fig. 8. In particular, the mean metallicity increases with the decrease of the lookback time of the merger, which suggests that a longer evolution of the dwarf galaxies prior to the merger allows them to be enriched to higher metallicities (see Fig. 1). However, the difference in the mean metallicity for different debris is substantial, especially in the case of the M31 galaxy in the 09–18 simulation, where the mergers occur on a long time scale (left-most column). On the opposite, the difference in the metallicity of the merger debris is almost negligible in the case of the MW 37-11 galaxy (right-most column), where all the significant mergers happen during ≈2 Gyr.

The most striking feature of the distributions in Fig. 8 is the metallicity gradient, optimally seen as a function of the total energy. Similarly to the total metallicity maps shown in Fig. 5, the mean stellar metallicity increases with lower energy, which suggests that more bound parts of the debris are more metal-rich. Since the most metal-rich stars of dwarf galaxies are located close to the center (before the merger), where the star formation takes longer and possibly is more intense; thus, we can clearly see that the structure of the accreted galaxies is imprinted in the metallicity distribution of their stellar debris (see e.g., Amarante & Debattista 2022). The metallicity gradients are quite prominent in the debris of the galaxies already accreted ≈10 Gyr ago, while for some earlier events, it is simply not enough time to create a gradient inside the dwarfs before they accreted. In all other cases, the gradients in the individual debris are quite prominent. This result again suggests that
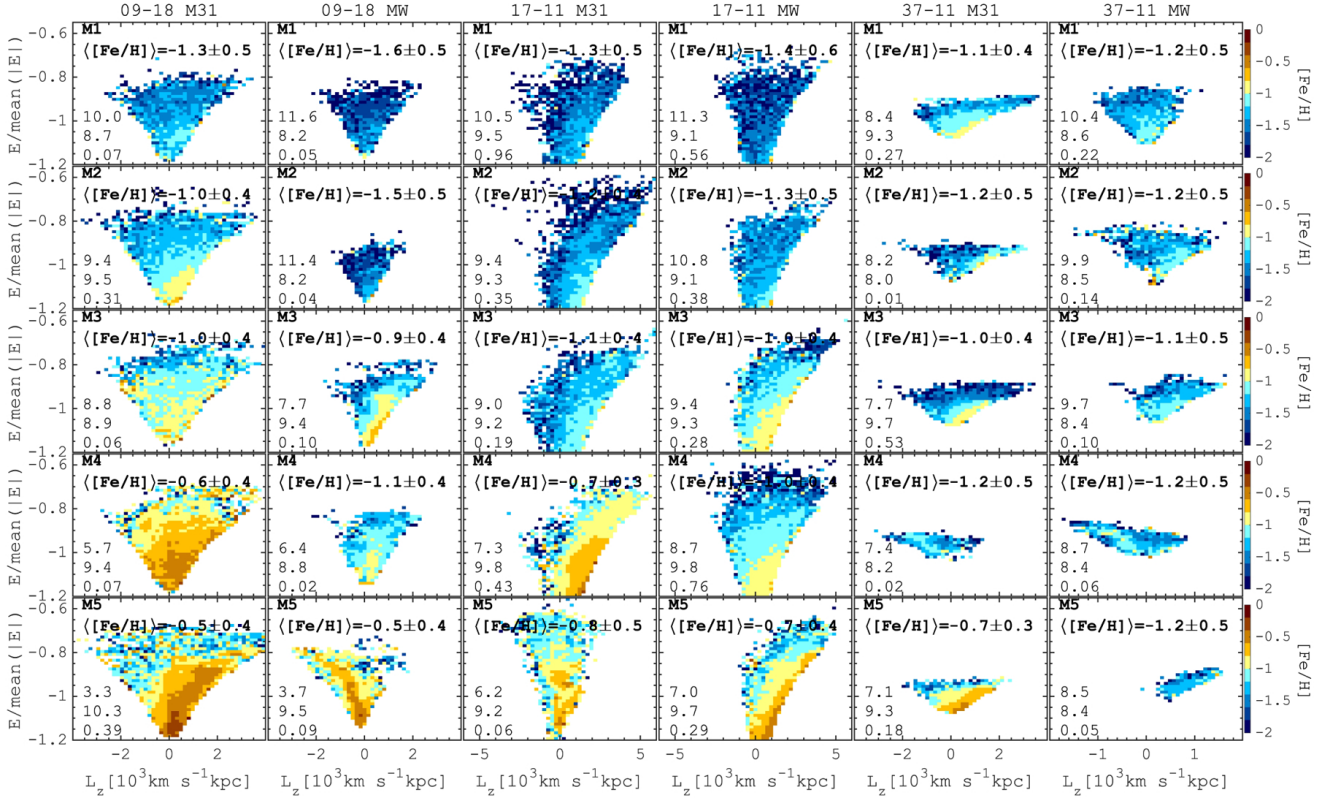
Fig. 8. Mean metallicity distribution for five of the most significant merger debris in $E - L_z$ coordinates. Different HESTA simulations are shown in different columns. The merger accretion time (Gyr), total stellar mass of the merger debris at the time of the merger ($\log_{10}(M_*/M_\odot)$), and the stellar mass ratio (μ_*) relative to the main M31/MW progenitor at the time of the merger are given in the bottom of each panel. The mean metallicity of debris associated with the given merger, together with its standard deviation, are shown in each panel. In most cases, individual merger debris show [Fe/H] gradients, where the most metal-rich stars, formed in the centers of accreting galaxies, have the lowest energies. Meanwhile, low-[Fe/H] stars formed in the outer regions of dwarf galaxies accreted first and, thus, have higher energy inside the host M31/MW galaxies.

the metallicity variations are expected in the debris of massive dwarfs accreted onto the MW, with parameters similar to the GSE system.

Next, we elaborate more on possible observational consequences that can be used to interpret the [Fe/H] variations in the $E - L_z$ space. As we mentioned already, in Fig. 8, we show the metallicity variations or gradients along with the total energy. To highlight this relation in Fig. 9 (top), we show the mean metallicity as a function of the total energy for the most significant merger debris (coloured lines) and the total ones (all the accreted stars in black). In all the HESTIA galaxies, we see a very prominent gradient of the metallicity and this gradient can be translated to the galactocentric distance because, on average, the lowest energy corresponds to the deepest parts of the potential well in the innermost parts of the host galaxy.

In Fig. 9 (bottom), we present the mean metallicity as a function of the galactocentric distance, where we also draw a vertical line at the Solar radius of 8.2 kpc. This figure demonstrates the metallicity gradients of the individual merger debris, together with the total accreted stars metallicity profiles, which show a clear decrease of the mean metallicity with radius (see also Font et al. 2011; Monachesi et al. 2016, 2019). In some cases, we observe a flattening of the negative gradient, but some of the debris have a very steep profile. This is in agreement with some previous models suggesting that the metallicity of stellar haloes is flat along radius (see e.g., De Lucia & Helmi 2008; Cooper et al. 2010; Gómez et al. 2012), however, models that include the contribution of in situ star formation reveal stronger

negative metallicity gradients (Font et al. 2011; Tissera et al. 2014). Our results suggest that once being accreted, more metal-rich parts of the disrupted galaxies are dragged towards the central parts of the host galaxy that creates the large-scale gradient (see also Figs. 6 and 7 in Paper II). This effect is more evident for the most massive accretion events. Also, the metallicity gradient suggests that the metallicity of the massive debris taken in the SNd-like region is not meant to be representative of the mean metallicity of the debris (see also Figs. 1 and 4), because the density of the debris increases towards the centre. Therefore, the measurements made in the MW should consider a possibility for the metallicity gradient of the stellar halo, even if it is dominated by a single debris (e.g., GSE). At the same time, the mean metallicity of the debris in the SNd-like region cannot be used to precisely date the merger and estimate the total stellar mass of the parental galaxy.

5. In situ and accreted populations in the [α/Fe]–[Fe/H] space

5.1. Chemical abundance space

In this section, together with [Fe/H], we include the abundances of [α/Fe] elements (in particular [Mg/Fe]) and combine this information with the kinematics of stars to describe the properties of accreted and in situ populations. In Fig. 10, we present the age-[Fe/H] and age-[Mg/Fe] relations for the in situ populations and five of the most significant merger debris. The in
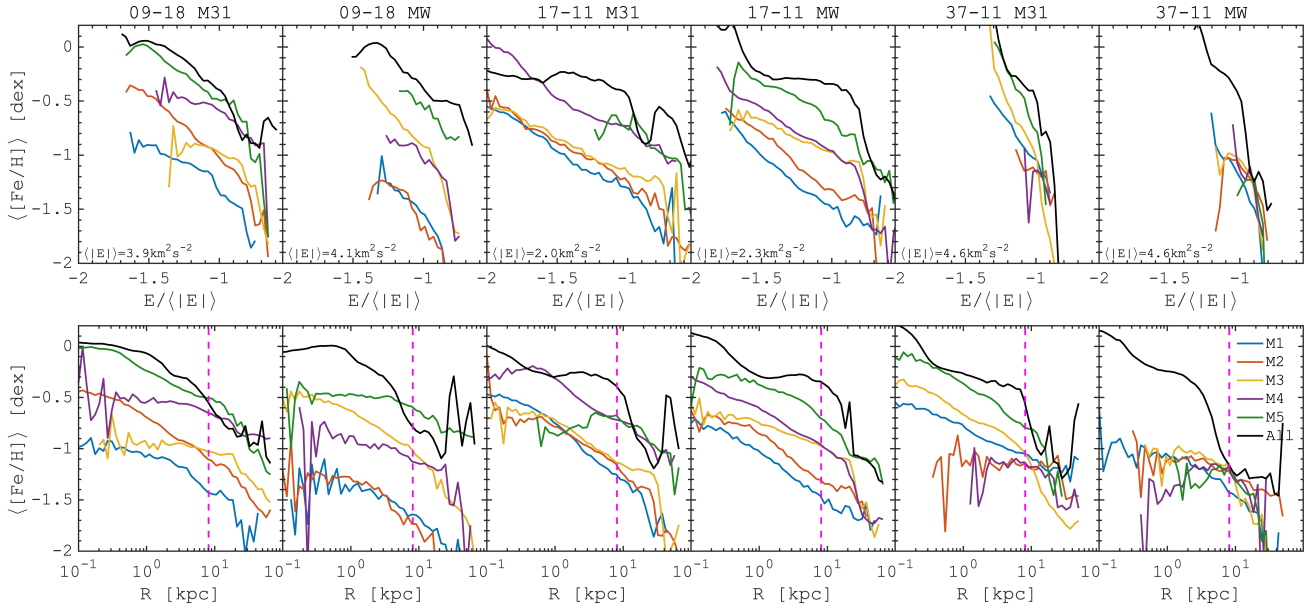
Fig. 9. Mean metallicity of the stellar debris for the five most significant mergers as a function of the total energy (top row) and galactocentric distance (bottom row). The chemical abundance variations within dwarf galaxies before the accretion result in the negative radial metallicity gradients in the main progenitors stellar halos once they are accreted.
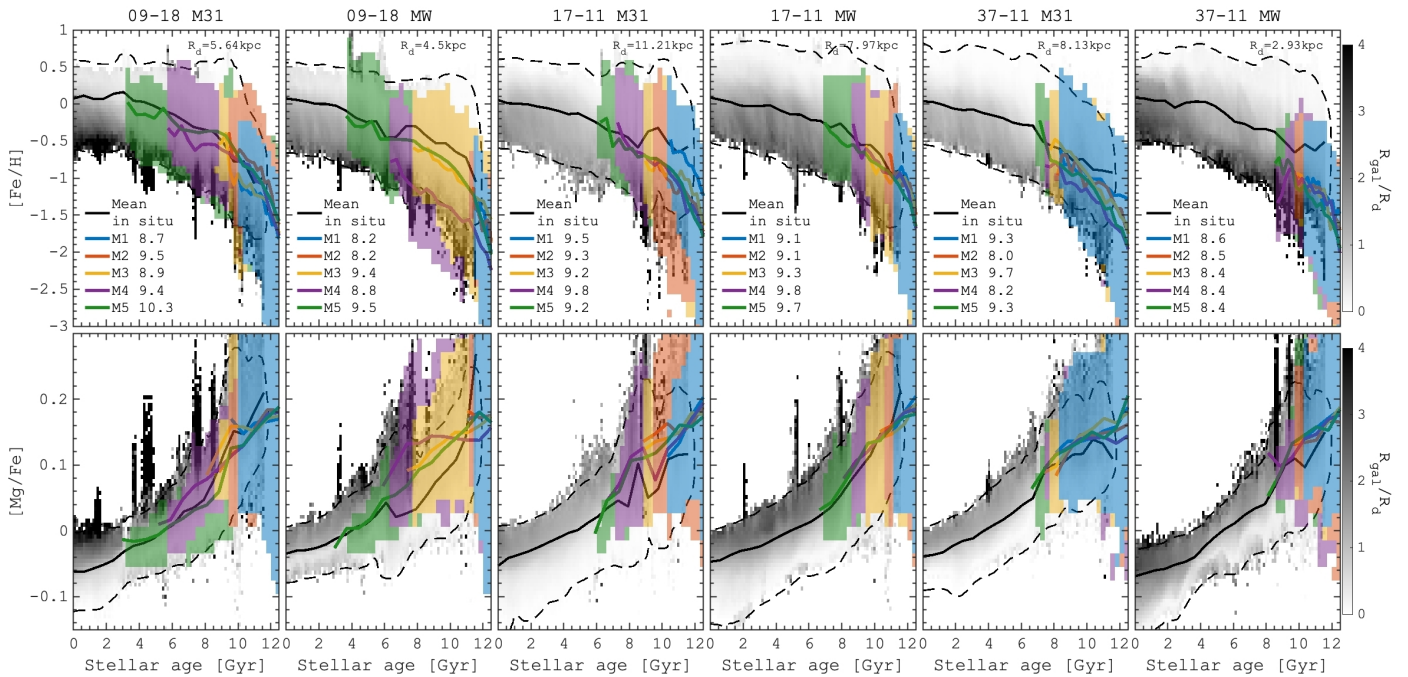

Fig. 10. Age-metallicity (top row) and age-[Mg/Fe] (bottom row) relations for in situ stars (grey shading) and the five most significant mergers (coloured areas). The total stellar mass of the merger debris at the time of the merger ($\log_{10}(M_*/M_\odot)$) is mentioned in the legend in each top panel. Mean values are shown by the coloured solid curves (except for the black one, which corresponds to the in situ stars). The grey-scale for the in situ stars represents the mean galactocentric distance in the units of the disc scale lengths from Libeskind et al. (2020), as seen in the colour bars on the right, while the black dashed curves correspond to the 3σ around the mean. For the accreted populations, there is no shading and the boundaries correspond to 3σ around the mean.

situ component is shown by the grey-scale distributions that are colour-coded by the mean galactocentric distance of stars at $z = 0$. The accreted stars are shown by the transparent filled areas where the mean age-abundance tracks are marked by the solid lines of the same colour. The yellow lines show the mean values for the in situ stars. The broad range of [Fe/H]/[Mg/Fe] at a given age represents the fact that disc galaxies exhibit a

gradient in metallicity with radius, which is often discussed in the context of inside-out galaxy formation (Larson 1976; Rupke et al. 2010; Pilkington et al. 2012; Brook et al. 2012). In the age-[Fe/H] relation, despite an overlap between the accreted and in situ stars of the same age, the accreted stars tend to have lower [Fe/H]; in this sense, it is similar to the outer disc stellar populations. The offset between the mean trends of accreted
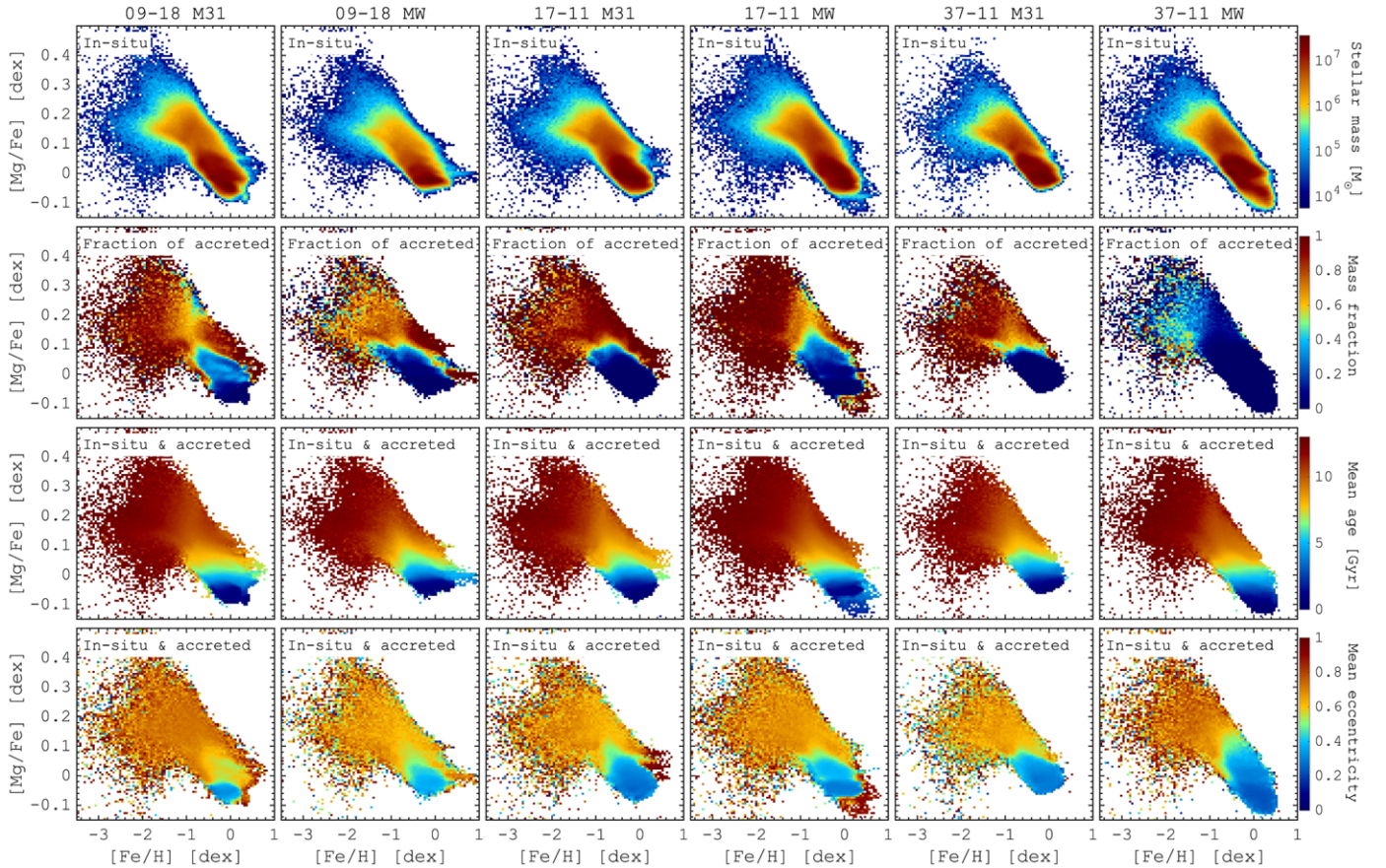
Fig. 11. [Mg/Fe]–[Fe/H] relations for the six M31/MW HESTIA galaxies (different columns), including all stars. From top to bottom: colour contours in different rows represent the stellar number density, the fraction of accreted stars, the mean stellar age, and the mean eccentricity.

and in situ stars at a given age does not exceed 0.2−0.3 dex for the low-mass debris (see Sect. 6 for more details about the offset between the accreted and in situ populations).

From Fig. 10 we can see that the maximum stellar metallicity of accreted dwarf galaxies reaches up the solar values at ≈10 Gyr ago, while at the same time, the stars in the center of the main could have even higher values. Nevertheless, it is quite intriguing that the accreted stars could have a chemical composition of the Sun. Of course, the amount of such stars is quite small, but searching for them in observations is very important because it will allow us to for better constrains to be placed on the chemical evolution, mass growth, and star formation history of accreted galaxies. One unfortunate detail, however, is the large scatter of age-[Fe/H]/[Mg/Fe] relations for the accreted populations, which also overlap with each other (see e.g., Mackereth et al. 2019; Brook et al. 2020), making it difficult to separate them when there are just a few abundances available.

In Fig. 11 we show the widely explored in the literature [Mg/Fe]–[Fe/H] relation for the simulated M31/MW HESTIA galaxies, selecting stars in the radial range $(0.5−2)R_d$ and $|z| < 10$ kpc. In the top row, we plot the stellar mass distributions, which show several substructures that are similar to the ones recently explored, for instance, in cosmological simulations (Brook et al. 2004, 2005; Grand et al. 2018; Mackereth et al. 2018). In particular, we can see a dichotomy of the [Mg/Fe] abundance at high metallicities ([Fe/H] > −0.5), similar to the MW high-[α/Fe] and low-[α/Fe] sequences. The exact mechanism of formation of the [α/Fe] bimodality in the in situ stars is beyond the scope of the present work, how-

ever, some recent results and competitive scenarios can be found in Grand et al. (2018), Clarke et al. (2019), Buck (2020), Renaud et al. (2021), Khoperskov et al. (2021). The most interesting in the context of the present study is the relation between the in situ and accreted stars.

The second row of Fig. 11 shows the fraction of accreted stars in the [Mg/Fe]–[Fe/H] plane, where the accreted stars occupy low-metallicity and high-[α/Fe] regions for all galaxies. More quantitatively, the high-[α/Fe] sequence is dominated by the accreted stars at [Mg/Fe] > 0.1, with the single exception of the MW analogue in galaxy 37-11 (rightmost panels), where the accreted component contribution is minor compared to the in situ (see also Fig. 4 in Paper I, which shows that this galaxy has a very quiet accretion history).

In some cases, however, we can see that the accreted stars can dominate in the super-solar metallicity regions, being likely formed in the cores of the dwarf galaxies, thus, they can have a longer enrichment history. This is suggested by the third row of Fig. 11, where we show that the distribution of the mean stellar age and the high-[Fe/H] knots of the accreted stars have ages comparable to the low-[α/Fe] sequence of in situ stars. More generally, regions with a higher fraction of accreted stars obviously contain older stars, and the mean stellar age decreases towards smaller [Mg/Fe] values and higher metallicity. Nevertheless, the orbital parameters of the accreted and in situ stars are very much different in the high-[Fe/H] regions, as illustrated in the bottom row of Fig. 11, colour-coded by the mean stellar eccentricity. The regions with a dominant fraction of accreted stars show higher orbital eccentricity (>0.5). At the
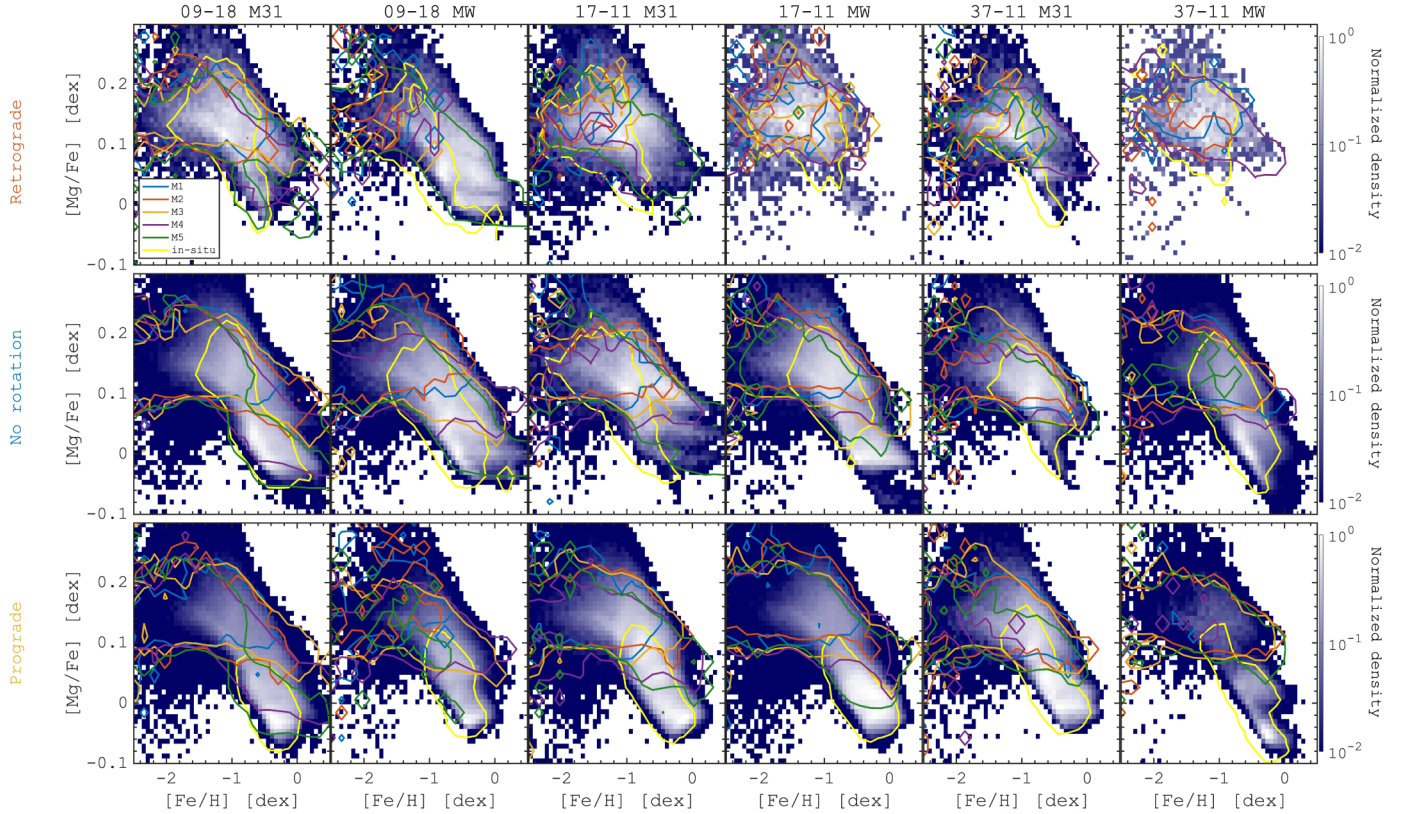
Fig. 12. [Mg/Fe]–[Fe/H] relation for retrograde (top), non-rotating (middle), and prograde (bottom) stars as selected in the blue, red, and orange rectangles in the top row of Fig. 5. The five most significant mergers are shown by colour-coded contours (75% density level), while the yellow contours correspond to the in situ stars. This figure shows that depending on the orbital properties, single merger debris can mimic chemically different substructures.

same time, the low-[α/Fe] in situ populations have disk-like kinematics[4].

As we have already shown, different merger debris overlap in the age-[Fe/H]/[Mg/Fe] relations (see Fig. 10). Thus, similar behaviour is expected in the [Mg/Fe]–[Fe/H] plane. In order to add more dimensions to purely chemical abundance relation in Fig. 12 we show the [Mg/Fe]–[Fe/H] for the stars taken in three different regions of the $E - L_z$ diagram. We also display in Fig. 5 three rectangles show the retrograde (blue), non-rotating (red), and prograde (yellow) regions. The contours in the panels of Fig. 12 show the distributions of stars from five of the most significant mergers (coloured contours), together with all accreted stars (colour map) and in situ stars (yellow line). We can see that the overall relations are somewhat different depending on the orbital parameters of stars. For instance, the metallicity peak of a given merger debris is shifting towards higher values from the retrograde to prograde stars. This also affects the shape of the relation, where for the retrograde stars, we missed a low-[α/Fe] sequence for the in situ stars suggesting that the high-[α/Fe] stars, being formed earlier, are the most contaminated by the mergers (see more details about the mergers impact on the in situ populations in Paper I). In most cases, a single merger debris have slightly different distributions in retrograde, non-rotating and prograde regions of $E - L_z$. In particular, in the retrograde region the merger debris seems not to have a prominent knee, mostly occupying the high-[α/Fe] sequence, while in the non-rotating and especially in the prograde part, the knee

is quite prominent, at least for the most massive merger debris. The explanation of such a differentiation is based on the accreting galaxy structure we already discussed above. Most likely, the retrograde part of the $E - L_z$ space is occupied by older stars formed earlier from the outer parts of the dwarf galaxies before they merged into the host. In Paper II (see Fig. 6), we showed that the retrograde parts of the $E - L_z$ are mainly made of stars accreted from the outer parts of the dwarf galaxies, which, in most cases, have a lower-[Fe/H] and do not thus experience long chemical enrichment history compared to the prograde parts of the debris, which are mainly made of the central regions of dwarf galaxies where the SNI were able to pull down the [α/Fe]-abundances to nearly solar values. Of course, there are some exceptions in the HESTIA galaxies sample, where the satellites are being accreted on the retrograde orbits (see e.g., MW in the 09–18 simulation). However, the main chemo-kinematic trends for these merger debris remain essentially the same.

5.2. Coupling chemical abundances with orbital parameters of in situ and accreted populations

In this section, we combine the orbital parameters and the kinematics of stellar populations and present their variations as a function of [Fe/H] abundances. In Fig. 13, we show how the stellar eccentricity depends on the metallicity for the in situ, accreted and individual merger debris. In these figures, we extend the eccentricity scale to the negative values by multiplying it by the sign of the angular momentum. As a result, the top parts (above the pink lines) in each panel show the relations between the stars on retrograde orbits. The in situ stellar populations (top) show

---

[4] For a detailed comparison of the orbital eccentricities of accreted and in situ stars we refer the reader to Paper II.
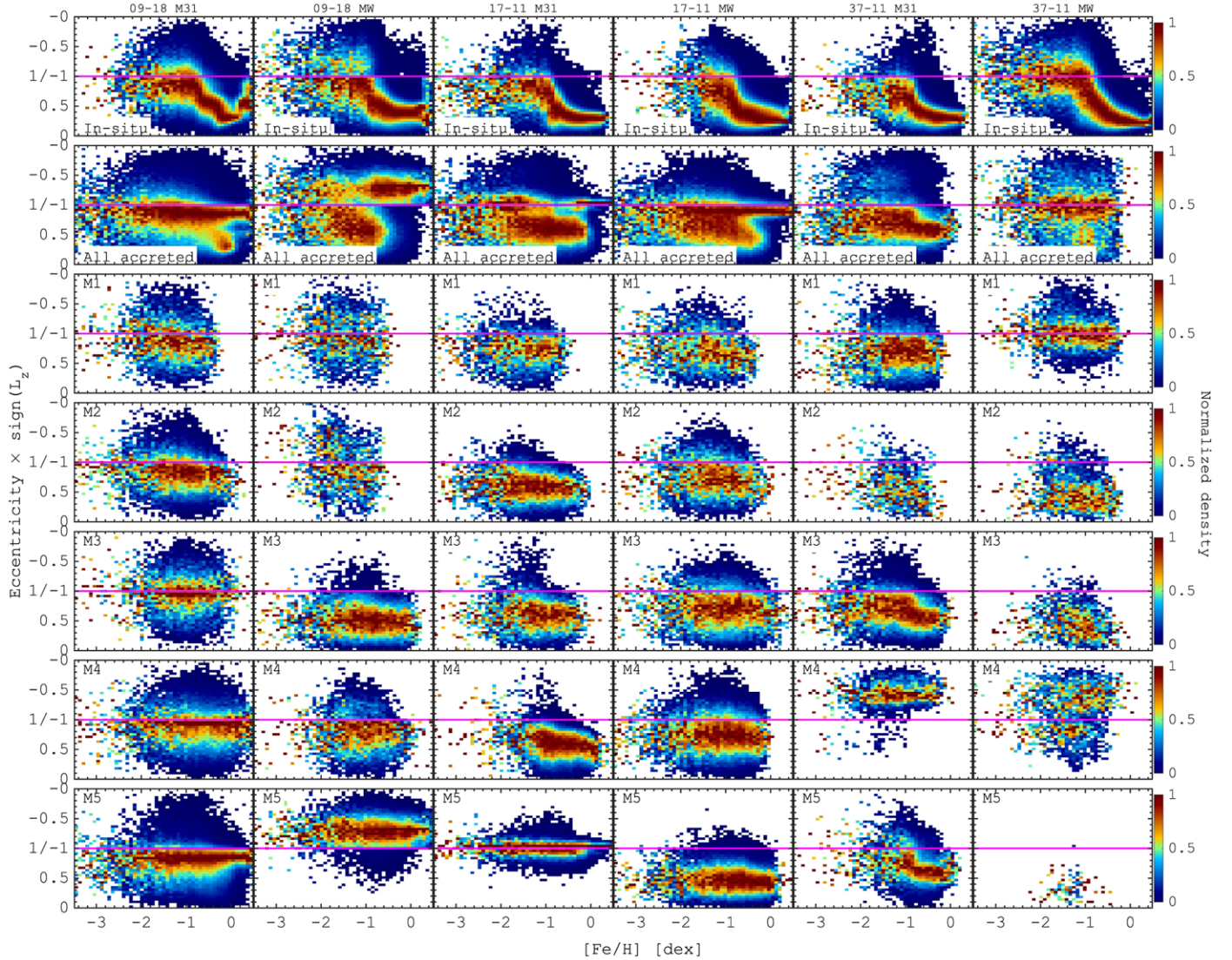
Fig. 13. Relation between orbital eccentricity and [Fe/H] for the in situ (first row), accreted (second row), and the five most significant merger debris (third through sixth rows). For improved visibility, the density maps are normalized in each metallicity bin. Stars with negative angular momentum have negative values for eccentricity. In other words, the stars above the pink line in each panel are counter-rotating. Although the eccentricity distribution of accreted stars shows a complex behaviour, the individual merger debris have nearly constant eccentricity as a function of [Fe/H]. The features of the in situ stars are linked to the merger events that heat up the disc and, thus, increase the eccentricity of the pre-existing populations.

a very complex behaviour of the eccentricity as a function of [Fe/H]. For the low-[Fe/H] stars the eccentricity is very high, reaching up to the unity with a non-negligible fraction of the retrograde stars. For the higher-[Fe/H], the eccentricity drops down to ≈0.2 in all the galaxies in our sample. Interestingly, in some cases, this transition is not smooth and we can see a few segments of nearly constant eccentricity in a range of [Fe/H]. In Paper I, we showed that these sharp changes of the eccentricity as a function of the stellar age are linked to the merger events (see Fig. 13 therein) where the oldest ones represent stellar populations with kinematics similar to the Splash/Plume discovered in the MW (Di Matteo et al. 2019; Belokurov et al. 2020).

In Fig. 13, the accreted stars show even more complex behaviour compared to the in situ populations. We can see bimodal distributions and broadening and thickening of the distributions in some ranges of abundances. In most of the abundance ranges, the eccentricity of accreted stars is higher

compared to the in situ stars, however, it is not always the case. In all the galaxies we can see the presence of rather regularly rotating accreted populations with the eccentricity down to ≈0.5. However, the complexity of the relation breaks once we decompose the accreted stars onto the individual merger debris. Starting from the third rows in Fig. 13 we present the eccentricity variations for the most significant debris. In this case, contrary to the entire sample of accreted stars, the individual debris are characterised by a nearly constant eccentricity along the [Fe/H]. Most of the debris have a broad range of eccentricities with some contribution to the retrograde regions and there are few individual merger debris that are predominantly counterrotating (e.g., M5 in MW 09–18 and M4 in M31 37-11). The results presented in Fig. 13 suggests that although the accreted stars have a complex distribution of the orbital eccentricities, the individual merger debris tend to have roughly constant eccentricity along the [Fe/H]. Therefore, once the scatter of about 0.1−0.2 of the eccentricity is taken into account, this behaviour
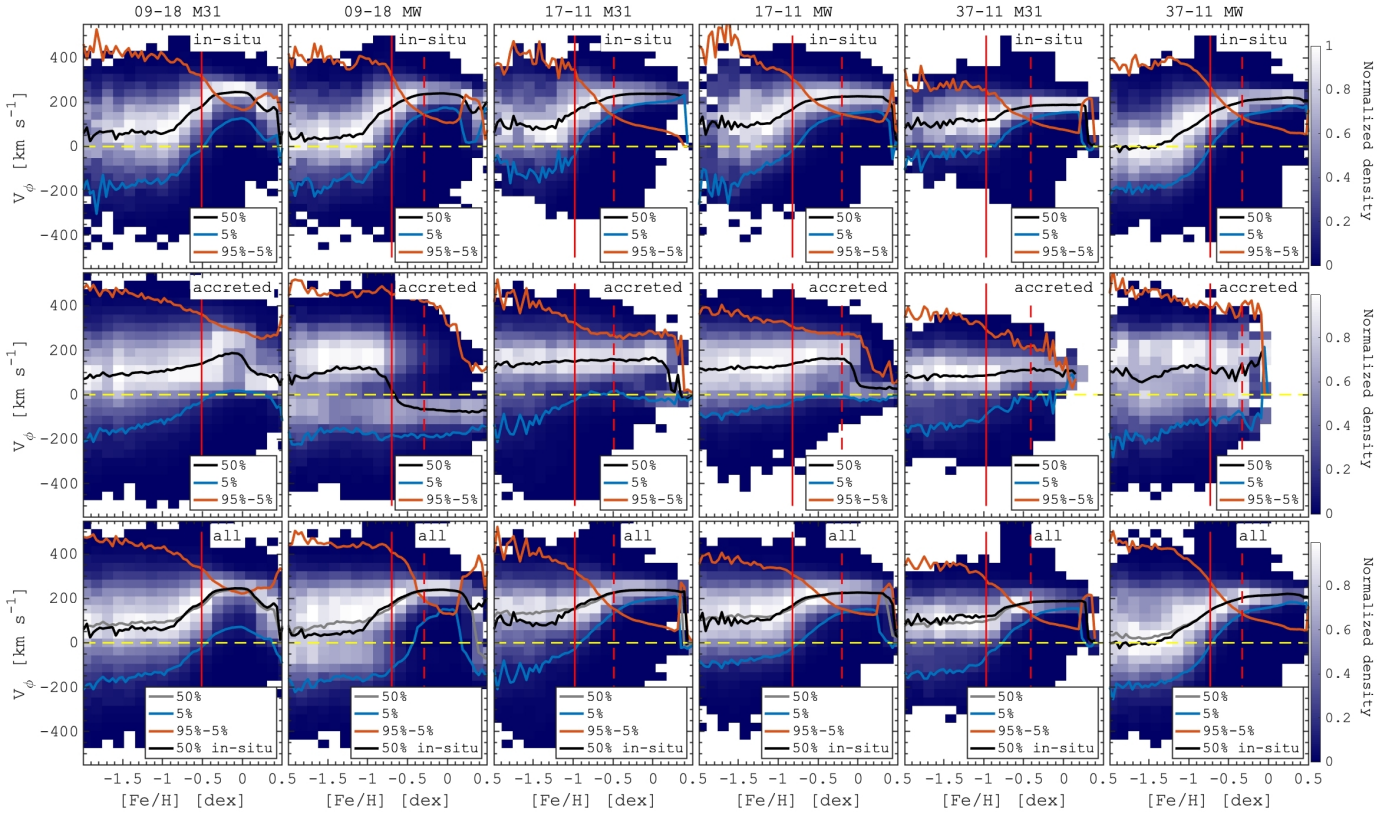
Fig. 14. Azimuthal velocity V_ϕ as a function of metallicity: In situ (top) and all accreted (bottom) stellar populations. Black lines show the mean azimuthal velocity values, while blue and red show 5% and 95−5% statistics of the velocity distribution. The red vertical solid and dashed lines highlight the boundaries similar to the ones used to identify the Aurora and Splash populations in the MW (Belokurov & Kravtsov 2022). The HESTIA galaxies show a behaviour similar to the MW with a little rotation of the low-[Fe/H] ($< -1.2\ldots -0.8$) stars and a sharp spin-up of the net rotation for more metal-rich stars ([Fe/H] < -0.5).

can be used to trace the stars associated with a single merger event.

Next, following a recent analysis of the APOGEE DR17 data by Belokurov & Kravtsov (2022), in Fig. 14, we show a variation of the azimuthal velocity components as a function of the stellar metallicity. Taking advantage of the simulations, we can separate the in situ and accreted populations in different rows. We recall that Belokurov & Kravtsov (2022) found a population of chemically pre-selected in situ stars that show no (or very little) net rotation at [Fe/H] < -1.3, dubbed as Aurora. These stars have a high velocity dispersion and a large scatter in element abundances which, at such low metallicities, can be interpreted as the manifestation of the chaotic pre-disk stages of MW's evolution. At a higher metallicity (between [Fe/H] $= -1.3$ and -0.9), the rotational velocity of the MW stars increases with metallicity up to $\approx 150\,\mathrm{km\,s^{-1}}$ suggesting a rapid transformation ("spin-up") from the turbulent stages to a coherently rotating disc of the MW.

Figure 14 suggests that all the HESTIA galaxies show similar behaviour of the azimuthal velocity as a function of [Fe/H]. In particular, below [Fe/H] < -1.2 (or [Fe/H] < -0.8 depending on the galaxy), the azimuthal velocity profile is flat with a modest net rotation ($\approx 0-80\,\mathrm{km\,s^{-1}}$). These populations, similar to the Aurora, have the highest rotational velocity dispersion. For more metal-rich ([Fe/H] $\lesssim -0.5 \pm 0.1$) in situ stars, the velocity dispersion decreases sharply together with increasing median azimuthal velocity which is in good agreement with the spin-up of the MW (see also Conroy et al. 2022).

Another piece of the analysis concerns the kinematics of the accreted stars. In the bottom row of Fig. 14, we show the

azimuthal velocity as a function of the stellar metallicity for all accreted stellar populations, without dissecting them onto individual merger debris. A striking feature here is that even accreted stars in all the galaxies show some rotation with the mean rotational velocity slowly increasing with the metallicity. There is a single exception here, namely: MW in the 09−18 simulation, which experienced a retrograde merger ≈ 3.7 Gyr ago (see the M5 merger). Nevertheless, our small statistics (albeit biased towards the MW-like galaxies) suggest that the accretion of dwarf galaxies at early times happens predominantly with the same direction of rotation as the host disc, which is likely due to dynamical friction and a torque from the disc (Gómez et al. 2017).

6. Multiple abundances

Recent progress in the understanding of the MW stellar populations is driven by large spectroscopic surveys: RAVE (Steinmetz et al. 2020), LAMOST (Luo et al. 2015), APOGEE (Majewski et al. 2017), GALAH (De Silva et al. 2015), and *Gaia*-ESO (Gilmore et al. 2012). They deliver a number of chemical abundances with statistical errors that allow for studies of different components and stellar populations not only in the disk but, thanks to the mounting growth of the measurements, in a less populated stellar halo as well. Concerning the stellar halo, using various combinations of the abundances (see, e.g., Buder et al. 2022; Myeong et al. 2022; Carrillo et al. 2022; Horta et al. 2023), it becomes possible to differentiate the halo substructures by their chemical composition, which can vary due
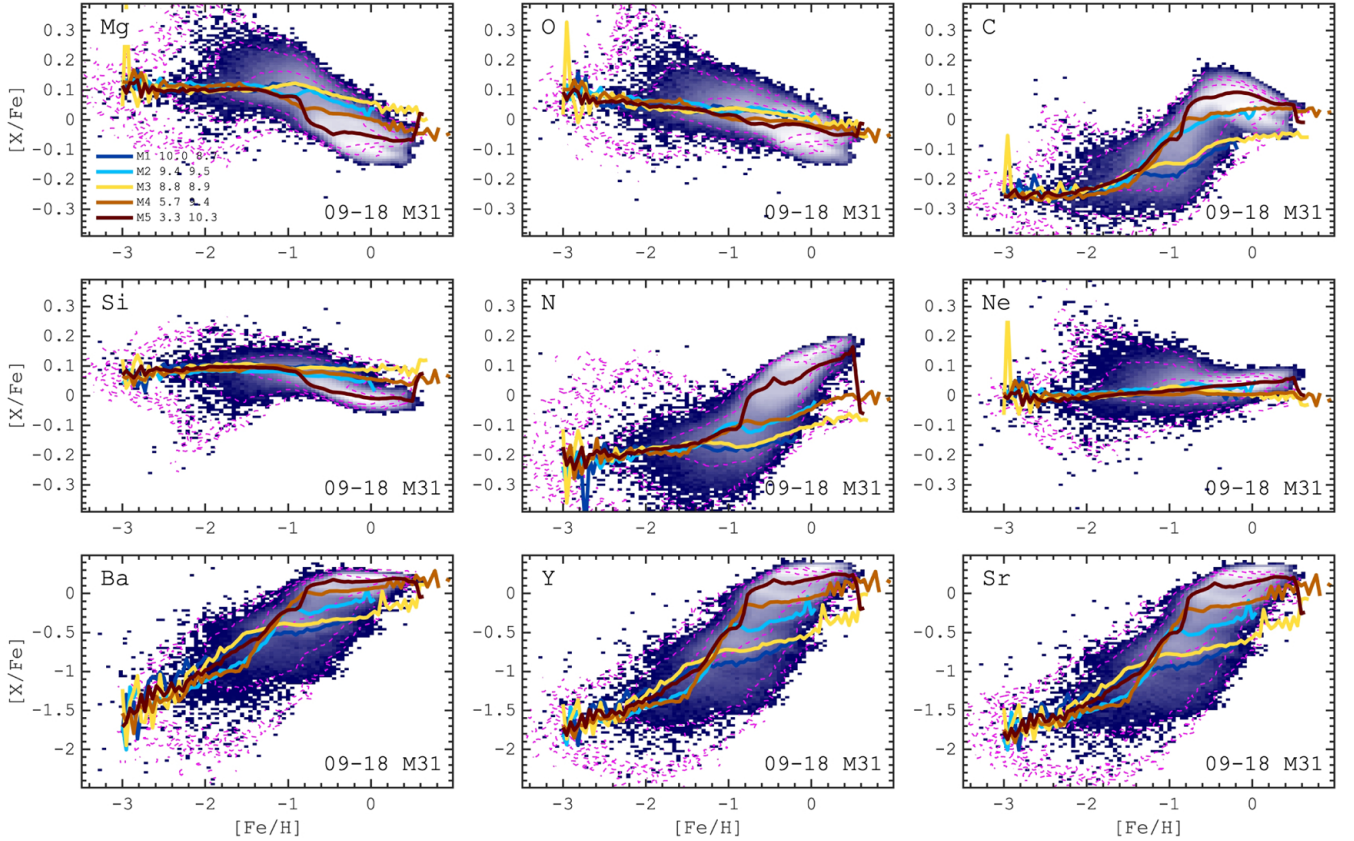
Fig. 15. Elemental abundance relations [X/Fe]–[Fe/H] for the M31 galaxy analogue from the 09–18 HESTIA simulation. X is indicated in each panel. Background maps correspond to the in situ stars, while the magenta contours (10^{-3}, 10^{-2} and 10^{-1} levels of the normalised density distribution) show all accreted stars distribution. The evolutionary tracks corresponding to the five most significant mergers are shown by the coloured lines. The merger accretion time (Gyr) and the total stellar mass of the merger debris at the time of the merger $\log_{10}(M_*/M_\odot)$ are shown in the top-left panel. The abundance patterns of both the accreted and in situ populations are very similar at very low metallicities ([Fe/H] $\lesssim -1$), while at larger metallicities different merger debris start to be distinguishable from each other.

to the initial conditions for the star formation, its efficiency, and the gas fraction. Meanwhile, cosmological and idealized tailored simulations have started to deliver reliable chemical abundances, not only for iron and some [α/Fe]-elements. In this case, of course, the models are still limited by the implementation of the subgrid physics and adopted yields (see e.g., Buck et al. 2021; which may not allow for reproducing observational data quantitatively. Nevertheless, the important information delivered by the simulations is the relative chemical abundance variation of different stellar populations. In particular, we suggest that even if the chemical abundances delivered by the simulations are not exactly on the same scale as the ones measured in the MW, we could still learn about the accreted systems once we compare their chemistry to the in situ stars.

In this section, we analyze a set of the chemical abundances delivered in the HESTIA simulations, which, for the reasons we stated above, may not perfectly match the MW data but still can guide further studies of the chemical abundance trends of accreted stellar populations in both simulations and observational data. First, in Fig. 15, we show the abundance ratio for X = Mg, O, C, Si, N, Ne, Ba, Y, Sr as a function of the iron abundance [Fe/H] in a single galaxy M31 from 09–18 HESTIA simulation. We note that the stellar yields for the s-process elements include a contribution from AGB stars only (Cristallo et al. 2015; Karakas & Lugaro 2016). Coloured lines correspond to the mean trends in [X/Fe]–[Fe/H] for five of the most significant mergers, while the magenta contours highlight the distribution of

all accreted stars and the background colour maps correspond to the distributions of the in situ stellar populations.

We can see that the difference between the merger debris is negligible at very low metallicities, [Fe/H] $\lesssim -2$. This is evidently because such low metallicities are typical for the stars formed very early on, $11-14$ Gyr ago (see Figs. 3 and 10) when the 'initial' conditions for the star formation are very similar inside progenitors of the dwarf galaxies, and the ISM is not enriched enough. At the same time, the main progenitor (MW or M31 in our case) is also similar to other proto-dwarf galaxies and, thus, the first stellar populations have nearly identical abundance relations. At slightly higher (but still low metallicities) of $-2 \lesssim$ [Fe/H] $\lesssim -1$, some elements (C, N, Ba, Y, Sr) start to show a certain level of differentiation but taking into account observational errors in this metallicity range these tiny differences can not be distinguished by the large-scale surveys. The most promising metallicity region is [Fe/H] $= -1 \ldots 0$, where the chemical abundance tracks for individual debris show the largest separation. Obviously, the difference is larger between the mergers the most distant in time. In this case, more recently accreted galaxies, tend to have higher stellar mass; in addition, the highest metallicity stars have lower-[α/Fe] abundances.

To better quantify the difference between the merger debris and the in situ stars, in Fig. 16, we present the difference between the mean individual merger debris abundance and the mean in situ values. We suggest that the residual abundances of the accreted stars relative to the in situ populations allow specific
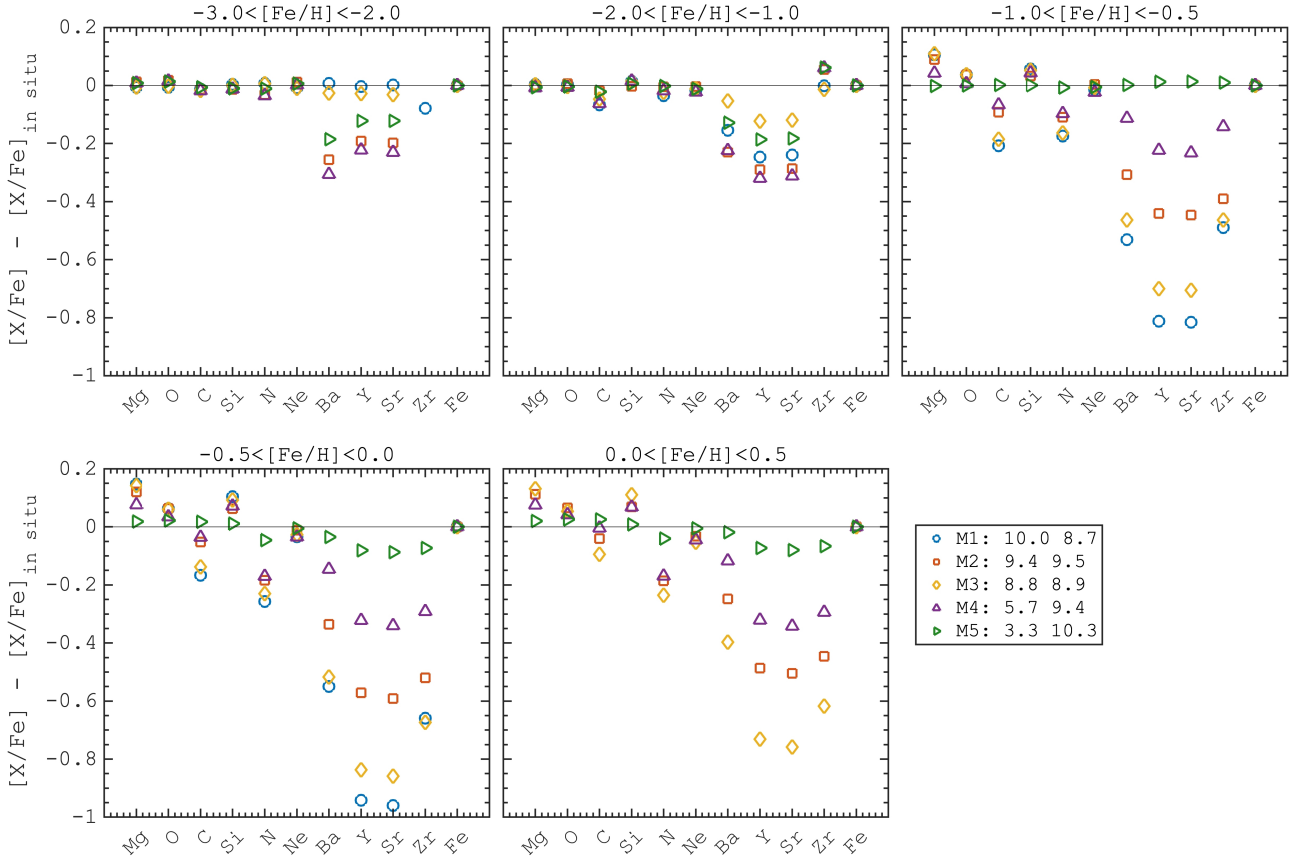
Fig. 16. Mean [X/H] values for the five most significant merger debris relative to the mean values of the in situ populations as a function of metallicity in the M31 09–18 galaxy. The merger accretion time (Gyr) and the total stellar mass of the merger debris at the time of the merger $\log_{10}(M_*/M_\odot)$ are shown in the legend.

trends of the merger debris to be captured more accurately. The relative abundances ([X/Fe] − [X/Fe]$_{\text{in situ}}$) are shown in several bins of the stellar metallicity [Fe/H]. The idea here is to provide typical parameters space where the chemical abundance differences of different merger remnants and the in situ stars are maximal. In other words, we set the criteria that will be used to explore the MW stellar halo in the chemical abundance space with the aim of disentangling merger remnants and in situ stellar populations – and the ultimate goal of recovering the assembly history of the MW. From this figure, we can see that the individual merger debris are very much similar to each other at the very low metallicities, although they are slightly [α/Fe]-enhanced relative to the in situ stars. This difference increases towards the higher-[Fe/H] for the some [α/Fe]-elements (Mg, O, Si), however, it does not exceed ≈0.1 dex. The next abundance value for the accreted stars that shows substantial deviation from the in situ population is carbon, where a significant gap is seen already at $-2 \lesssim$ [Fe/H] $\lesssim -1$ that can be traced out to super-solar metallicities. Another group of elements showing the biggest difference, up to 1 dex, is Ba, Y, Zr. These abundances seem to be the most promising in the differentiation of the merger debris and the in situ populations.

Finally, we summarise the trends presented for a single galaxy (M31 from 09–18 simulation) in Figs. 15 and 16 by combining all the merger debris we identified in all the HESTIA galaxies. In Fig. 17, we show the residual abundances (relative to the in situ populations) of accreted stars as a function of the total stellar mass of the debris where the relations are shown for the stars in three age ranges: <8 Gyr (blue), 8−12 Gyr (red) and >12 Gyr (yellow). We already showed that separating the in situ

and different accreted debris using chemical abundances poses a great challenge due to substantial overlap. Nevertheless, the broad statistics of various galaxies accreted onto the M31/MW allows us to find some systematic differences in the chemical abundances. The most striking feature of the relations we present here is the dependence of the residual abundances on the stellar age. Therefore, one of the key elements in the future studies of the MW assembly history is the age measurements for larger samples of stars.

7. Summary

This paper, as part of a series based on the analysis of the HESTIA cosmological simulations of the Local Group, explores the chemical abundance variations among the accreted populations together with in situ stellar populations. Our main conclusions are as follows:

– We show that the mean metallicity of the entire merger debris increases with the total stellar mass and decreases with the lookback time of accretion (t_{accr}). In particular, for the stellar masses larger than $10^6 M_\odot$, the mean stellar metallicity can be fitted as [Fe/H] $\propto -0.06 \times t_{\text{accr}} + C$ where C increases with the total stellar mass (see Fig. 1). We also find that the chemical abundances of stellar merger debris differ from the properties of the surviving dwarf galaxies. For the same stellar mass, the merger debris are more metal-poor and more [α/Fe]-enhanced (see, Figs. 2 and 4). In analysing the star formation histories of the satellites and merger debris, we show that the chemical abundances of the merger debris are
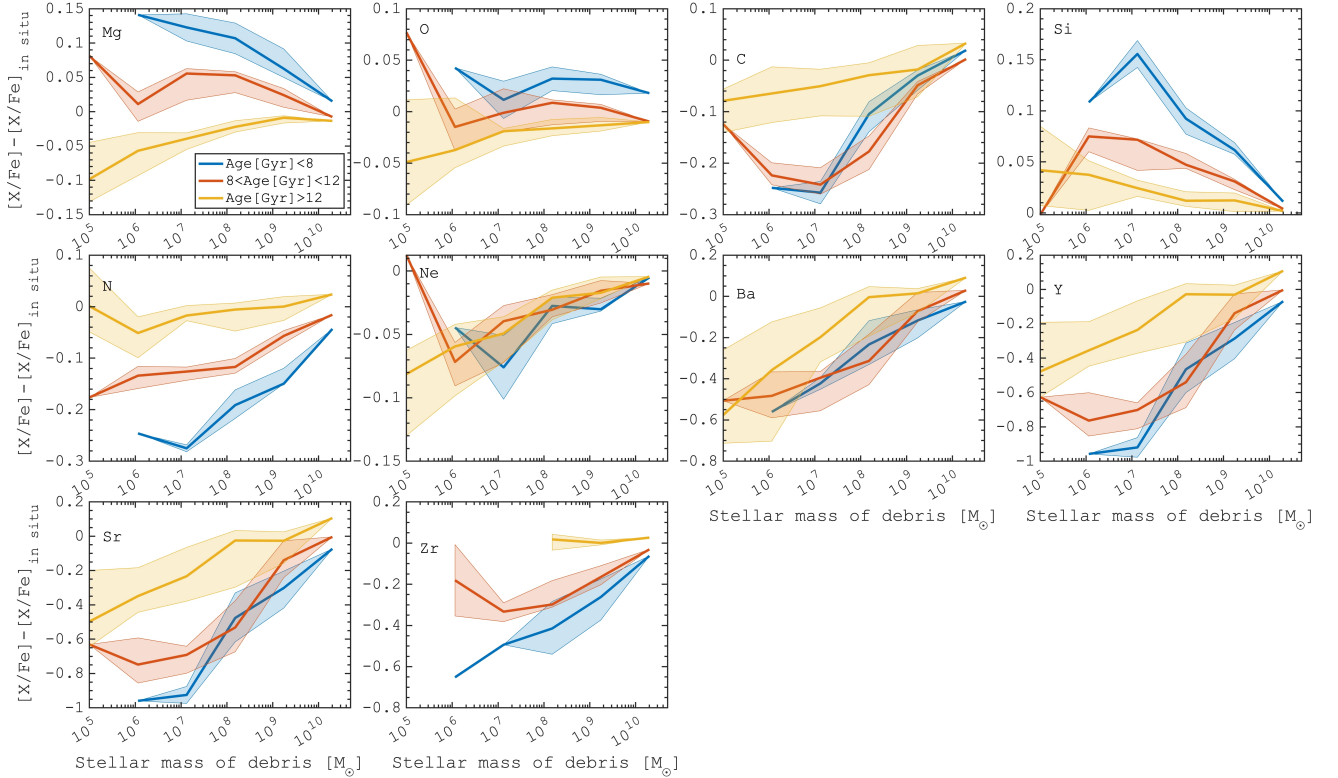
Fig. 17. Relative chemical abundance of accreted stars and in situ populations versus the total stellar mass of the merger debris. The statistics is based on all the merger debris from the six M31/MW HESTIA galaxies. Lines of different colour correspond to different stellar age bins (see top-left panel), while the shaded area shows the 3σ level. The most striking feature of the relations is the dependence of the residual abundances on the stellar age.

the most similar to the dwarf galaxies that quench their star formation early on (see Fig. 2).

- By analysing the metallicity patterns in $E - L_z$ space, we find a prominent diagonal gradient of the mean stellar metallicity, which increases towards the prograde regularly rotating stellar populations (see Fig. 5). The observed gradient does not only result from the in situ stars' fraction variations, as the individual merger debris depict similar metallicity patterns (see Fig. 13). More generally, the mean stellar metallicity of the individual debris increases towards the stars formed in the innermost parts of the disrupted dwarf galaxy. The latter results in the negative metallicity gradient as a function of the galactocentric distance for the massive ($M_{star} > 10^8 M_\odot$) merger debris (see Fig. 9). This result also suggests that the metallicity of the debris measured at the Solar radius in the Milky Way may underestimate the mean metallicity of the debris and, thus, it may either underestimate the total mass or overestimate the accretion lookback time.

- We find that the MDF of some individual merger debris reveal several peaks (see Fig. 7). This result is more prominent for the stars selected in the retrograde part of $E - L_z$ space. The complex composition of MDFs is likely linked to the differentiation of stellar populations accreted from a single dwarf galaxy, according to the initial kinematics inside the dwarf prior to the merger.

- We show that, despite the substantial overlap with the in situ stars, the accreted populations dominate in the low-[Fe/H] regions and in the high-[α/Fe] sequence of the [Mg/Fe]–[Fe/H] plane (see Fig. 11). The [Mg/Fe]–[Fe/H] relations for the individual debris vary significantly depending on the mean angular momentum. We find that the retrograde

stars of the merger debris usually do not show a knee while the prograde populations of the same debris show a more advanced chemical enrichment with prominent 'knees' (see Fig. 12). Therefore, depending on the orbital properties, individual merger debris can mimic chemically different substructures.

- By coupling the chemical abundances with the kinematics of stars we found that the eccentricity of stars from the individual merger debris while having a large scatter in a given chemical abundance range, while remaining nearly constant along [Fe/H] and [Mg/Fe]. At the same time, when we consider the accreted halo made up of all the merger debris, the eccentricity distributions are the complex functions of the [Fe/H] and [Mg/Fe]; for instance, there are regions of abundances with a bimodal distribution of the eccentricity. In situ stars also show a non-featureless distribution of the eccentricity as a function of the chemical composition of stars (see Fig. 13). In this case, the variations of the eccentricity are aptly explained by the impact of the mergers (see also Paper II), where the more metal-poor stars formed before the most massive accretion events were heated up by the mergers and, thus, have higher mean eccentricity.

- In agreement with a recent analysis of the APOGEE data (Belokurov & Kravtsov 2022), the most metal-poor in situ stars ([Fe/H] $\lesssim -1$) in the HESTIA galaxies have a constant azimuthal velocity with a net rotation of $0-80$ km s^{-1} (for different models), which reproduces the Aurora stellar population of the MW well. At higher metallicities (up to [Fe/H] $\approx -0.5 \pm 0.1$), similar to the case of the MW, we detected the rapid spin-up, where the rotational velocity sharply increases up to $150-200$ km s^{-1}. The latter is a

manifestation of the formation of rotating stellar components in the HESTIA galaxies early on (see also Paper I).

- In the [Fe/H]–[Mg/Fe] plane, different merger debris show a similar behaviour, thus making it difficult to capture various individual events. Nevertheless, we show that combining a set of the abundances allows us to capture specific patterns corresponding to different merger debris (see Figs. 15 and 16), thus making it possible to uncover the assembly history of the MW, with the data being delivered by current and forthcoming spectroscopic surveys. According to the HESTIA simulations, a larger differentiation of stellar debris appears for metallicities [Fe/H] ≈ −1 and [Fe/H] ≈ 0.
- The HESTIA simulations predict systematic variations of the residual abundances of accreted stars (relative to the mean abundances of the in situ stars) from the total stellar mass of the debris (see Fig. 17). The relations show different behaviour depending on the age of the stellar populations. This makes the determination of the ages vital for larger samples of the MW stellar populations, and once they are used together with the chemical abundances, it will allow for better constraints to be placed on the merger history of the MW galaxy.

The chemical composition of accreted and in situ stars in the MW reveals distinct characteristics that provide clues about their origins. Accreted stars often exhibit different elemental abundance patterns compared to in situ stars, indicating their diverse stellar birthplaces. These differences in chemical composition arise from the accretion of dwarf galaxies or satellite systems with distinct enrichment histories and enrichment mechanisms. In analyzing the suite of the HESTIA simulations, we highlight the importance of chemical abundances in tracing accreted stars. This, together with growing samples of stars with precise ages, will allow us to trace the history of Galactic growth, along with the various sources that have contributed to the MW stellar halo.

References

Abadi, M. G., Navarro, J. F., & Steinmetz, M. 2006, MNRAS, 365, 747
Agertz, O., Renaud, F., Feltzing, S., et al. 2021, MNRAS, 503, 5826
Amarante, J. A. S., Debattista, V. P., Beraldo e Silva, L., Laporte, C. F. P., & Deg, N. 2022, ApJ, 937, 12
Amorisco, N. C. 2017, MNRAS, 464, 2882
Arentsen, A., Starkenburg, E., Martin, N. F., et al. 2020, MNRAS, 496, 4964
Astropy Collaboration (Robitaille, T. P., et al.) 2013, A&A, 558, A33
Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, AJ, 156, 123
Belokurov, V., & Kravtsov, A. 2022, MNRAS, 514, 689
Belokurov, V., Zucker, D. B., Evans, N. W., et al. 2006, ApJ, 642, L137
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
Brammer, G. B., Whitaker, K. E., van Dokkum, P. G., et al. 2011, ApJ, 739, 24
Brook, C. B., Kawata, D., Gibson, B. K., & Freeman, K. C. 2004, ApJ, 612, 894
Brook, C. B., Gibson, B. K., Martel, H., & Kawata, D. 2005, ApJ, 630, 298
Brook, C. B., Stinson, G. S., Gibson, B. K., et al. 2012, MNRAS, 426, 690
Brook, C. B., Kawata, D., Gibson, B. K., Gallart, C., & Vicente, A. 2020, MNRAS, 495, 2645
Buck, T. 2020, MNRAS, 491, 5435
Buck, T., Rybizki, J., Buder, S., et al. 2021, MNRAS, 508, 3365
Buder, S., Lind, K., Ness, M. K., et al. 2022, MNRAS, 510, 2407
Bullock, J. S., & Johnston, K. V. 2005, ApJ, 635, 931
Carlin, J. L., Garling, C. T., Peter, A. H. G., et al. 2019, ApJ, 886, 109
Carrillo, A., Hawkins, K., Jofré, P., et al. 2022, MNRAS, 513, 1557
Clarke, A. J., Debattista, V. P., Nidever, D. L., et al. 2019, MNRAS, 484, 3476
Conroy, C., Naidu, R. P., Garavito-Camargo, N., et al. 2021, Nature, 592, 534
Conroy, C., Weinberg, D. H., Naidu, R. P., et al. 2022, ApJ, submitted [arXiv:2204.02989]
Cooper, A. P., Cole, S., Frenk, C. S., et al. 2010, MNRAS, 406, 744
Cristallo, S., Straniero, O., Piersanti, L., & Gobrecht, D. 2015, ApJS, 219, 40
Cunningham, E. C., Sanderson, R. E., Johnston, K. V., et al. 2022, ApJ, 934, 172
Deason, A. J., Mao, Y.-Y., & Wechsler, R. H. 2016, ApJ, 821, 5
Deason, A. J., Belokurov, V., & Sanders, J. L. 2019, MNRAS, 490, 3426
de Jong, R. S., Agertz, O., Berbel, A. A., et al. 2019, The Messenger, 175, 3
De Lucia, G., & Helmi, A. 2008, MNRAS, 391, 14
De Silva, G. M., Freeman, K. C., Bland-Hawthorn, J., et al. 2015, MNRAS, 449, 2604
Di Matteo, P., Haywood, M., Lehnert, M. D., et al. 2019, A&A, 632, A4
D'Souza, R., & Bell, E. F. 2018, Nat. Astron., 2, 737
Dupuy, A., Libeskind, N. I., Hoffman, Y., et al. 2022, MNRAS, 516, 4576
Erkal, D., Deason, A. J., Belokurov, V., et al. 2021, MNRAS, 506, 2677
Fakhouri, O., Ma, C.-P., & Boylan-Kolchin, M. 2010, MNRAS, 406, 2267
Fardal, M. A., Guhathakurta, P., Babul, A., & McConnachie, A. W. 2007, MNRAS, 380, 15
Fattahi, A., Deason, A. J., Frenk, C. S., et al. 2020, MNRAS, 497, 4459
Fernández-Alvar, E., Fernández-Trincado, J. G., Moreno, E., et al. 2019, MNRAS, 487, 1462
Feuillet, D. K., Sahlholdt, C. L., Feltzing, S., & Casagrande, L. 2021, MNRAS, 508, 1489
Font, A. S., Johnston, K. V., Bullock, J. S., & Robertson, B. E. 2006, ApJ, 646, 886
Font, A. S., McCarthy, I. G., Crain, R. A., et al. 2011, MNRAS, 416, 2802
Forbes, D. A. 2020, MNRAS, 493, 847
Frenk, C. S., & White, S. D. M. 2012, Annalen der Physik, 524, 507
Gallart, C., Bernard, E. J., Brook, C. B., et al. 2019, Nat. Astron., 3, 932
Gilmore, G., Randich, S., Asplund, M., et al. 2012, The Messenger, 147, 25
Gómez, F. A., Minchev, I., Villalobos, Á., O'Shea, B. W., & Williams, M. E. K. 2012, MNRAS, 419, 2163
Gómez, F. A., Helmi, A., Cooper, A. P., et al. 2013, MNRAS, 436, 3602
Gómez, F. A., Grand, R. J. J., Monachesi, A., et al. 2017, MNRAS, 472, 3722
Gonzalez, O. A., Mucciarelli, A., Origlia, L., et al. 2020, The Messenger, 180, 18
Grand, R. J. J., Gómez, F. A., Marinacci, F., et al. 2017, MNRAS, 467, 179
Grand, R. J. J., Bustamante, S., Gómez, F. A., et al. 2018, MNRAS, 474, 3629
Grand, R. J. J., Deason, A. J., White, S. D. M., et al. 2019, MNRAS, 487, L72
Hasselquist, S., Shetrone, M., Smith, V., et al. 2017, ApJ, 845, 162
Haywood, M., Di Matteo, P., Lehnert, M. D., et al. 2018, ApJ, 863, 113
Helmi, A. 2004, ApJ, 610, L97
Helmi, A. 2020, ARA&A, 58, 205
Helmi, A., White, S. D. M., de Zeeuw, P. T., & Zhao, H. 1999, Nature, 402, 53
Helmi, A., Veljanoski, J., Breddels, M. A., Tian, H., & Sales, L. V. 2017, A&A, 598, A58
Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, Nature, 563, 85
Horta, D., Schiavon, R. P., Mackereth, J. T., et al. 2023, MNRAS, 520, 5671
Hunter, J. D. 2007, Comput. Sci. Eng., 9, 90
Ibata, R. A., Gilmore, G., & Irwin, M. J. 1994, Nature, 370, 194
Jean-Baptiste, I., Di Matteo, P., Haywood, M., et al. 2017, A&A, 604, A106
Johnston, K. V., Spergel, D. N., & Hernquist, L. 1995, ApJ, 451, 598
Johnston, K. V., Sackett, P. D., & Bullock, J. S. 2001, ApJ, 557, 137
Karakas, A. I., & Lugaro, M. 2016, ApJ, 825, 26
Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, MNRAS, 341, 33
Khoperskov, S., Haywood, M., Snaith, O., et al. 2021, MNRAS, 501, 5176
Khoperskov, S., Minchev, I., Libeskind, N., et al. 2023a, A&A, 677, A90 (Paper II)
Khoperskov, S., Minchev, I., Libeskind, N., et al. 2023b, A&A, 677, A89 (Paper I)
Kirby, E. N., Cohen, J. G., Guhathakurta, P., et al. 2013, ApJ, 779, 102
Knebe, A., Gill, S. P. D., Kawata, D., & Gibson, B. K. 2005, MNRAS, 357, L35
Knollmann, S. R., & Knebe, A. 2009, ApJS, 182, 608
Koposov, S. E., Belokurov, V., Li, T. S., et al. 2019, MNRAS, 485, 4726
Koppelman, H. H., Helmi, A., Massari, D., Price-Whelan, A. M., & Starkenburg, T. K. 2019, A&A, 631, L9
Koppelman, H. H., Bos, R. O. Y., & Helmi, A. 2020, A&A, 642, L18

Kruijssen, J. M. D., Pfeffer, J. L., Reina-Campos, M., Crain, R. A., & Bastian, N. 2019, MNRAS, 486, 3180
Kruijssen, J. M. D., Pfeffer, J. L., Chevance, M., et al. 2020, MNRAS, 498, 2472
Lacey, C., & Cole, S. 1993, MNRAS, 262, 627
Laporte, C. F. P., Johnston, K. V., Gómez, F. A., Garavito-Camargo, N., & Besla, G. 2018, MNRAS, 481, 286
Larson, R. B. 1976, MNRAS, 176, 31
Li, Y.S., & Helmi, A. 2008, MNRAS, 385, 1365
Li, T. S., Koposov, S. E., Erkal, D., et al. 2021, ApJ, 911, 149
Libeskind, N. I., Carlesi, E., Grand, R. J. J., et al. 2020, MNRAS, 498, 2968
Lu, Y., Minchev, I., Buck, T., et al. 2022, Nat. Lett., submitted [arXiv:2212.04515]
Luo, A. L., Zhao, Y.-H., Zhao, G., et al. 2015, Res. Astron. Astrophys., 15, 1095
Mackereth, J. T., Crain, R. A., Schiavon, R. P., et al. 2018, MNRAS, 477, 5072
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
Malhan, K., Ibata, R. A., Sharma, S., et al. 2022, ApJ, 926, 107
Marchesini, D., van Dokkum, P. G., Förster Schreiber, N. M., et al. 2009, ApJ, 701, 1765
Marchesini, D., Whitaker, K. E., Brammer, G., et al. 2010, ApJ, 725, 1277
Martin, N. F., Ibata, R. A., Rich, R. M., et al. 2014, ApJ, 787, 19
Martínez-Delgado, D., Romanowsky, A. J., Gabany, R. J., et al. 2012, ApJ, 748, L24
Mateu, C. 2023, MNRAS, 520, 5225
Matsuno, T., Aoki, W., & Suda, T. 2019, ApJ, 874, L35
Matteucci, F. 2021, A&ARv, 29, 5
Monachesi, A., Gómez, F. A., Grand, R. J. J., et al. 2016, MNRAS, 459, L46
Monachesi, A., Gómez, F. A., Grand, R. J. J., et al. 2019, MNRAS, 485, 2589
Monty, S., Venn, K. A., Lane, J. M. M., Lokhorst, D., & Yong, D. 2020, MNRAS, 497, 1236
Myeong, G. C., Vasiliev, E., Iorio, G., Evans, N. W., & Belokurov, V. 2019, MNRAS, 488, 1235
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., Conroy, C., Bonaca, A., et al. 2022, ApJ, submitted [arXiv:2204.09057]
Necib, L., Ostdiek, B., Lisanti, M., et al. 2020, ApJ, 903, 25
Newberg, H. J., Yanny, B., & Willett, B. A. 2009, ApJ, 700, L61
Nidever, D. L., Hasselquist, S., Hayes, C. R., et al. 2020, ApJ, 895, 88

Pagnini, G., Di Matteo, P., Khoperskov, S., et al. 2023, A&A, 673, A86
Pakmor, R., Springel, V., Bauer, A., et al. 2016, MNRAS, 455, 1134
Panithanpaisal, N., Sanderson, R. E., Wetzel, A., et al. 2021, ApJ, 920, 10
Perez, F., & Granger, B. E. 2007, Comput. Sci. Eng., 9, 21
Pilkington, K., Few, C. G., Gibson, B. K., et al. 2012, A&A, 540, A56
Planck Collaboration XVI. 2014, A&A, 571, A16
Queiroz, A. B. A., Chiappini, C., Perez-Villegas, A., et al. 2021, A&A, 656, A156
Reichert, M., Hansen, C. J., Hanke, M., et al. 2020, A&A, 641, A127
Renaud, F., Agertz, O., Andersson, E. P., et al. 2021, MNRAS, 503, 5868
Roederer, I. U., Hattori, K., & Valluri, M. 2018, AJ, 156, 179
Rupke, D. S. N., Kewley, L. J., & Barnes, J. E. 2010, ApJ, 710, L156
Sanders, J. L., & Binney, J. 2013, MNRAS, 433, 1813
Sanders, J. L., Belokurov, V., & Man, K. T. F. 2021, MNRAS, 506, 4321
Shetrone, M., Venn, K. A., Tolstoy, E., et al. 2003, AJ, 125, 684
Shipp, N., Drlica-Wagner, A., Balbinot, E., et al. 2018, ApJ, 862, 114
Simpson, C. M., Gargiulo, I., Gómez, F. A., et al. 2019, MNRAS, 490, L32
Smith, M. C., Evans, N. W., Belokurov, V., et al. 2009, MNRAS, 399, 1223
Springel, V. 2005, MNRAS, 364, 1105
Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629
Steinmetz, M., Matijevič, G., Enke, H., et al. 2020, AJ, 160, 82
Taylor, M. B. 2005, ASP Conf. Ser., 347, 29
Tissera, P. B., Beers, T. C., Carollo, D., & Scannapieco, C. 2014, MNRAS, 439, 3128
Tissera, P. B., Machado, R. E. G., Carollo, D., et al. 2018, MNRAS, 473, 1656
Tolstoy, E., Hill, V., & Tosi, M. 2009, ARA&A, 47, 371
Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898
van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Comput. Sci. Eng., 13, 22
van Dokkum, P., Gilhuly, C., Bonaca, A., et al. 2019, ApJ, 883, L32
Vasiliev, E. 2019, MNRAS, 482, 1525
Vincenzo, F., Spitoni, E., Calura, F., et al. 2019, MNRAS, 487, L47
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nat. Methods, 17, 261
Vogelsberger, M., White, S. D. M., Helmi, A., & Springel, V. 2008, MNRAS, 385, 236
Wes McKinney 2010, in Proceedings of the 9th Python in Science Conference, eds. S. van der Walt, & J. Millman, 56
White, S. D. M., & Frenk, C. S. 1991, ApJ, 379, 52