



Observing the Stellar Halo of Andromeda in Cosmological Simulations: The AURIGA2PANDAS Pipeline

Guillaume F. Thomas^{1,2,3} , Nicolas F. Martin^{4,5} , Azadeh Fattahi⁶, Rodrigo A. Ibata⁴ , John Helly^{6,7}, Alan W. McConnachie¹ , Carlos Frenk⁶ , Facundo A. Gómez^{8,9} , Robert J. J. Grand¹⁰ , Stephen Gwyn¹,

Dougal Mackey¹¹ , Federico Marinacci¹² , and Rüdiger Pakmor¹⁰

¹ NRC Herzberg Astronomy and Astrophysics, 5071 West Saanich Road, Victoria, BC, V9E 2E7, Canada; guillaume.thomas@nrc-cnrc.gc.ca

² Instituto de Astrofísica de Canarias, E-38205 La Laguna, Tenerife, Spain

³ Universidad de La Laguna, Dpto. Astrofísica, E-38206 La Laguna, Tenerife, Spain

⁴ Université de Strasbourg, CNRS, Observatoire astronomique de Strasbourg, UMR 7550, F-67000 Strasbourg, France

⁵ Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117, Heidelberg, Germany

⁶ Department of Physics, Institute for Computational Cosmology, Durham University, South Road, Durham, DH1 3LE, UK

⁷ Carnegie Observatories, 813 Santa Barbara Street, Pasadena, CA 91101, USA

⁸ Instituto de Investigación Multidisciplinar en Ciencia y Tecnología, Universidad de La Serena, Raúl Bitrán 1305, La Serena, Chile

⁹ Departamento de Astronomía, Universidad de La Serena, Av. Juan Cisternas 1200 Norte, La Serena, Chile

¹⁰ Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, D-85748 Garching, Germany

¹¹ Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia

¹² Department of Physics & Astronomy, University of Bologna, via Gobetti 93/2, I-40129 Bologna, Italy

Received 2020 August 28; revised 2021 January 12; accepted 2021 January 22; published 2021 March 30

Abstract

We present a direct comparison of the Pan-Andromeda Archaeological Survey (PAndAS) observations of the stellar halo of M31 with the stellar halos of six galaxies from the Auriga simulations. We process the simulated halos through the AURIGA2PANDAS pipeline and create PAndAS-like mocks that fold in all observational limitations of the survey data (foreground contamination from the Milky Way stars, incompleteness of the stellar catalogs, photometric uncertainties, etc.). This allows us to study the survey data and the mocks in the same way and generate directly comparable density maps and radial density profiles. We show that the simulations are overall compatible with the observations. Nevertheless, some systematic differences exist, such as a preponderance for metal-rich stars in the mocks. While these differences could suggest that M31 had a different accretion history or has a different mass compared with the simulated systems, it is more likely a consequence of an underquenching of the star formation history of galaxies, related to the resolution of the AURIGA simulations. The direct comparison enabled by our approach offers avenues to improve our understanding of galaxy formation as they can help pinpoint the observable differences between observations and simulations. Ideally, this approach will be further developed through an application to other stellar halo simulations. To facilitate this step, we release the pipeline to generate the mocks, along with the six mocks presented and used in this contribution.

Unified Astronomy Thesaurus concepts: [Andromeda Galaxy \(39\)](#); [Galaxy stellar halos \(598\)](#); [Galaxy structure \(622\)](#); [Optical observation \(1169\)](#); [Astronomical simulations \(1857\)](#)

1. Introduction

How were L^* galaxies like the Milky Way (MW) or Andromeda formed? Behind this relatively simple question hides a very complex picture, starting from cosmological scales and trickling down to subparsec scales. In the standard Λ -cold dark matter cosmological paradigm, the formation of structures happens hierarchically, with dark matter halos growing through a succession of mergers, their inhabitant galaxies absorbing the smaller galaxies located at the center of these accreted halos (Searle & Zinn 1978; White & Rees 1978; Moore et al. 1999). Due to dynamical friction, a large fraction of the accreted stars, coming mostly from major mergers, sinks rapidly toward the Galactic disk (e.g., Boylan-Kolchin et al. 2008; Pillepich et al. 2015; Gómez et al. 2017), where dynamical timescales are of order a few hundred million years, removing the traces of these accretion events. Fortunately, the imprints of the merger history are conserved over a very long period of time (many gigayears; Bullock & Johnston 2005; Johnston et al. 2008) in the most distant, and faintest (reaching surface brightnesses of $\mu_V \sim 35$ mag arcsec⁻²), component of a galaxy, its stellar halo. Therefore, this component is key to unveiling the details of the formation history of L^* galaxies, such as the number of

past accretion events, the mass of the accreted galaxies, and the epoch of these events. The stellar halos of MW-like galaxies are very diffuse and extended, reaching out to their virial radius (e.g., Helmi & White 1999; Cooper et al. 2010; Gómez et al. 2013). They are often complex structures, very inhomogeneous and clumpy (e.g., Bell et al. 2008; Martínez-Delgado et al. 2010; Duc et al. 2015; Merritt et al. 2016; McConnachie et al. 2018), with the presence of numerous satellite galaxies, globular clusters, and their tidal debris (streams, shells, plumes, etc.) visible for several gigayears (Johnston et al. 2008).

Because stellar halos are very extended and only account for a few percent of the light emitted from a galaxy, it is very challenging to observe them. Despite earlier studies that detected faint components around nearby galaxies and discovered several stellar streams (e.g., Malin & Hadley 1997; Mihos et al. 2005; Martínez-Delgado et al. 2010; Martínez-Delgado & Mediavilla 2013), it is only recently that integrated light photometry can efficiently be used to measure the surface brightness profiles of galactic halos (Duc et al. 2015; Merritt et al. 2016, 2020; Trujillo & Fliri 2016; D’Souza & Bell 2018; Wang et al. 2019). A tremendous amount of work has been invested to better characterize the point-spread function of

photometric observations and to improve the optics of the instruments, allowing for a better separation of the scattered light from galaxies and contaminating scattered light produced by galactic cirrus, stars, and other compact objects (Slater et al. 2009). These relatively inexpensive observations¹³ allow for the observation of a great number of galaxies with a great diversity of halo profiles stemming from the stochasticity of the galaxy formation process. Yet, with this technique, it is very challenging to study the detailed properties of individual halos, such as their radial age and metallicity distributions, or, even harder, to obtain their dynamical properties.

On the other hand, resolved stellar photometry is more informative and allows the measurement of the physical properties of individual halos, at the cost of very expensive observations. In particular, the MW halo has been extensively studied due to its relatively close distance that allows main-sequence stars to be observed out to a few dozen kiloparsecs (e.g., Carollo et al. 2007, 2010; Ivezić et al. 2008; Jurić et al. 2008; Bell et al. 2008, 2010; Watkins et al. 2009; Sesar et al. 2011; Deason et al. 2012; Gómez et al. 2012; Deason et al. 2013; Xue et al. 2011, 2015; Ibata et al. 2017; Fukushima et al. 2018, 2019; Hernitschek et al. 2018; Thomas et al. 2018). However, it is extremely costly to study the halo of the MW, since a large area of the sky has to be observed and, ideally, in a broad range of photometric bands to facilitate the separation of different stellar populations. Furthermore, to study the profile of the halo of our Galaxy, it is necessary to use three-dimensional distances due to our central position, but only a few percent of the halo stars have distances known with good enough accuracy (Ibata et al. 2017, but see Thomas et al. 2019). Thus, in some respects, the halos of nearby external galaxies are easier to study than the halo of our Galaxy. They cover a smaller area on the sky, and, with some assumptions, it is possible to use the projected distances to infer their profile with a better accuracy and at larger distances than for the MW.

In this endeavor, the Hubble Space Telescope (HST) has been a powerful tool to study the halo of nearby galaxies, in particular thanks to the ACS Nearby Galaxy Survey Treasury (Dalcanton et al. 2009, ANGST) and to the Galaxy Halos, Outer Disks, Substructure, Thick Disks, and Star Clusters (GHOSTS Radburn-Smith et al. 2011) programs. These surveys have revealed a great diversity of masses, of metallicity distributions, and of stellar populations in the stellar halos of L^* galaxies that otherwise have similar disk morphologies, masses, and luminosities (e.g., Ibata et al. 2009; Monachesi et al. 2016; Harmsen et al. 2017). However, these surveys are often pencil beam and do not offer a complete view of the halo of the target galaxies.

Therefore, the halo of the Andromeda galaxy (M31) is an important and unique object due to our ability to probe the stellar populations of an L^* galaxy at very faint stellar magnitudes, as was done with HST (e.g., Brown et al. 2006, 2007; Richardson et al. 2009) and, in parallel, over a very large area, as done by the Pan-Andromeda Archaeological Survey (PAndAS, McConnachie et al. 2009, 2018). This latter survey, covering >400 deg² around M31, has been extremely important to study the morphology and chemistry of the stellar halo of a galaxy similar to the MW (i.e., Ibata et al. 2014) but also to quantify its level of substructures (stellar streams, dwarf galaxies, or globular clusters; i.e., Mackey et al. 2010;

Collins et al. 2013; Lewis et al. 2013; Martin et al. 2013; Huxor et al. 2014; McConnachie et al. 2018). This survey has shown, for instance, that the Andromeda galaxy has a stellar halo that is about 15 times more massive than that of the MW ($\sim 10^{10} M_\odot$ for M31 and $4 - 8 \times 10^8 M_\odot$ for the MW; Bell et al. 2008; Deason et al. 2011, 2019; Ibata et al. 2014), with a shallower profile slope, and displays a metallicity gradient that drops from $\langle [\text{Fe}/\text{H}] \rangle = -0.7$ at 30 kpc to $\langle [\text{Fe}/\text{H}] \rangle = -1.5$ at 150 kpc (Ibata et al. 2014, hereafter I14), contrary to what is observed in the MW (e.g., Ivezić et al. 2008; Jurić et al. 2008; Sesar et al. 2011; Xue et al. 2015; Ibata et al. 2017). These differences between the two galaxies tend to indicate that they have been subjected to two very different formation histories (e.g., Deason et al. 2013; Gilbert et al. 2014; Harmsen et al. 2017; D’Souza & Bell 2018).

From a simulation point of view, analytical models and dark matter-only simulations (with different prescriptions to include baryons) that take into account only the accreted component of a halo showed very early that the broad diversity of halos observed in nearby galaxies is a natural consequence of the stochasticity of the merger/accretion history for each galaxy (e.g., Bullock & Johnston 2005; Renda et al. 2005; Font et al. 2006; Purcell et al. 2007; De Lucia & Helmi 2008; Cooper et al. 2010). Moreover, these simulations also demonstrated that the majority of the halo material of a given galaxy has been contributed by the few most massive satellite galaxies accreted through its history (e.g., Robertson et al. 2005; Cooper et al. 2010; Deason et al. 2016; Amorisco 2017). However, as shown by Bailin et al. (2014), the simplifying assumptions made by dark matter-only simulations tend toward a systematic underestimate of the halo concentration and an incorrect quantification of the level of substructure of the halo. More recent cosmological hydrodynamical simulations such as Eris (Pillepich et al. 2015), APOSTLE (Sawala et al. 2016; Oman et al. 2017), Auriga (Grand et al. 2017), Illustris TNG (Pillepich et al. 2018a), Latte/FIRE (Wetzel et al. 2016; Hopkins et al. 2018), or ARTEMIS (Font et al. 2020) can model more representative stellar halos by self-consistently including the baryonic (stellar) distribution. With these simulations, it has been possible to confirm that a correlation exists among the number of significant progenitors, the metallicity, and the mass of a halo. For instance, they show that halos made by a few significant progenitors tend to be more massive, more concentrated, and with a significant negative metallicity gradient (Deason et al. 2016; D’Souza & Bell 2018; Pillepich et al. 2018b; Monachesi et al. 2019). In addition, these simulations include an in situ stellar component of the stellar halo, composed of stars born in the galactic disk that have been ejected by the interactions with subhaloes or molecular clouds. They also include stars formed in streams of gas stripped from infalling satellites and dominating the inner halo of L^* galaxies (e.g., Purcell et al. 2010; Font et al. 2011; Pillepich et al. 2015; Cooper et al. 2015).

Therefore, important information on the formation history of galaxies like ours can be gained by comparing the simulations to the observations. The observations are useful to constrain the assumptions and limitations of the simulations, and the simulations can provide useful physical context to explain the formation of the galaxies. However, only a rigorous apples-to-apples comparison between the simulations and the observations, using the same tools and the same methods, with consistent biases and limitations and with the same

¹³ In comparison to observations that aim to resolve individual stars.

assumptions, are meaningful to improve the underlying physical models used by the simulations, especially concerning the baryonic physics.

The aim of this paper is to pursue a direct comparison between simulations and observations of Andromeda’s stellar halo, by transforming six galaxies from the AURIGA simulations into PAndAS-like mocks, as presented in Section 2. The mocks are compared to the observations using the same tools in Section 3, including a discussion on the importance of the in situ population in the simulations in Section 3.3. The implications of the results are analyzed and discussed in Section 3.4, and we conclude in Section 4.

2. Method

In this section, we describe how we transform the simulated stellar halos of six of the 30 Andromeda-like galaxies from the Auriga suite of simulations (Grand et al. 2017) to “realistic” stellar halo mocks as if they were observed by the PAndAS (McConnachie et al. 2009, 2018).

The PAndAS survey was a large program of the Canada–France–Hawaii Telescope (CFHT) that observed the surrounding of the Andromeda (M31) and Triangulum (M33) galaxies. This program comprises 406 fields of $1^\circ \times 1^\circ$ obtained in the g and i bands with the MegaPrime/MegaCam camera between 2008 and 2011 and also includes fields from a pilot survey between 2003 and 2008 with the same instrumental setup. A detailed description of the PAndAS data, their acquisition and reduction, and the resulting catalogs of resolved stars can be found in Ibata et al. (2014) and McConnachie et al. (2018). In particular, the location of the survey fields can be seen in Figure 1 of I14.

The AURIGA simulations are a suite of 30 cosmological magnetohydrodynamical zoom-in simulations of MW-like galaxies, made with the moving mesh magnetohydrodynamics code AREPO (Springel 2010; Pakmor et al. 2016). These galaxies were selected from the parent dark matter-only cosmological simulation EAGLE (the EAGLE team 2017) to have a similar halo mass to the MW and to satisfy being isolated at $z = 0$. We refer the reader to Grand et al. (2017) for a detailed description of these simulations.

2.1. Generation of the Initial Mock Stellar Catalogs

The mock stellar catalogs used as inputs of this pipeline are computed from the Auriga simulations (Grand et al. 2017) in a very similar manner to what was previously presented by Grand et al. (2018a) to produce the AURIGA mock stellar catalogs. To generate these stellar mocks, two methods were presented, the HITS-MOCKS and the ICC-MOCKS. These two methods differ in how the “stars” generated from the stellar particles of the simulations are distributed in phase space, as well as by the choice of stellar evolution models. For the rest of this paper, we use exclusively the ICC-MOCKS, based on Lowing et al. (2015). The reason is that this method produces smoother distributions of “stars” and avoids discrete clumps at the coordinates of the parent stellar particles, while still preserving the phase-space distribution. Therefore, with this method, the presence of artificial features, such as fake clumps, in the distribution of “stars”¹⁴ from the stellar halo in the

Auriga simulations is minimized. With this method, each parent stellar particle is split into N “stars,” assuming a Chabrier initial mass function (IMF) and drawing from a model stellar population of a given age and metallicity according to the PARSEC isochrones (Bressan et al. 2012), which take into account the evolution of the massive end of the IMF for stellar particles with old ages. The stellar parameters and the absolute magnitudes in the CFHT g and i bands¹⁵ are computed for each mock star from the PARSEC isochrones (Bressan et al. 2012), using the age and metallicity of the parent stellar particle.

The foreground MW extinction was added at a later step, as will be described later in Section 2.2. Indeed, our study is focused on the region of M31 observed by the PAndAS survey, that is, the stellar halo, where the extinction caused by the interstellar medium (ISM) of M31 is negligible. The large majority of the dimming of stars observed in the stellar halo of M31 are due to the foreground extinction caused by the ISM of the MW and are included later in the pipeline.

It is important to note here that for the analysis of the simulations, detailed in Section 3, the central projected 20 kpc (or 1.47° at the chosen distance of M31) are not taken into account. This region is severely incomplete in the PAndAS survey, and a study of this region would require dedicated work (e.g., HST PHAT survey Dalcanton et al. 2012). Therefore, from this stage and for each mock, we decided to completely avoid “stars” within the a central sphere of 20 kpc radius. This action considerably reduces the size of the mocks and the computation time to apply the AURIGA2PANDAS pipeline since the large majority of the simulation “stars” reside in this region. By applying this cut to the three-dimensional distances instead of the projected distances at this step, we keep the possibility to project the galaxy using a random point of view around the galaxy (see below).

2.2. The AURIGA2PANDAS Pipeline

From these raw mocks that include all “stars” down to an absolute magnitude of $M_i = 2$ (~ 6 magnitude fainter than the tip of the red giant branch [RGB] at the distance of M31), we degrade the data to be as close as possible to the observed PAndAS data. The different steps we perform are as follows:

1. Placing the simulation at the distance of M31, orienting it similarly to M31 and projecting it on the sky;
2. Masking out “stars” that are not in a PAndAS field or behind saturated foreground stars;
3. Computing the apparent magnitude of the “stars” and their associated uncertainties;
4. Making the data incomplete following the observed, field-specific completeness functions;
5. Selecting the RGB “stars” with a color–magnitude cut;
6. Adding the contamination from foreground MW stars and background unresolved galaxies.

The effects of these steps on the distribution of “stars” for one of the halos is visible in Figure 1. We now discuss each of these stages in turn.

¹⁴ In the rest of the paper, the mock stars generated from a simulation’s star particles will be referred as “stars” to differentiate them from the real stars of the PAndAS observations.

¹⁵ Here we use the pre-2014 CFHT/MegaCam photometric system, before the current set of filters was built. It is worth noting that the fields obtained before 2007 have been observed with a slightly different i -band filter (see Figure 3 of McConnachie et al. 2018), but this change is subtle enough that it should not affect our results.

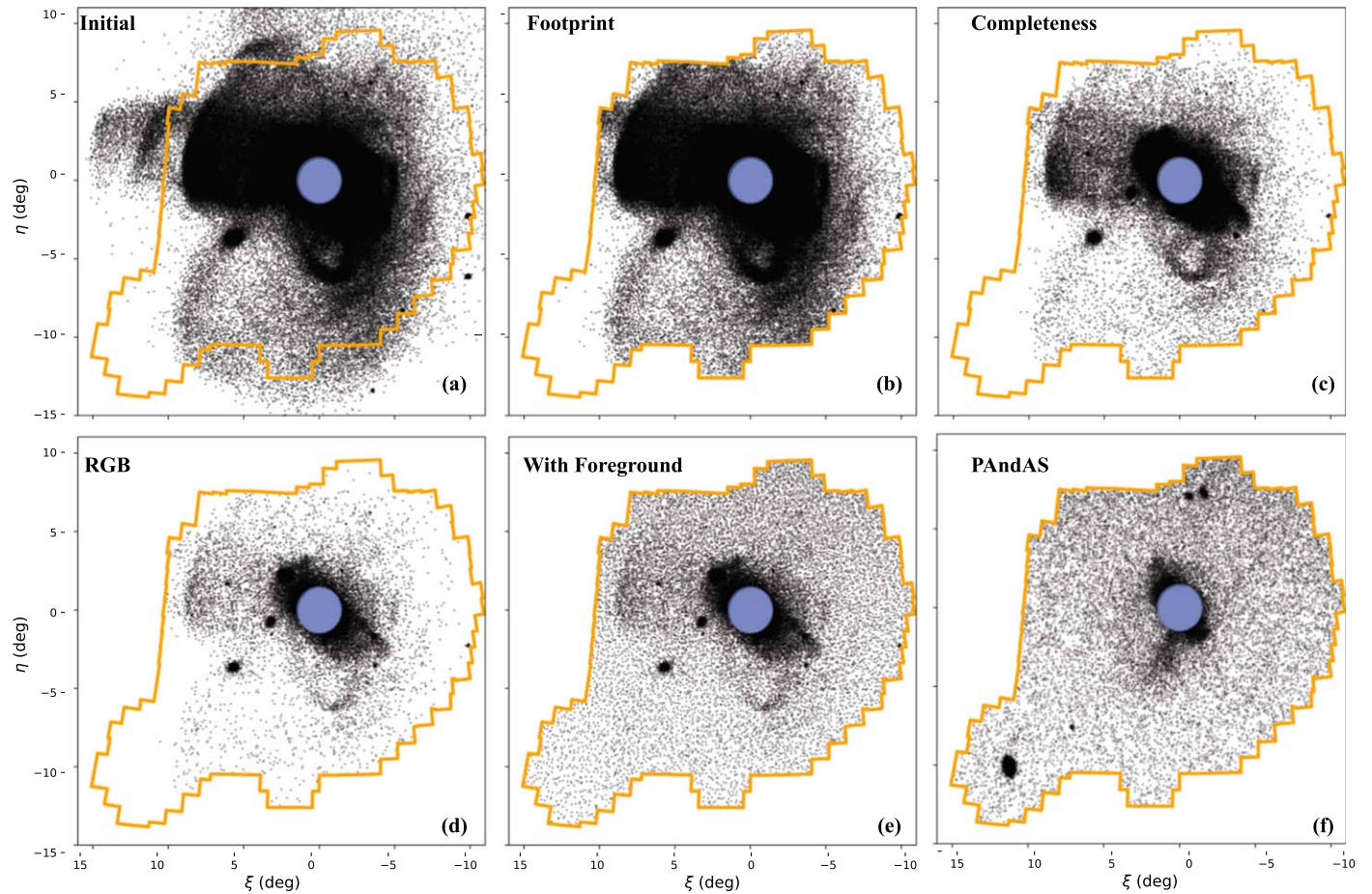


Figure 1. Example of the projection of 1,000,000 initial “stars” from halo H23 into the observational space. In all panels, we show the plane of the sky tangent to the celestial sphere at the location of M31. Panel (a) presents the initial distribution of mock “stars.” Panel (b) shows the same stars after application of the PAndAS footprint. In panel (c), we further apply the completeness criteria, while panel (d) shows the selection of “stars” within the RGB box of Martin et al. (2013). Finally, panel (e) displays the final state of the mock, after adding “stars” from the MW foreground contamination model. In all panels, the orange polygon traces the external border of the PAndAS footprint, and the blue circle highlights the inner 20 kpc that have been initially removed from the mock to mask the majority of the disk stars in the simulations since they are irrelevant to this study. Panel (f) shows a subselection in PAndAS that has a similar density of stars as in panel (e) 6° outside from the center of M31 and 2° outside M33. The gray scale is the same in all panels.

2.2.1. Projection

Once the initial mock stellar catalogs are built, we project them on the sky, placing the center of the simulated galaxy at the position of M31 (R.A., decl.) = $(00^{\text{h}}42^{\text{m}}44^{\text{s}}.330, +41^\circ16'07''.50)$ (Skrutskie et al. 2006) and at a heliocentric distance of 778 kpc (Conn et al. 2011, 2012). The simulations are oriented such that their Galactic disks have a similar orientation to the disk of the Andromeda galaxy, with an inclination $i = 77.5^\circ$ and a position angle $\theta = 37.7^\circ$, following Metz et al. (2007). Even though throughout this paper we only show and analyze a single realization of each simulated galaxy, the AURIGA2PANDAS pipeline allows the user to rotate the disk around the vertical galactic axis with a random angle for each realization, to allow for future work to study statistically the properties of a given “observed” simulation with different points of view. The coordinates of each “star” are computed in the plane tangential to the celestial sphere at the location of M31, (ξ, η) . As per convention, ξ is increasing toward the west and η toward the north. Throughout the rest of the paper, only one realization for each Auriga halo is considered. A statistical analysis of the simulations, incorporating different points of view for each halo will be performed in the future.

The apparent velocities of the “stars” are also computed from their physical velocity, to which we add the global motion of M31, $(V_{\text{los}}, \mu_\alpha^*, \mu_\delta) = (-300 \text{ km s}^{-1}, 65 \mu\text{as yr}^{-1}, -57 \mu\text{as yr}^{-1})$ (McConnachie 2012; van der Marel et al. 2019), assuming a solar radius of 8.1 kpc (Gravity Collaboration et al. 2018), a circular velocity at the solar radius of 229 km s^{-1} (Eilers et al. 2019), and a solar peculiar motion of $(U_\odot, V_\odot, W_\odot) = (11.1, 12.24, \text{ and } 7.25) \text{ km s}^{-1}$ in local standard of rest coordinates (Schönrich et al. 2010).

In Figure 1(a), we show the impact of this step to mock H23. Here we only show a random subsample of 1,000,000 “stars” in the figure for clarity.

2.2.2. PAndAS Footprint Mask

We apply the mask of the PAndAS coverage to the mocks to remove “stars” outside the footprint. This includes the “stars” that are outside the external border of the survey but also “stars” that fall in the small number of holes between some of the observed fields or in the gaps between the lines of CCDs in the MegaPrime/MegaCam camera (see Figure 2 of McConnachie et al. 2018). “Stars” at locations in the (ξ, η) plane that correspond to bright foreground stars in the PAndAS data are also removed from the mock catalog, similar to how we treat the observed data

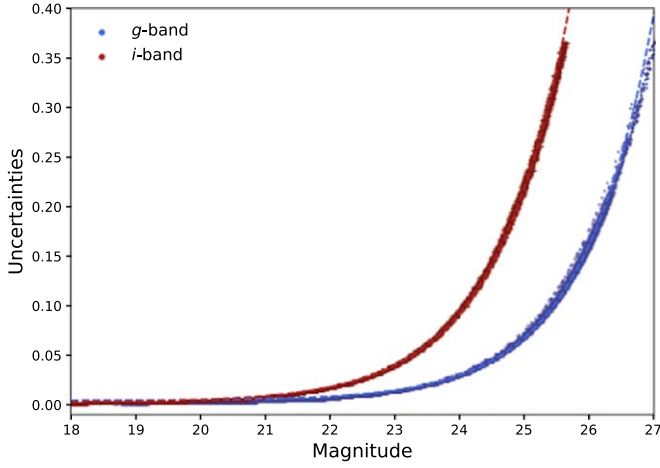


Figure 2. Photometric uncertainties as a function of the observed magnitude in the g (blue) and i bands (red) for the point sources observed in reference field 10. The fitted relations of Equation (2) are represented by the two dashed lines.

(see Section 2 of Ibata et al. 2014). Figure 1(b) shows the impact of this step of the procedure on the 1,000,000 stars of mock H23 shown in panel (a) of the same figure.

2.2.3. Apparent Magnitude

For all unmasked “stars,” their apparent magnitudes in the g and i bands are determined from the absolute magnitudes, given in the initial mock stellar catalog, and from their individual heliocentric distances, computed in the previous step. To account for the foreground extinction produced by the ISM of the MW, the absolute magnitudes are reddened using the $E(B - V)$ extinction map of Schlegel et al. (1998). We further assume a 10% uncertainty on these values to mimic our likely imperfect extinction correction for the PAndAS data and to avoid reddening the data by the exact same amount we will later deredden them by when we study the mocks like we study the PAndAS data. With the knowledge of $E(B - V)$ at a given location in the survey, we redden the data, using the coefficients from Martin et al. (2013) (hereafter referred to as M13), such that

$$\begin{aligned} g &= g_0 + 3.793 E(B - V), \\ i &= i_0 + 2.086 E(B - V). \end{aligned} \quad (1)$$

Here g_0 and i_0 refer to the perfect apparent magnitude of a “star” contained in the mock catalog, while g and i are their reddened equivalent, comparable to the observed and calibrated magnitudes in the PAndAS catalog.

The photometric uncertainties that we assign to each mock “star” are then computed from their apparent magnitude. In order to build a model for the uncertainties as a function of magnitude, we first chose an observed reference field (field 10), for which we isolate all point sources.¹⁶ We use those to build a model of the photometric uncertainties as a function of magnitude (Figure 2) that we model with the

¹⁶ As in I14, we define as point sources objects that have a classification flag from the Cambridge Astronomical Survey Unit pipeline (Irwin & Lewis 2001) of either -1 or -2 in both the g and i bands.

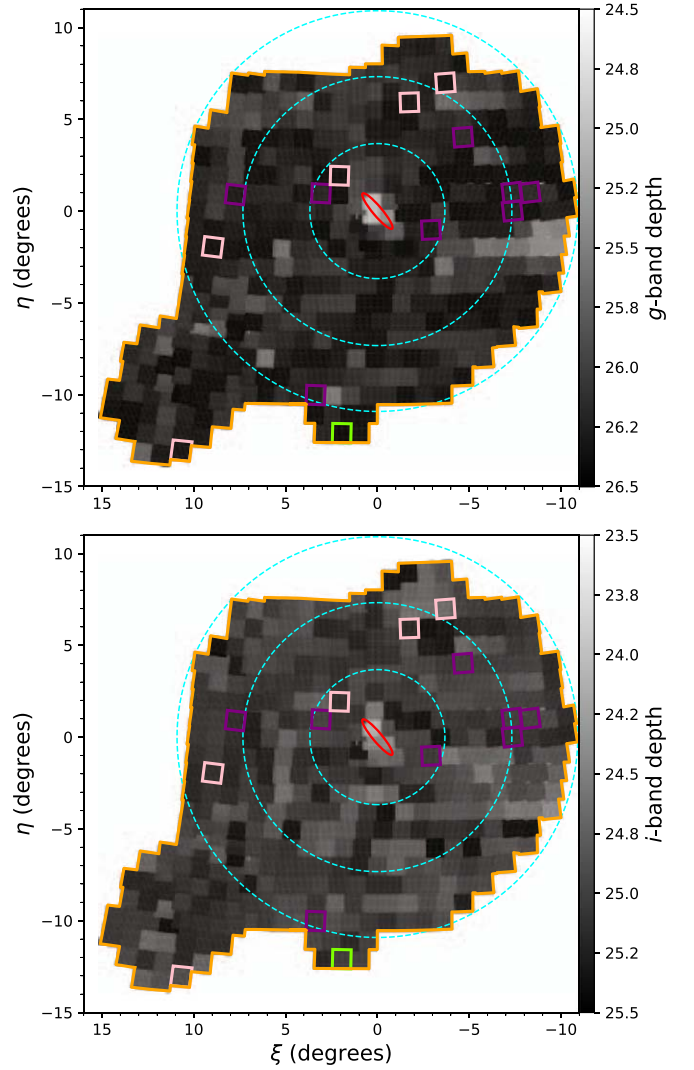


Figure 3. Map of the photometric depths of all PAndAS fields for point sources in the g (top) and i bands (bottom). The cyan circles correspond to projected distances of 50, 100, and 150 kpc, and the red ellipse shows the approximate edge of the main stellar disk of M31. The reference field used to calibrate the photometric uncertainties is represented by the green square. The fields around the globular clusters used to calibrate the completeness function are highlighted by the purple squares, while the pink squares show the fields used to validate this calibration.

following functions:

$$\begin{aligned} \delta g &= 0.032 \exp\left(\frac{g - 24.25}{1.10}\right) + 0.004, \\ \delta i &= 0.112 \exp\left(\frac{i - 24.25}{1.13}\right) + 0.003. \end{aligned} \quad (2)$$

However, it is important to note that the depth, and so the related photometric uncertainties, is different for each field, due to changes in the observing conditions of any PAndAS field. These depths, $\lim_{g,j}$ and $\lim_{i,j}$, as defined by the magnitude for which the observed photometric uncertainties reach 0.2 mag, are determined for all PAndAS fields and shown in Figure 3 for the two observed bands. Substituting $g - (\lim_{g,j} - \lim_{g,10})$ for g (respectively $i - (\lim_{i,j} - \lim_{i,10})$ for i) in Equation (2) allows us to shift the uncertainty models from reference field 10 to any field j . Only moderate inhomogeneities are apparent over the survey footprint, and only the two fields covering the M31 disk

are noticeable outliers, with a shallower depth in both the g and i bands.

For every “star” in the mock, we determine the PAndAS field it falls in given its (ξ, η) location and we randomly draw an uncertainty in g and in i based on the models described above. We then update the apparent magnitude of this “star” by adding random Gaussian deviates based on these modeled δg and δi . The “noisy” measurements are those stored in the mock catalogs.

Finally, the g and i magnitudes are corrected from the foreground extinction using the exact $E(B - V)$ from Schlegel et al. (1998).

2.2.4. Completeness

With the next step, we aim to take the photometric completeness of the survey into account. To estimate the completeness of PAndAS in each band, we compare the number of PAndAS point-source objects as a function of magnitude with the number of point sources observed in the deeper HST fields obtained by Mackey et al. (2007) and Mackey et al. (2013) around globular clusters of M31. These deep HST fields have been observed in the $F606W$ and $F814W$ bands¹⁷ by the Advanced Camera for Surveys (ACS) and are ~ 2.5 – 3 magnitudes deeper than the PAndAS survey. The data reduction of these fields and the star/galaxy separation will be described in a future contribution (Mackey et al. in prep.).

Over the 46 fields observed by HST, we selected the 14 fields located around globular clusters B517, H1, PA 02, PA 03, PA 06, PA 11, PA 18, PA 43, PA 44, PA 45, PA 46, PA 47, PA 49, and PA 56. These HST fields have been selected because they are located in PAndAS fields, for which the depth is similar in the g and i bands, namely, in fields 35, 229, 243, 261, 263, 267, 274, and 335, with a mean depth in these field of 26.03 ± 0.03 in g and 24.85 ± 0.04 in i , close to the mean depth of the overall PAndAS survey. The transformation from the HST photometric system to the MegaPrime/MegaCam system is performed using objects detected as point sources in the V , I , g , and i bands in these 14 fields. It is worth noting here that the objects located in the inner $25''$ of each clusters are not taken into account in the rest of the analysis as these regions can suffer from crowding. By cross-matching the HST fields to the PAndAS ones, we find that the relations between these two photometric systems for point sources are the following:

$$\begin{aligned} g &= \begin{cases} -0.18 + 0.98(V - I) + V, & \text{if } (V - I) < 1.30 \\ 0.92 + 0.13(V - I) + V, & \text{if } (V - I) \geq 1.30, \end{cases} \\ i &= 0.33 + 0.18(V - I) + I. \end{aligned} \quad (3)$$

Using these color transformations, the completeness of PAndAS in those fields is determined independently in g and i , by comparing the number of objects identified in PAndAS as point sources, in this specific band, per bin of 0.5 mag, to the number of point sources objects identified in V and I in the HST fields. This method was preferred over an artificial star-test experiment directly into the PAndAS images because this latter method is very computationally expensive and would not have improved the quality of our work or of the pipeline, especially regarding other steps where the assumptions made have

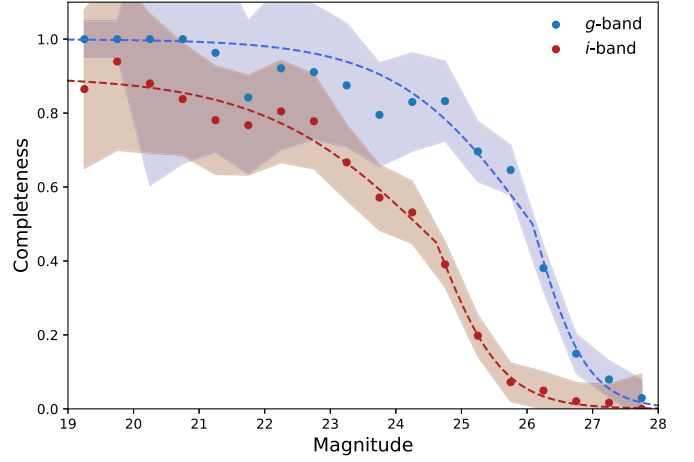


Figure 4. Completeness in the g (blue) and i bands (red) of the PAndAS survey, built from the data in reference fields 35, 229, 243, 261, 263, 267, 274, and 335. The shaded area represents the uncertainties on the completeness values, including a constant value of 0.05 added to the Poissonian uncertainties. The dashed lines show the best-fit relations detailed in Equation (4).

stronger impacts on the final representation of the simulations (e.g., the split of stellar particles into “stars”). The assumption made by the method chosen here is that all stars down to the PAndAS depth are present in the HST observations, which seems reasonable since the HST fields are ~ 2.5 – 3 magnitudes deeper than the PAndAS observations. The completeness of the g and i bands determined in this way is shown in Figure 4. We find that the completeness functions, $C_g(g)$ and $C_i(i)$ for the g and i bands, respectively, can be fit reasonably with the following functions:

$$\begin{aligned} C_g(g) &= \begin{cases} \left[1 + \exp\left(\frac{g - 26.08}{1.04}\right) \right]^{-1}, & \text{if } (g < 26.08) \\ \left[1 + \exp\left(\frac{g - 26.08}{0.41}\right) \right]^{-1}, & \text{if } (g \geq 26.08), \end{cases} \\ C_i(i) &= \begin{cases} 0.9 \left[1 + \exp\left(\frac{i - 24.62}{1.31}\right) \right]^{-1}, & \text{if } (i < 24.62) \\ 0.9 \left[1 + \exp\left(\frac{i - 24.62}{0.50}\right) \right]^{-1}, & \text{if } (i \geq 24.62). \end{cases} \end{aligned} \quad (4)$$

As for the photometric uncertainties, the completeness of a large survey like PAndAS that has been observed over many years varies from field to field, reflecting the specific observational conditions of each field. To account for this spatial variation, it is possible to replace g and i in Equation (4) by $g' = g - (\text{lim}_{g,j} - 26.03)$ and $i' = i - (\text{lim}_{i,j} - 24.85)$, respectively, where $\text{lim}_{g,j}$ and $\text{lim}_{i,j}$ are the depth of field j in the g and i bands found previously and 26.03 and 24.85 are the mean depths of the reference fields for the completeness determination. We validated this method with additional HST fields located around the dwarfs galaxies And X, And XVII, And XXI, And XXIV, And XXV, and And XXVI (Martin et al. 2017). These HST fields are located in fields 3, 194, 296, 376, and 390, for which the mean depth is 26.12 ± 0.15 in g and 24.83 ± 0.14 in i . The results for these fields are presented in Figure 5, with the gray lines showing the completeness determined by our method and the blue/red dashed lines

¹⁷ The Hubble $F606W$ and $F814W$ filters correspond roughly to V and I in the Johnson system, so we will abusively refer to them as such in the rest of the paper.

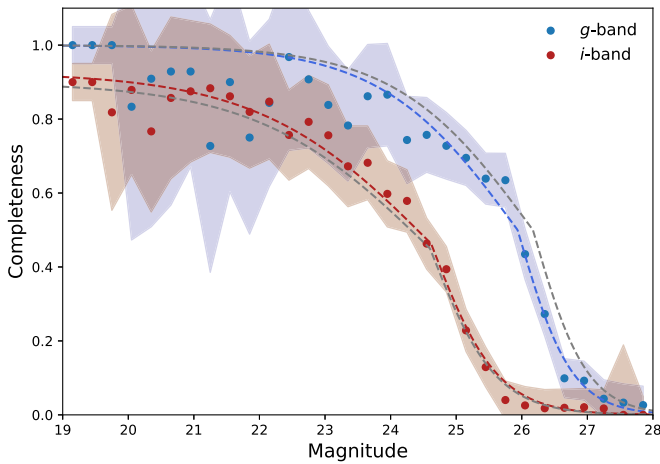


Figure 5. Completeness of the PAndAS fields around dwarf galaxies And X, And XVII, And XXI, And XXIV, And XXV, and And XXVI. The gray lines show the fits of the models described in Equation (4), shifted to the depth of the fields, while the red and blue lines show the best fit of the completeness for these fields. The shaded areas represent the 1σ of the completeness.

representing the best fit of the completeness in this field in g and i . Our method gives similar results to the best fit, with differences of only a few percent, especially for $i < 23.5$ and for $g < 25.5$, where RGB stars at the distance of Andromeda are present (M13).

Once we have a completeness model, we apply this model to the mocks by using an acceptance-rejection method such that “stars” are removed from the catalog if $\text{Rand}(0, 1) > C_g(g, \text{lim}_{g,j})$ and $\text{Rand}(0, 1) > C_i(i, \text{lim}_{i,j})$. Here, $\text{Rand}(0, 1)$ is a number drawn from a uniform distribution between 0 and 1. The result of this operation yields the “star” distribution shown in Figure 1(c) for halo H23.

2.3. Selection of the RGB Stars and Inclusion of the Foreground Contamination

Following M13, we keep stars that have a color and magnitude compatible with those of M31 RGB stars and remove “stars” outside of the selection box whose $(i_0, (g - i)_0)$ vertices are $(21, 0.7)$, $(21, 2.3)$, $(23.5, 0.4)$, $(23.5, 1.6)$. For halo H23, this step yields the distribution shown in Figure 1(d).

Despite this selection, a large fraction of PAndAS stars present in this region of the color–magnitude diagram (CMD) are actually contaminant objects. The source of this contamination is mostly due to the foreground MW dwarf stars and to unresolved background galaxies that appear as point sources. Therefore, in our attempt to produce realistic “observed” stellar halo mocks, it is important to add this contamination component, especially since the density of contaminants is not constant over the survey because of the increasingly dense MW disk toward the north. The addition of this component is extremely important, in order to treat the observations and the “realistic” mocks with the same methods, since most analyses will remove, or at least take into account, the contamination to enhance the signal produced by the stars from the stellar halo of M31.

The model of the contamination used by the pipeline is based on the four-dimensional (spatial and color–magnitude) model of M13, for which the density of contaminant objects at a position (ξ, η) and at a given color and magnitude $(g - i, i)$

follow an exponential such that

$$\Sigma_{(g-i,i)}(\xi, \eta) = \exp(\alpha_{(g-i,i)} \xi + \beta_{(g-i,i)} \eta + \gamma_{(g-i,i)}). \quad (5)$$

The $\alpha_{(g-i,i)}$, $\beta_{(g-i,i)}$, and $\gamma_{(g-i,i)}$ parameters were determined for any location $(g - i, i)$ in the CMD from a region outside ~ 120 kpc from Andromeda’s center, assuming that the density of objects in this external region is produced uniquely by contamination (the reader is referred to M13 for a detailed description of the model). The parameters $\alpha_{(g-i,i)}$, $\beta_{(g-i,i)}$, and $\gamma_{(g-i,i)}$ used here are slightly different from the ones of M13 since we now use the new public reduction of PAndAS presented in McConnachie et al. (2018).

Practically, and following M13, the number of contaminant objects per spatial pixel of 15×15 arcmin², N_j , is computed using Equation (5) for the value of $\alpha_{(g-i,i)}$, $\beta_{(g-i,i)}$, and $\gamma_{(g-i,i)}$ and integrating the model counts in the CMD region delineated by the M31 RGB selection box mentioned above. This number is multiplied by $\mathcal{F} = 0.93$, since I14 show that $\simeq 7\%$ of the objects of the external region over which the contamination model was constructed, and which are in a color–magnitude region similar to the M31 RGB box, are actually stars from the halo of M31. Moreover, to incorporate statistical fluctuations to the contamination model, the actual number of contaminants for each pixel is drawn randomly, assuming a Poissonian distribution centered on N_j . The spatial locations of the N_j contaminant particles of a given pixel are then randomly distributed following a uniform distribution over the pixel.

The g and i magnitudes of the contaminant particles are then randomly distributed following the probability distribution function of the contaminant CMD at the location of the spatial pixel of the “stars” (see M13). It is important to note here that the magnitudes given by the model are already corrected for the extinction. The photometric uncertainties of each contaminant “star” are then computed using Equation (2). Finally, only the contaminant “stars” that are contained in the M31 RGB box are kept. This removes the few particles from pixels cut by the M31 RGB box that are outside this box. Depending on the science goal with the mocks, the M31 RGB box criterion can be removed from the overall pipeline (i.e., for particles from the stellar mock and for contaminants). However, in that case the number of contaminants will be slightly overestimated, as mentioned earlier, but by less than 7%, since this number was obtained in the M31 RGB box, which contains most of the M31 halo stars.

The final distribution of the “stars” for mock H23, including “stars” from the stellar mock and from the contamination model, is shown in Figure 1(e).

The density of stars for a full realization of this halo is presented in Figure 6. The lower left panel shows the total density of “stars” that are in the color–magnitude space delimited by the M31 RGB box. This includes the “stars” from the stellar mock, whose distribution is shown in the upper left panel, and contaminants, shown in the upper right panel. While the details of the substructures are of course different because of the difference in the accretion history of the simulated galaxy and Andromeda, we note qualitative consistency between the mock and the PAndAS data.

3. Results

The AURIGA2PANDAS pipeline is applied to the six simulated galaxies selected by Grand et al. (2018a), H6, H16, H21, H23, H24, and H27, whose parameters are listed in Table 1. These

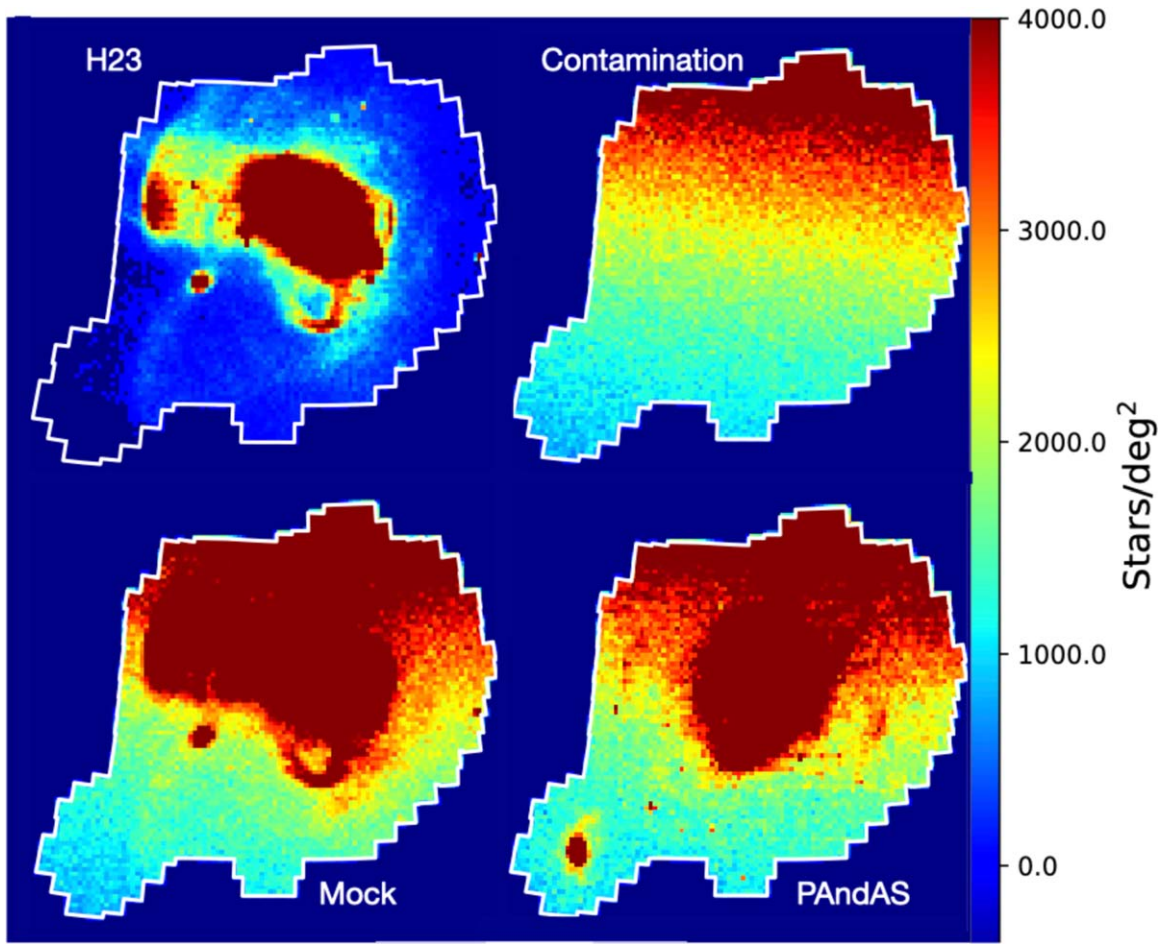


Figure 6. Upper left: density of remaining particles from the H23 stellar mock in the M31 RGB box at the end of the pipeline. Upper right: density of contaminant particles, integrated over the M31 RGB box. Lower left: total density of particles of the final “observed” mock over the M31 RGB box. Lower right: density of actual observed objects in the M31 RGB box over the PAndAS footprint.

Table 1
Table of Parameters of the Stellar Halos from AURIGA

No. Sim.	R_{vir} (kpc)	M_{vir} ($10^{12} M_{\odot}$)	M_{stellar} ($10^{10} M_{\odot}$)	M_{acc} ($10^{10} M_{\odot}$)	$M_{\text{in-situ}}$ ($10^{10} M_{\odot}$)
H6	213.82	1.04	5.41	0.38	0.64
H16	241.48	1.50	7.01	0.50	0.85
H21	238.64	1.45	8.65	1.17	1.03
H23	245.27	1.58	9.80	0.90	0.79
H24	240.85	1.49	7.66	0.64	0.74
H27	253.80	1.75	10.27	0.85	1.02

Note. The values are from Monachesi et al. (2019).

galaxies have virial masses¹⁸ ranging from 1.0 to $1.8 \times 10^{12} M_{\odot}$ and have virial radii of ~ 240 kpc. They cover the range of virial masses found for M31 using different tracers (0.8 to $2.0 \times 10^{12} M_{\odot}$; Chemin et al. 2009; Watkins et al. 2010; Tollerud et al. 2012; Fardal et al. 2013; Veljanoski et al. 2014; Peñarrubia et al. 2016; Kafle et al. 2018). The stellar mass of the simulated halos inside their virial radii varies from 6.6 to $10.3 \times 10^{10} M_{\odot}$. Their stellar masses inside the inner 30 kpc range from 6.2 to $9.6 \times 10^{10} M_{\odot}$, which is slightly less massive than the $10.3^{+2.3}_{-1.7} \times 10^{10} M_{\odot}$ found by Sick et al. (2015) for

¹⁸ Here we define the virial radius as the radius where the mean density is equal to 200 times the critical density of the universe.

M31 inside this radius. However, in the outer halo (>30 kpc), the simulations have stellar masses of 0.8 to $1.0 \times 10^{10} M_{\odot}$, similar to the mass of the stellar halo of M31 found by I14 ($\simeq 1 \times 10^{10} M_{\odot}$). The only exception is the simulation H6, where the stellar mass outside 30 kpc is smaller than for the other simulated galaxies ($0.3 \times 10^{10} M_{\odot}$) but is the consequence of this galaxy being the less massive of this sample.

Prior to any analysis and following M13 and I14, we fill the holes of the PAndAS coverage present in the observations and included in the mocks by the pipeline. To do so, we duplicate stars from neighboring regions, both for the mocks and for the observations. To select the duplicated stars, the observations of the mocks are shifted by $0^{\circ}.15$ along both the ξ and η axes, and the stars falling in the gaps are kept.

The fully processed mocks that are directly comparable to the PAndAS observations are distributed with this publication. A full description of the catalogs is provided in the Appendix, and the data themselves are accessible online.

3.1. Qualitative Description of the Simulations

Figure 7 shows a mapping of the structures present in the stellar halo of M31 and in the six mocks. Each of these maps is a red-green-blue image, made from the combination of matched filter (MF) maps whose filters are CMD models of old RGB stars at the distance of M31, convolved by the photometric uncertainties. The background model of the MF technique is

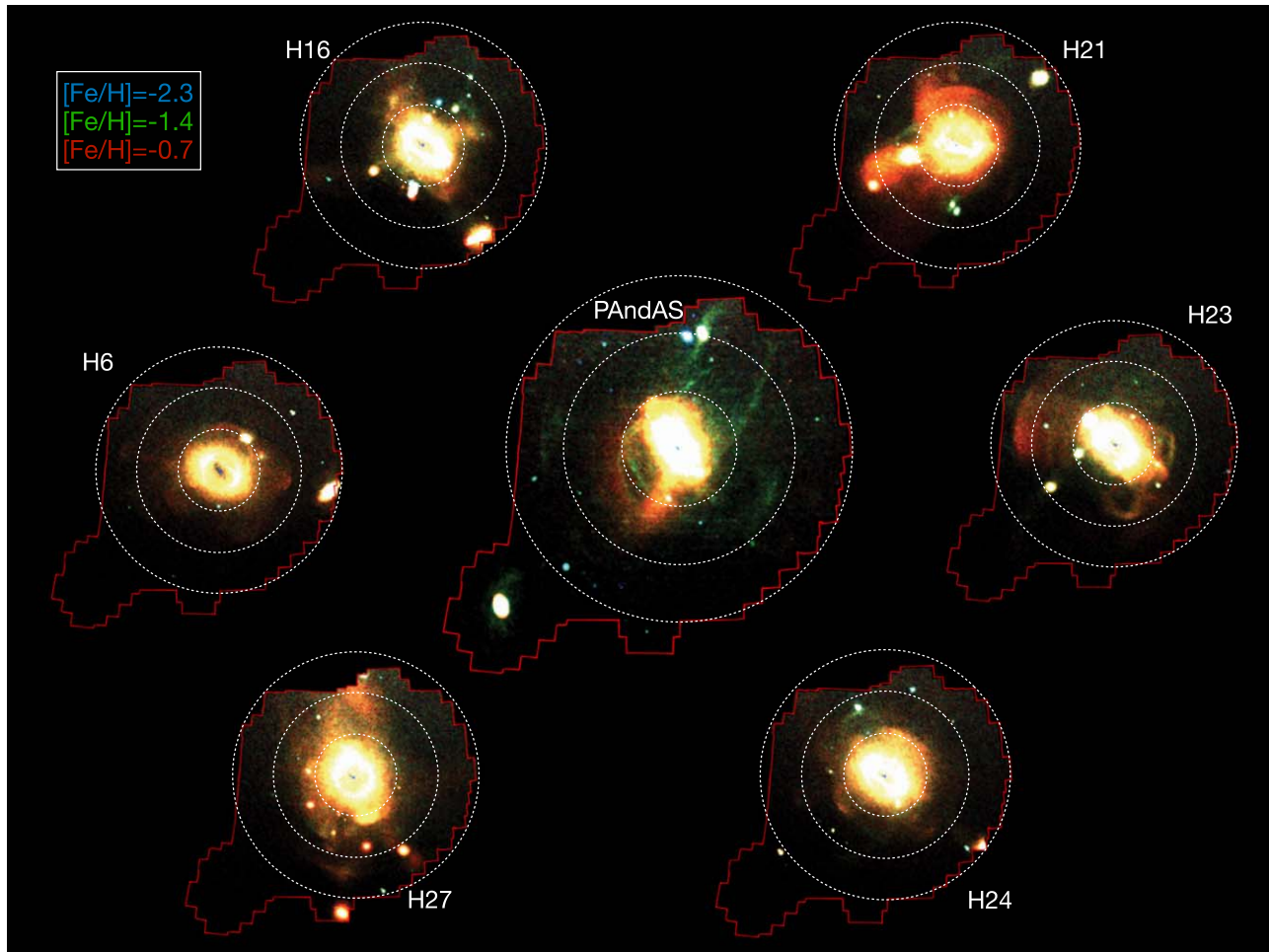


Figure 7. Red-green-blue maps of the observed PAndAS survey (at the center) and of the six AURIGA mock simulations. Each color channel is a MF map using, as the signal, the CMD of an old RGB population at the distance of M31 with metallicities of $[\text{Fe}/\text{H}] = -2.3$ (blue), $[\text{Fe}/\text{H}] = -1.4$ (green), and $[\text{Fe}/\text{H}] = -0.7$ (red) and with a local background CMD that follows the contamination model of M13. The white circles correspond to projected distances of 50, 100, and 150 kpc.

the contamination model of M13. The blue image corresponds to a signal from RGB stars with a metallicity of $[\text{Fe}/\text{H}] = -2.3$, the green to $[\text{Fe}/\text{H}] = -1.4$, and the red to $[\text{Fe}/\text{H}] = -0.7$. The maps are made of spatial pixels of $3' \times 3'$ and have been smoothed by a Gaussian kernel of 1 pixel width. This is similar to the MF technique used by M13 to produce the map of their Figure 2. The holes visible at the very center of the maps for the mocks are caused by the cut made to mask the stars within a three-dimensional radius of 20 kpc, as described in Section 2.1.

A visual inspection of these maps shows that the simulated stellar halos present numerous, well-defined substructures on a similar scale to those observed around M31 (central panel). In the central 50 kpc, the simulations and observations are very similar, with an inner stellar halo dominated by metal-rich stars ($[\text{Fe}/\text{H}] \sim -0.7$) and a small number of identifiable structures. Most of the mocks, with the exception of H6 and H24, have a large and metal-rich stream or shell, a sign of a recent, or ongoing, massive accretion. These structures are similar to the Giant Stream¹⁹ (Ibata et al. 2001) visible in PAndAS that is the consequence of the accretion of a galaxy with a mass similar to the Large Magellanic Cloud 2–3 Gyr ago (Fardal et al. 2013).

On average, the mock halos are redder, and so more metal rich, than the observations. The halo of M31 presents more structures, such as stellar streams or clouds, at intermediate/low metallicity than the simulations. It is important to note here that the absence of a galaxy similar to M33 in some of the simulations is not surprising. Indeed, M33 would not be visible in the PAndAS footprint if it were located at the same projected distance but at a different angle around M31, as is likely to be the case in the mocks. Furthermore, not all simulated halos presently have a satellite that is as massive as M33.

For each of the halos, we compute the surface density of RGB stars in spatial pixels of $0.1^\circ \times 0.1^\circ$ for different ranges of metallicities, in the same way as for the observations. In each of these pixels, the number of contaminant objects has been statistically removed assuming the model of M13. To compare directly the mocks to the observations, the metallicities have been derived photometrically for each star by comparing their colors and magnitudes to the Dartmouth isochrones (Dotter et al. 2008) assuming that the stars are at the distance of M31 and have an age of 10 Gyr and an alpha enhancement of $[\alpha/\text{Fe}] = +0.2$, typical of the populations in the halo of M31/MW (i.e., Helmi 2008; Kilic et al. 2019). The surface density maps in four ranges of metallicities (all, metal-poor, intermediate-metallicity, and metal-rich stars) for the halo of M31 and for mock H23 are shown on Figure 8. The information this

¹⁹ We follow Lewis et al. (2013) and McConnachie et al. (2018) for the nomenclature of the halo structures of M31.

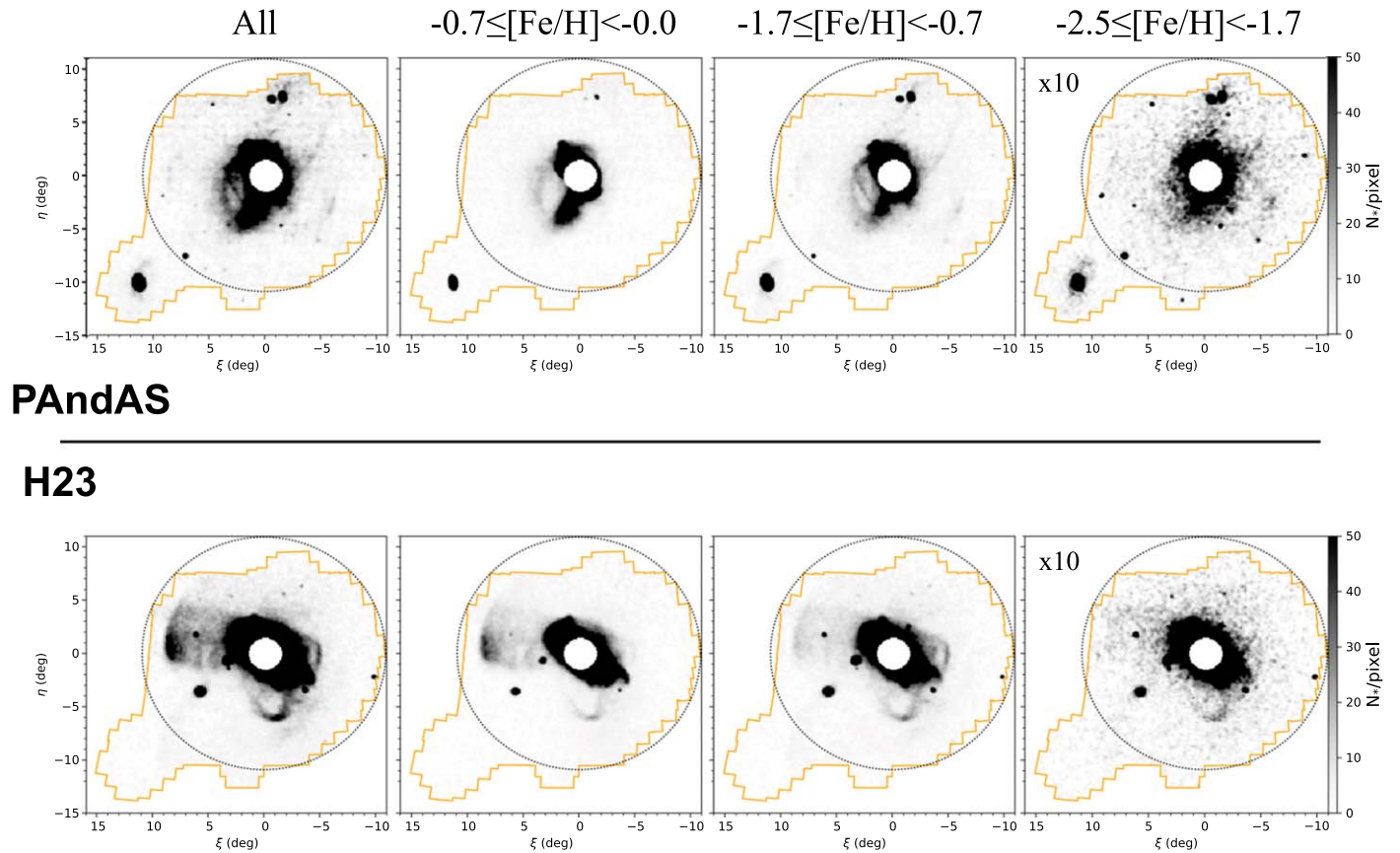


Figure 8. Map of stars in PAndAS (top panels) and in the mock of halo H23 (bottom panels) for different ranges of photometric metallicities. The density has been multiplied by 10 for the lower metallicities panel (right-hand panel) for better visibility. The contamination from the foreground MW dwarfs and from the unresolved background galaxies has been removed statistically, assuming the contamination model of M13. The dotted circle represents a projected distance of 150 kpc.

figure contains is similar to that visible in Figure 7, but it is easier to see the contribution of each structure in the different ranges of metallicity. For instance, it clearly shows that H23 has a similar number of dwarf galaxies at intermediate metallicities compared with M31 but has only about half its number of metal-poor satellite galaxies, similar to AndXXI or AndXXIII ($M_V \sim -9.5$). This characteristic is present in all mocks, especially for H6 and H21, for which the metal-poor galaxies are far less numerous but also more centrally distributed than those observed around M31.

We use these surface density maps to derive the median surface density radial profile of RGB stars in the four ranges of metallicities for all the halos. The median profile is favored over the mean one since it limits the contribution from compact substructures, such as satellite galaxies. The resulting profiles are shown in Figure 9 and fitted with a single power-law profile in the range of projected distances of 23–120 kpc (9° at the distance of M31). The resulting power-law slope values are listed on Table 2. We purposefully avoid the region beyond 120 kpc since it was used to construct the contamination model of M13, leading to profiles that could be systematically biased at these radii, despite the correction mentioned in Section 2.3.

When considering the surface density profiles for the full metallicity range, all mocks apart from H27 have profile slopes relatively similar to M31. However, most of the mocks are more centrally populated than M31, with higher surface densities, up to a factor ~ 4 , until ~ 75 kpc. Beyond ~ 100 kpc, most of the mocks are less densely populated than the observed M31 halo. It is also interesting to note here that

the profile of H21 shows a clear drop at ~ 75 kpc, also visible in the metal-rich range. This is caused by the two major accretions that are ongoing and largely dominate the rest of the halo population up to ~ 75 kpc.

By comparing the density profiles in the different metallicity intervals, it is clear that the global surface density profiles of the stellar halos are dominated by metal-rich and intermediate-metallicity stars, both in M31 and in the mocks. However, the difference in density between the observations and the simulations are most prominent for metal-rich stars ($-0.7 \leq [\text{Fe}/\text{H}] < 0.0$), for which the density of stars in the mocks is higher than the observations at all the radii. In this metallicity range, H16 is the exception, with surface density profiles that are similar to the ones observed for M31 beyond 30 kpc. At intermediate metallicities ($-1.7 \leq [\text{Fe}/\text{H}] < -0.7$), the differences are smaller than in the metal-rich regime, with the profile of the mocks being, on average, $\simeq 1.5$ more populated than that of M31 at intermediate metallicities, against $\simeq 2.2$ at metal-rich metallicities. Overall, the mocks are still more densely populated than observed until a projected radius of ~ 100 kpc. Beyond that distance, the density in most of the simulated halos drops below the observed density.

For the metal-poor star selection ($-2.5 \leq [\text{Fe}/\text{H}] < -1.7$), half of the mocks are ~ 1.7 times more populated than M31. The other half have a similar inner density to what is observed in M31. However, in this metallicity range, the surface density profile of the mocks is systematically steeper than observed, and for most of them, the distant halo (> 60 kpc) shows a clear deficit of metal-poor stars compared to the observation. H27 is

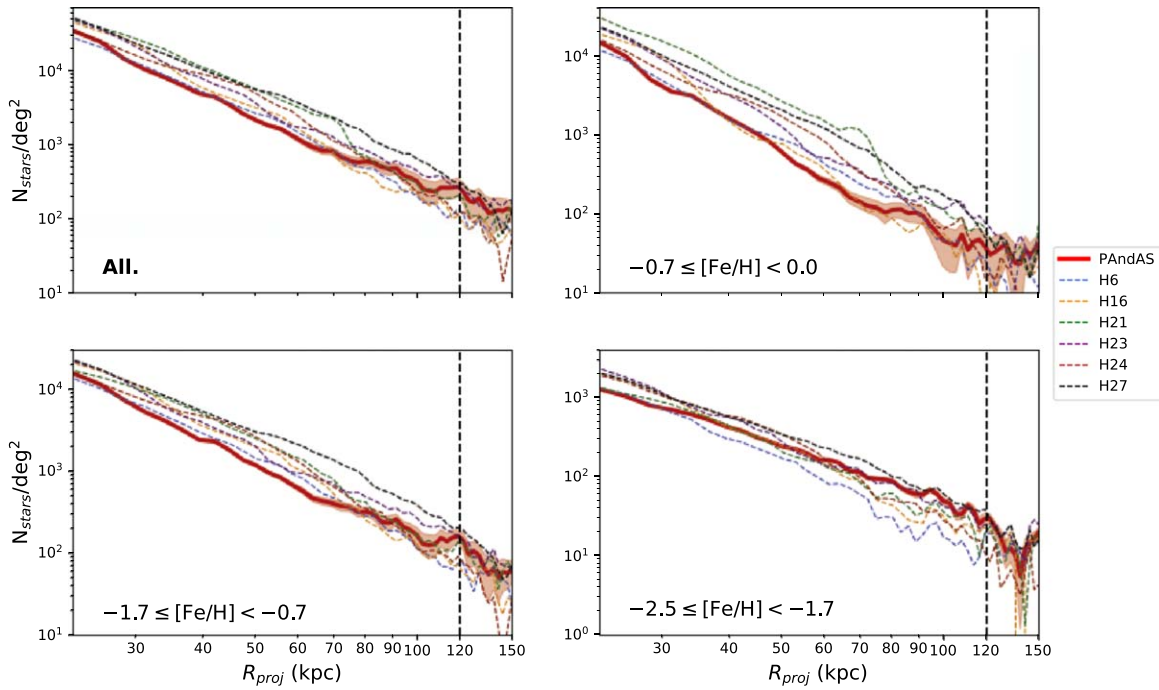


Figure 9. Profile of the median surface density of the different halos in four ranges of metallicities: all metallicities, metal poor ($-2.5 \leq [\text{Fe}/\text{H}] < -1.7$), intermediate metallicity ($-1.7 \leq [\text{Fe}/\text{H}] < -0.7$), and metal rich ($-0.7 \leq [\text{Fe}/\text{H}] < 0.0$). The scales are similar for all panels, except for the metal-poor range, where the scale is divided by 10 for easy comparison with the other panels. In each panel, the red shaded area shows the variation in the observed surface profile that we derived in PAndAS accounting for the 1σ uncertainties on the correction factor applied to the M13 decontamination model, as found by I14 ($\mathcal{F} = 0.93 \pm 0.03$).

an exception compared to the other simulations since, for all metallicity intervals, the surface density of the mock is always higher than observed in M31 at any radius.

Given these observations, we can conclude that the halos of the simulated galaxies are, in general, more populated than the M31 halo, especially in the central region. In particular, the simulations are more populated by metal-rich stars than is observed around M31 at any radius. Moreover, the metal-poor stars in the mocks are more centrally concentrated and are less extended than observed by PAndAS around M31.

3.2. Precision of the Photometric Metallicities

In the previous section, the analysis is partly based on photometric metallicities derived using the Dartmouth isochrones, while the Padova isochrones are used to give an absolute magnitude to the “stars” in the simulations. In this section, we show that the choice of the set of isochrones has a negligible impact on our analysis and that this choice does not alter the conclusions that we draw from the analysis of the mocks.

For the comparison between the mocks and the observations, the Dartmouth isochrones were preferred over the Padova ones for mainly two reasons. The first is that, historically, the previous works done by the PAndAS collaboration used mainly the Dartmouth isochrones (e.g., McConnachie et al. 2009; Collins et al. 2010; Martin et al. 2013; Ibata et al. 2014; McConnachie et al. 2018). Thus, using the same set of isochrones facilitates a direct comparison with these previous analyses. The second is that using a set of isochrones that is independent from the set used to generate the stellar particles in the simulations removes any suspicion of possible biases in the analysis.

Figure 10 displays the median surface density profile for mock H23 in the four previous ranges of metallicities for the

case where the photometric metallicities are derived using the 10 Gyr old Dartmouth isochrones with $[\alpha/\text{Fe}] = +0.2$ (purple lines), like in the previous section, and with the Padova isochrones (blue dashed lines) corresponding to a population of 8 Gyr, the median age of the “stars” in this specific halo. This figure also includes the median surface profile of the uncontaminated “stars” (i.e., it does not include the “stars” from the foreground model), using the metallicities directly provided by the simulations (red dots). These different profiles are very similar in all metallicity ranges, except for the metal-poor regime. In this range, the two profiles obtained using the photometric metallicities are still very similar to each other but are systematically higher than the “true” profile by a factor ~ 2 . This is explained by the fact that isochrones of metal-poor stars are very close to each other in a CMD. Thus, with the photometric uncertainties and the intrinsic scatter of a stellar population, metal-poor stars overlap each other in a CMD, leading to a lower accuracy of the photometric metallicities in that regime. Actually, it seems that a photometric metallicity range of $-2.5 \leq [\text{Fe}/\text{H}]_{\text{photo}} < -1.7$ corresponds to a range of $-2.5 \leq [\text{Fe}/\text{H}]_{\text{sim}} < -1.5$ for the metallicities of the mocks, for which the profiles are represented by the green dots on the lower right panel.

However, it is important to keep in mind that our analysis is based on relative metallicities, with four broad ranges. Therefore, whether a given “star” has $[\text{Fe}/\text{H}] = -1.7$ or $[\text{Fe}/\text{H}] = -1.5$ does not significantly impact the fact that metal-poor stars are less numerous than the intermediate or metal-rich ones. In addition, we perform the exact same analysis on the PAndAS observations and on the AURIGA mocks. The conclusions that we draw are therefore based on a relative comparison of these results and do not rely on the absolute metallicity values.

Table 2

Slopes of the Power-law Fit to the Projected Density Profile for the Different Halo in the Four Different Ranges of Metallicity

	[Fe/H] range	All	Accreted
PAndAS	All	$-3.39^{+0.01}$...
	$-2.5 \leq [\text{Fe}/\text{H}] < -1.7$	$-2.23^{+0.06}$...
	$-1.7 \leq [\text{Fe}/\text{H}] < -0.7$	$-3.24^{+0.01}$...
	$-0.7 \leq [\text{Fe}/\text{H}] < 0.0$	$-4.04^{+0.02}$...
H6	All	$-3.22^{+0.01}$	$-3.13^{+0.01}$
	$-2.5 \leq [\text{Fe}/\text{H}] < -1.7$	$-2.95^{+0.09}$	$-2.88^{+0.10}$
	$-1.7 \leq [\text{Fe}/\text{H}] < -0.7$	$-3.07^{+0.01}$	$-3.06^{+0.01}$
	$-0.7 \leq [\text{Fe}/\text{H}] < 0.0$	$-3.57^{+0.02}$	$-3.39^{+0.02}$
H16	All	$-3.67^{+0.01}$	$-3.49^{+0.01}$
	$-2.5 \leq [\text{Fe}/\text{H}] < -1.7$	$-2.95^{+0.06}$	$-2.84^{+0.07}$
	$-1.7 \leq [\text{Fe}/\text{H}] < -0.7$	$-3.38^{+0.01}$	$-3.30^{+0.01}$
	$-0.7 \leq [\text{Fe}/\text{H}] < 0.0$	$-4.39^{+0.02}$	$-4.09^{+0.02}$
H21	All	$-3.20^{+0.01}$	$-2.98^{+0.01}$
	$-2.5 \leq [\text{Fe}/\text{H}] < -1.7$	$-2.57^{+0.06}$	$-2.51^{+0.07}$
	$-1.7 \leq [\text{Fe}/\text{H}] < -0.7$	$-2.95^{+0.01}$	$-2.88^{+0.01}$
	$-0.7 \leq [\text{Fe}/\text{H}] < 0.0$	$-3.48^{+0.01}$	$-3.19^{+0.01}$
H23	All	$-3.35^{+0.01}$	$-3.26^{+0.01}$
	$-2.5 \leq [\text{Fe}/\text{H}] < -1.7$	$-2.76^{+0.05}$	$-2.73^{+0.07}$
	$-1.7 \leq [\text{Fe}/\text{H}] < -0.7$	$-3.17^{+0.01}$	$-3.21^{+0.01}$
	$-0.7 \leq [\text{Fe}/\text{H}] < 0.0$	$-3.76^{+0.01}$	$-3.50^{+0.01}$
H24	All	$-3.05^{+0.01}$	$-2.80^{+0.01}$
	$-2.5 \leq [\text{Fe}/\text{H}] < -1.7$	$-2.78^{+0.05}$	$-2.71^{+0.06}$
	$-1.7 \leq [\text{Fe}/\text{H}] < -0.7$	$-2.91^{+0.01}$	$-2.74^{+0.01}$
	$-0.7 \leq [\text{Fe}/\text{H}] < 0.0$	$-3.38^{+0.01}$	$-3.03^{+0.01}$
H27	All	$-2.91^{+0.01}$	$-2.70^{+0.01}$
	$-2.5 \leq [\text{Fe}/\text{H}] < -1.7$	$-2.40^{+0.04}$	$-2.37^{+0.05}$
	$-1.7 \leq [\text{Fe}/\text{H}] < -0.7$	$-2.69^{+0.01}$	$-2.61^{+0.01}$
	$-0.7 \leq [\text{Fe}/\text{H}] < 0.0$	$-3.28^{+0.01}$	$-2.96^{+0.02}$

Note. For the mocks, the middle column lists the slopes of the overall populations in the halos (accreted and in situ), and the right column lists the slopes of the accreted population only.

3.3. Comparison of the In Situ/Accreted Components

It has been shown with semianalytical and hydrodynamical simulations of L^* galaxies that their stellar halo can be decomposed into two populations with very different origins, the in situ and the accreted components (also referred to as the ex situ component in the literature, e.g., Bullock & Johnston 2005; Cooper et al. 2010; Font et al. 2011; Purcell et al. 2011; Pillepich et al. 2014; Monachesi et al. 2019). The accreted component is made of stars initially hosted by dwarf galaxies and globular clusters that have since been disrupted by tidal effects and that have deposited their stars into the main halo. The in situ component is made of stars from the galactic disk that have been ejected into the halo due to the dynamical heating produced by the interactions with subhalos, galaxies, or massive molecular clouds and by stars that were born in the halo of the protogalaxy.

To differentiate the contribution of these two populations in the mocks, we redo the analysis done in Section 3.1 but select only the *accreted* “stars” in the mocks. Mock “stars” are classified as accreted or in situ according to their parent stellar particle, following the definition presented in Section 2.3 of

Monachesi et al. (2019). In this definition, stellar particles bound to the host galaxy at birth are considered as in situ stars.

As visible in Figure 11, considering only the accreted stars, the surface density profiles of the simulated galaxies are generally in better agreement with the profile of M31. In all the different metallicity intervals, the shapes of the surface density profiles of the accreted “stars” alone are closer to the profiles observed in M31 than if we consider the full stellar halo populations (in situ and ex situ), especially for simulation H23. The strongest deviations between the surface density profiles of the accreted populations and the overall population (accreted+in situ) are visible with the most metal-rich stars, followed by the intermediate metallicities range. This is not surprising since the accreted stars are on average more metal poor than the stars formed in situ, as seen in the MW (i.e., Carollo et al. 2007, 2010; Belokurov et al. 2020) and in different suites of cosmological simulations (i.e., Purcell et al. 2011; Pillepich et al. 2014; Monachesi et al. 2019).

As visible on Figure 12, the accreted stars make up the majority of the stellar halos (at least beyond 23 kpc) for all metallicity ranges, as noted by Monachesi et al. (2019). However, the radial profile of the fraction of accreted stars is very different in the different metallicity intervals. Indeed, for metal-rich stars, $\simeq 50\%$ – 65% of them are accreted at 23 kpc, and this fraction increases slowly to reach $\simeq 80\%$ – 90% at 120 kpc, while for the intermediate and metal-poor stars, the fraction of accreted stars is almost constant across all radii, accounting for $\sim 70\%$ – 80% and $\sim 90\%$ of the stars in these respective ranges. This confirms that the in situ component is mostly composed of centrally concentrated metal-rich stars, similar to what is observed in the MW or in different simulations (i.e., Purcell et al. 2010; Cooper et al. 2015; Pillepich et al. 2015; Monachesi et al. 2019; Sanderson et al. 2018; Belokurov et al. 2020). Note that the ratio shown in this figure corresponds to the ratio of stars from the smooth halo component (i.e., not in substructures). Taking into account the overall population, the ratios are similar for all mocks, except for H16 and H24 (referred to as the full population of Figure 12), for which the fraction of in situ stars is significantly higher, up to ~ 35 and 25 kpc, respectively. This is caused by the presence of very extended galactic disk in these simulations, as noticed by Grand et al. (2018b). However, the goal of this section is to compare the fraction of accreted stars in the halos of the different simulated galaxies, so we purposefully do not take into account the stars in the disk or those present in substructures.

By comparing the surface density profile observed in M31 in the different ranges of metallicity and the profiles of the accreted and of the accreted+in situ populations in the mocks, it seems that the mocks show a density of in situ stars that is about twice as large as that observed in M31. Indeed, by reducing by roughly a factor 2 the contribution of the in situ component, the surface density profiles will be closer to the profile observed in PAndAS, especially for H16, H21, and H23. This is most visible when comparing the accreted and overall metal-rich surface density profile in the inner halo (< 30 kpc). This conclusion is in agreement with the observation of Monachesi et al. (2019) using the AURIGA simulations, and similar results have recently been found by Merritt et al. (2020) comparing the ILLUSTRIS TNG100 simulations (Springel et al. 2018; Pillepich et al. 2018b) to the Dragonfly Nearby Galaxies Survey (Merritt et al. 2016).

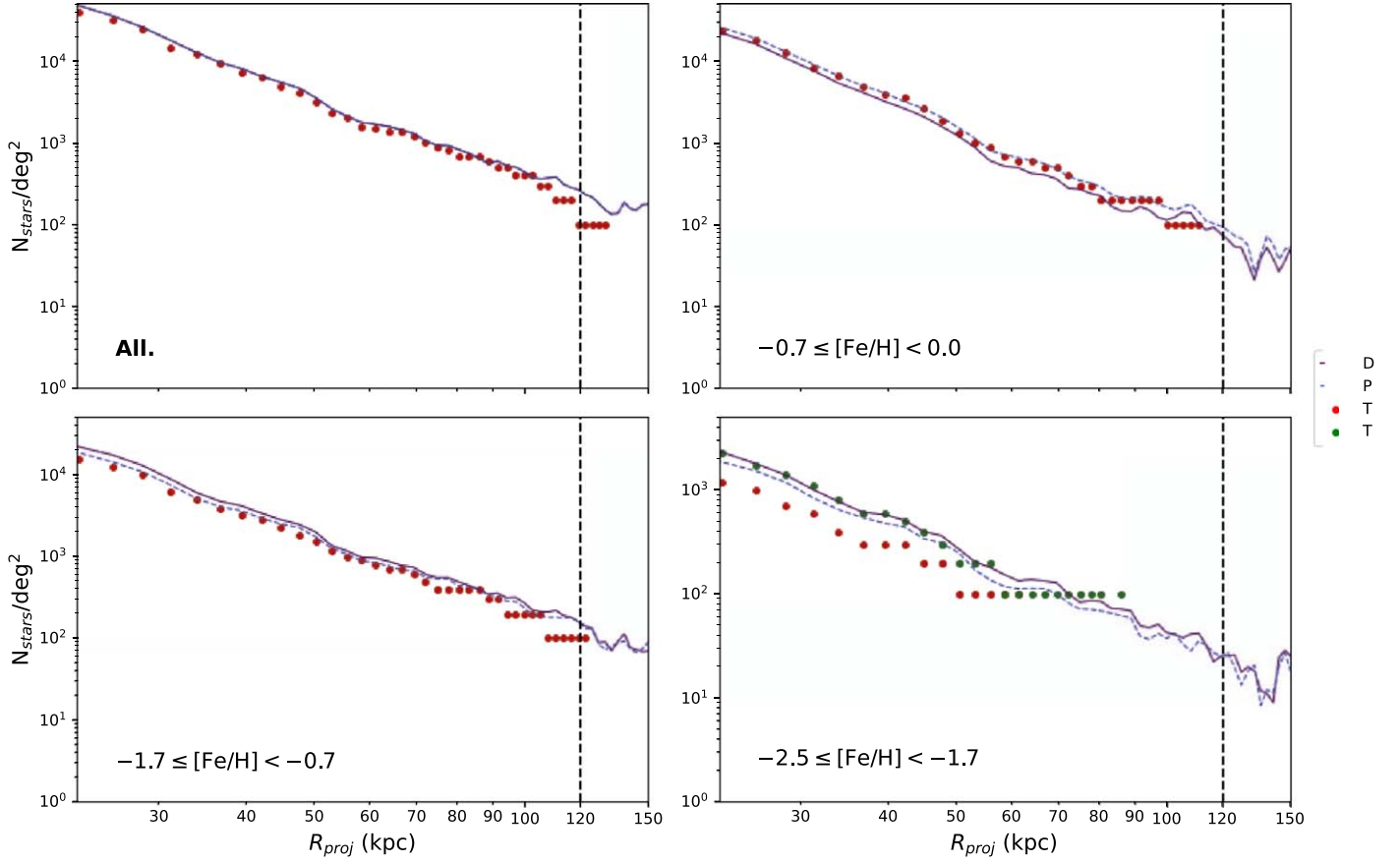


Figure 10. Comparison of the median surface density profile of mock H23 with the photometric metallicities determined from the Dartmouth isochrones used for Figure 9 (purple), the 8 Gyr old Padova isochrones (blue), and the original “true” metallicities of the simulations’ stellar particles (red dots). The profile shown by the green dots corresponds to the profile of the stellar particles in the range $-2.5 \leq [\text{Fe}/\text{H}]_{\text{sim}} < -1.5$. The plateau at $100 \text{ stars deg}^{-2}$ for this profile is due to the fact that we take the median of the number of actual stellar particles per pixels of $0.1^\circ \times 0.1^\circ$, which are by definition an integer. The profiles determined by the photometric metallicities do not have this behavior because the number of foreground stars per pixel is removed statistically.

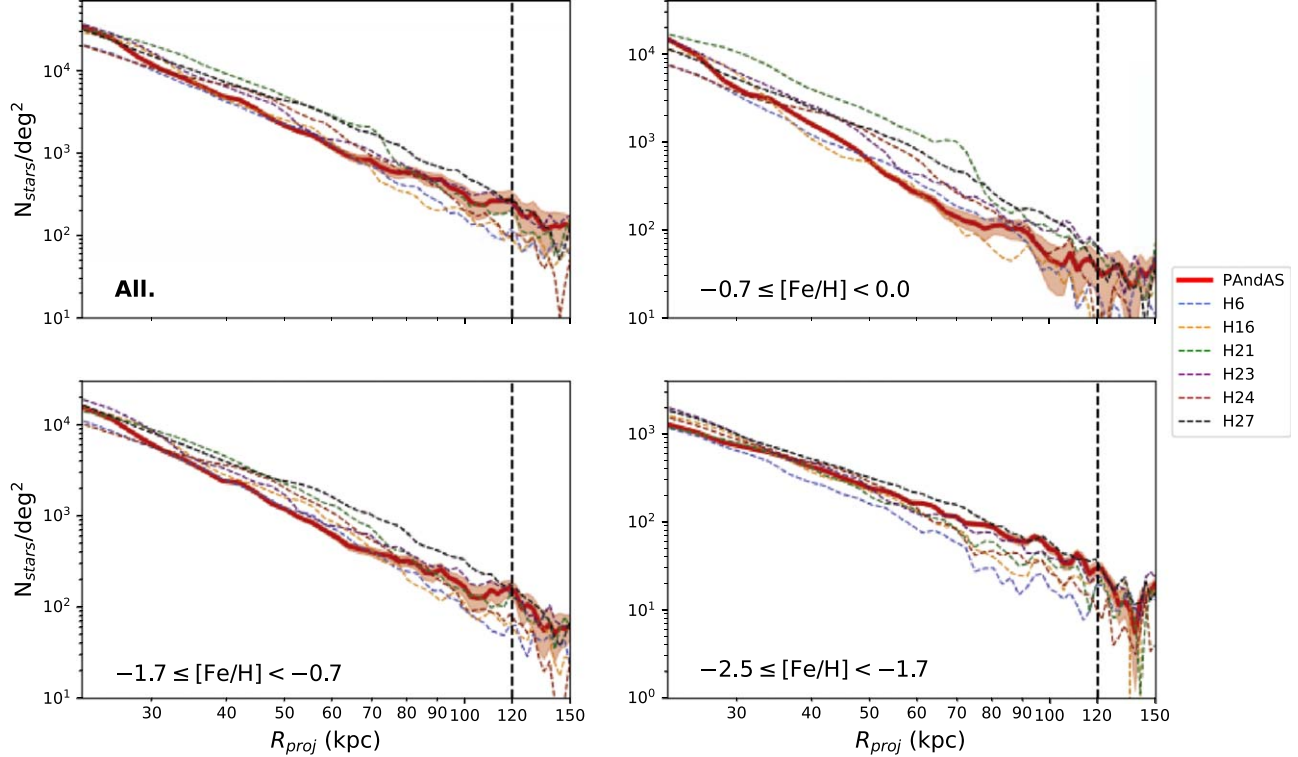


Figure 11. Median surface profile, similar to Figure 9 but just for the accreted “stars.”

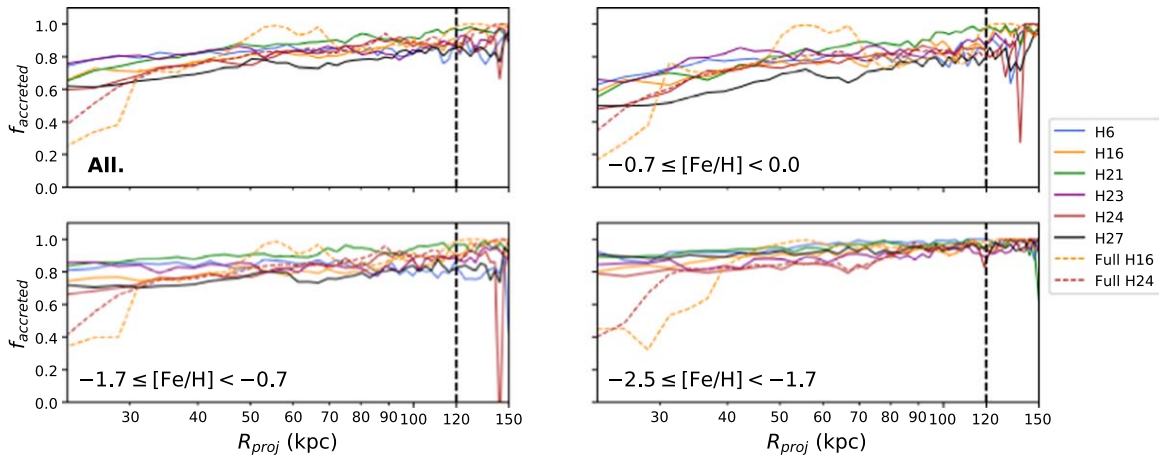


Figure 12. Projected profile of the fraction of accreted “stars” in the mocks.

However, even by reducing the number of the stars formed in situ by a factor of 2, most of the mocks have halos that are overall more populated than M31’s, and their profiles present a great diversity. Therefore, outside the inner halo, where the majority of the in situ stars are located, this scenario is not the only explanation for the observed difference between M31 and the simulated galaxies. Because galaxies like H23 have a very similar profile to the observed one, this overpopulation of most of the galaxies and the diversity of profiles is possibly driven by the stochastic process on the merger histories of each galaxies.

From all the simulated halos analyzed here, the halo H23 is the most representative of the halo of M31. Not only does this halo have a higher number of dwarf galaxies similar to AndXXI or AndXXIII than the other simulations, but its surface density profile is the closest to the profile constructed from PAndAS in all metallicity ranges, especially for the accreted “stars.” However, as for all the simulated halos, the in situ component is ~ 2 times more populated compared with M31.

3.4. Analysis of the Simulations

As we have seen in the previous sections, the AURIGA-PANDAS mocks are, overall, reasonable approximations of the stellar halo of M31 and its diversity of structures. Qualitatively, the inner halos of the simulations are very similar to the inner halo of M31. Moreover, most of the halos are populated by a number of intermediate-metallicity dwarf galaxies that are similar to what we see around M31. They cover a broad range of luminosities, from galaxies with a size and a metallicity similar to NGC 147 and NGC 185 (corresponding to a stellar mass of $\simeq 5 \times 10^7 M_\odot$), to galaxies similar to AndXXI and AndXXIII (corresponding to stellar masses of $\sim 1 \times 10^6 M_\odot$). However, in certain aspects, the mocks differ from what is observed around M31. Although the goal of the paper is not to analyze the reasons for the (overall) small differences between the simulations and the observations, we mention hereafter a few points that could be explored in future works.

For instance, we can notice that the simulated stellar halos are on average more metal rich than observed in M31. As explained in the previous section, this is likely a consequence of an overrepresentation of the in situ stars in the simulations, these stars being more metal rich than the accreted ones. As mentioned by Monachesi et al. (2019), this overpopulation of

stars formed in situ could be a consequence of the disks of the host galaxy, which, in the AURIGA simulations, are typically larger than observed for M31 (see Grand et al. 2017 and Monachesi et al. 2019 for a discussion on the size of the disk). However, even with considering only the accreted stars, the simulations tend to be more metal rich than observed for all radii, which might be a consequence of the properties of the accreted dwarf galaxies that formed these structures but also, potentially, an underquenching of the star formation feedback in the last billion years, especially in the most massive satellites, as proposed by Monachesi et al. (2019). Moreover, the surface density profile is lower than the majority of the simulation out to ~ 70 – 80 kpc, even when considering only the accreted population. This is surprising given that the simulated galaxies have a similar or lower stellar mass than observed in M31. This could also be a consequence of the underquenching of the star formation feedback, of the way in which the stellar particles are split into “stars” in the simulations, and/or of the IMF choice or the stellar populations models.

One of the other noticeable differences between the simulations and the observations of M31 is the relatively small number of metal-poor dwarf galaxies present around the simulated galaxies, compared to the numbers observed in PAndAS. Galaxies similar to And XIV or And XII are not visible at all in the simulations. Simpson et al. (2018) show that the number of dwarf galaxy AURIGA simulations converge down to 2 – $3 \times 10^4 M_\odot$ for different resolutions. However, it is possible that the small number of metal-poor dwarf galaxies in the simulation might be due to an incomplete sampling of the total phase space of these galaxies, despite the fact that, with the ICC method, “stars” conserve the phase space of the initial stellar particle. It will be interesting to quantify these apparent differences by searching and characterizing the dwarf galaxies in the mocks in the same way they were found and characterized in PAndAS.

4. Conclusions

We presented a pipeline to transform the output of *top-down* AURIGA simulations as if they were observed by the CFHT telescope with the MegaCam instrument as part of the PAndAS, a survey that observed the stellar halos of M31 and M33 out to a projected radius of 150 kpc and 50 kpc, respectively. Such a method is a basic requirement if we are to directly compare state-of-the-art cosmological simulations to

exquisite observations of the Local Group in order to better constrain and refine the galaxy formation processes used in these simulations. The transformation of the simulations into “observable” mocks allows the use of the exact same tools, with all the assumptions that they include, on the observations and on the simulations.

We have performed a qualitative comparison between the simulated stellar halo of six L^* galaxy mocks from the AURIGA simulations and the observed stellar halo of the Andromeda galaxy. We find that overall, the mock halos are similar to the halo of M31, with the presence of numerous structures similar to those observed. This is especially true in the inner halo (<30 kpc), where mocks and observations have very comparable structure and metallicity. Moreover, most of the mocks present the sign of a recent massive accretion, similar to the Giant Stream observed in M31. Some of the mocks have a very similar density profile of accreted stars compared to observations, the variations between the different simulated galaxies being a consequence of the stochasticity of the hierarchical galaxy formation process. However, although it is challenging to conclude definitively with only six simulations to compare to, the mocks present some systematic differences with the observations. We see that the in situ populations are overly represented by a factor $\simeq 2$, that the faintest dwarf galaxies visible in PAndAS are absent from the mocks, and that metal-poor structures like the NGC 147 stream are also absent. We also find that the metal-poor component of the different mocks is more concentrated than that observed in M31. We interpret these differences as a consequence of underquenching by stellar feedback, increasing the average metallicity of the simulated galaxies, but also as a limitation of the resolution of the stellar particles in the simulations. It warns us against currently pushing the simulations into the very faint regimes probed by the PAndAS observations.

In future work, we will quantify the level of structure present in the simulations using the method presented in McConnachie et al. (2018), and we will compare it to the level of the different structures seen in PAndAS. Moreover, the distribution and the characteristics of the simulated satellite galaxies will be studied

more in detail to compare to the distribution of the satellites of M31 and determine how faint the comparison is possible.

The AURIGA2PANDAS pipeline presented here and one realization of each mock are publicly available online (<https://github.com/GFThomas/Auriga2PAndAS>). The different columns of the catalogs are described in the Appendix. We encourage the application of this pipeline to other state-of-the-art cosmological simulations such as EAGLE (the EAGLE team 2017), Illustris TNG (Pillepich et al. 2018a), Latte/FIRE (Wetzel et al. 2016; Hopkins et al. 2018), or ARTEMIS (Font et al. 2020) to compare the results obtained from these simulations with the observations and between each other. Only then will we be able to efficiently constrain the different parameters of the simulations, especially those relating to the complicated baryonic physics.

G.T. acknowledges support from the Agencia Estatal de Investigación (AEI) of the Ministerio de Ciencia e Innovación (MCINN) under grant with reference (FJC2018-037323-I).

F.A.G. acknowledges financial support from CONICYT through the project FONDECYT Regular Nr. 1181264 and funding from the Max Planck Society through a Partner Group grant. Based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada–France–Hawaii Telescope (CFHT), which is operated by the National Research Council (NRC) of Canada, the Institut National des Science de l’Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii.

Appendix

Description of the Online Catalog

Table 3 provides a description of each column in the online catalog available at <https://wwwmpa.mpa-garching.mpg.de/auriga/gaiamock.html#pandas>. There are six high-resolution simulations available in FITS files. Each contains the data products of a single realization of an AuriPAndAS mock.

Table 3
Catalog Column Descriptions

No.	Column name	Description
0	ID	Identifiant of the star from the Auriga simulations. -99 if from the background model
1	RA	R.A. (deg)
2	Dec	Decl. (deg)
3	xki	ξ tangential coordinates centered on M31 (deg)
4	eta	η tangential coordinates centered on M31 (deg)
5	rhelio	Heliocentric distance of each star (kpc). 0 if from the background model
6	pmra	Proper motion in the R.A. direction (mas yr ⁻¹). 0 if from the background model
7	pmdec	Proper motion in the decl. direction (mas yr ⁻¹). 0 if from the background model
8	Vrad	Heliocentric radial velocity (km s ⁻¹). 0 if from the background model
9	g	Apparent magnitude in the <i>g</i> -band
10	dg	Uncertainty on the <i>g</i> -band magnitude
11	g0	Deredded <i>g</i> -band magnitude
12	i	Apparent magnitude in the <i>i</i> -band
13	di	Uncertainty on the <i>i</i> -band magnitude
14	i0	Deredded <i>i</i> -band magnitude
15	EBV	Extinction from (Schlegel et al. 1998)
16	nb_field	PAndAS field in which is the star (from 1 to 406)
17	x	Galactocentric cartesian coordinate of the star from Auriga (kpc). 0 if from the background model
18	y	Galactocentric cartesian coordinate of the star from Auriga (kpc). 0 if from the background model
19	z	Galactocentric cartesian coordinate of the star from Auriga (kpc). 0 if from the background model
20	Mg	Absolute magnitude in the <i>g</i> -band from the Auriga simulation. 0 if from the background model
21	Mi	Absolute magnitude in the <i>i</i> -band from the Auriga simulation. 0 if from the background model
22	feh_sim	Metallicity of the stars from the Auriga simulation. 0 if from the background model
23	Acc	Flag the origin of the simulated stars (-1 = formed in situ, 0 = accreted, 1 = formed in a satellite after infall). 0 for the stars from the background model.

Note. Table 3 describes the contents of the catalogs available on the AuriPAndAS website.

ORCID iDs

Guillaume F. Thomas  <https://orcid.org/0000-0002-2468-5521>

Nicolas F. Martin  <https://orcid.org/0000-0002-1349-202X>

Rodrigo A. Ibata  <https://orcid.org/0000-0002-3292-9709>

Alan W. McConnachie  <https://orcid.org/0000-0003-4666-6564>


Carlos Frenk  <https://orcid.org/0000-0002-2338-716X>

Facundo A. Gómez  <https://orcid.org/0000-0002-1947-333X>

Robert J. J. Grand  <https://orcid.org/0000-0001-9667-1340>

Dougal Mackey  <https://orcid.org/0000-0002-6529-8093>

Federico Marinacci  <https://orcid.org/0000-0003-3816-7028>

Rüdiger Pakmor  <https://orcid.org/0000-0003-3308-2420>

References

- Amorisco, N. C. 2017, *MNRAS*, **469**, L48
- Bailin, J., Bell, E. F., Valluri, M., et al. 2014, *ApJ*, **783**, 95
- Bell, E. F., Xue, X. X., Rix, H.-W., Ruhland, C., & Hogg, D. W. 2010, *AJ*, **140**, 1850
- Bell, E. F., Zucker, D. B., Belokurov, V., et al. 2008, *ApJ*, **680**, 295
- Belokurov, V., Sanders, J. L., Fattahi, A., et al. 2020, *MNRAS*, **494**, 3880
- Boylan-Kolchin, M., Ma, C.-P., & Quataert, E. 2008, *MNRAS*, **383**, 93
- Bressan, A., Marigo, P., Girardi, L., et al. 2012, *MNRAS*, **427**, 127
- Brown, T. M., Smith, E., Ferguson, H. C., et al. 2006, *ApJ*, **652**, 323
- Brown, T. M., Smith, E., Ferguson, H. C., et al. 2007, *ApJ*, **658**, L95
- Bullock, J. S., & Johnston, K. V. 2005, *ApJ*, **635**, 931
- Carollo, D., Beers, T. C., Chiba, M., et al. 2010, *ApJ*, **712**, 692
- Carollo, D., Beers, T. C., Lee, Y. S., et al. 2007, *Natur*, **450**, 1020
- Chemin, L., Carignan, C., & Foster, T. 2009, *ApJ*, **705**, 1395
- Collins, M. L. M., Chapman, S. C., Irwin, M. J., et al. 2010, *MNRAS*, **407**, 2411
- Collins, M. L. M., Chapman, S. C., Rich, R. M., et al. 2013, *ApJ*, **768**, 172
- Conn, A. R., Ibata, R. A., Lewis, G. F., et al. 2012, *ApJ*, **758**, 11
- Conn, A. R., Lewis, G. F., Ibata, R. A., et al. 2011, *ApJ*, **740**, 69
- Cooper, A. P., Cole, S., Frenk, C. S., et al. 2010, *MNRAS*, **406**, 744
- Cooper, A. P., Parry, O. H., Lowing, B., Cole, S., & Frenk, C. 2015, *MNRAS*, **454**, 3185
- Dalcanton, J. J., Williams, B. F., Lang, D., et al. 2012, *ApJS*, **200**, 18
- Dalcanton, J. J., Williams, B. F., Seth, A. C., et al. 2009, *ApJS*, **183**, 67
- De Lucia, G., & Helmi, A. 2008, *MNRAS*, **391**, 14
- Deason, A. J., Belokurov, V., & Evans, N. W. 2011, *MNRAS*, **416**, 2903
- Deason, A. J., Belokurov, V., Evans, N. W., et al. 2012, *MNRAS*, **425**, 2840
- Deason, A. J., Belokurov, V., Evans, N. W., & Johnston, K. V. 2013, *ApJ*, **763**, 113
- Deason, A. J., Belokurov, V., & Sanders, J. L. 2019, *MNRAS*, **490**, 3426
- Deason, A. J., Mao, Y.-Y., & Wechsler, R. H. 2016, *ApJ*, **821**, 5
- Dotter, A., Chaboyer, B., Jevremović, D., et al. 2008, *ApJS*, **178**, 89
- D'Souza, R., & Bell, E. 2018, *MNRAS*, **474**, 5300
- Duc, P.-A., Cuillandre, J.-C., Karabal, E., et al. 2015, *MNRAS*, **446**, 120
- Eilers, A.-C., Hogg, D. W., Rix, H.-W., & Ness, M. K. 2019, *ApJ*, **871**, 120
- Fardal, M. A., Weinberg, M. D., Babul, A., et al. 2013, *MNRAS*, **434**, 2779
- 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
- Font, A. S., McCarthy, I. G., Poole-Mckenzie, R., et al. 2020, *MNRAS*, **498**, 1765
- Fukushima, T., Chiba, M., Homma, D., et al. 2018, *PASJ*, **70**, 69
- Fukushima, T., Chiba, M., Tanaka, M., et al. 2019, *PASJ*, **71**, 72
- Gilbert, K. M., Kalirai, J. S., Guhathakurta, P., et al. 2014, *ApJ*, **796**, 76
- Gómez, F. A., Grand, R. J. J., Monachesi, A., et al. 2017, *MNRAS*, **472**, 3722
- Gómez, F. A., Minchev, I., O'Shea, B. W., et al. 2012, *MNRAS*, **423**, 3727
- Gómez, F. A., Minchev, I., O'Shea, B. W., et al. 2013, *MNRAS*, **429**, 159
- Grand, R. J. J., Bustamante, S., Gómez, F. A., et al. 2018b, *MNRAS*, **474**, 3629
- Grand, R. J. J., Gómez, F. A., Marinacci, F., et al. 2017, *MNRAS*, **467**, 179
- Grand, R. J. J., Helly, J., Fattahi, A., et al. 2018a, *MNRAS*, **481**, 1726
- Gravity Collaboration, Abuter, R., Amorim, A., et al. 2018, *A&A*, **615**, L15
- Harmsen, B., Monachesi, A., Bell, E. F., et al. 2017, *MNRAS*, **466**, 1491
- Helmi, A. 2008, *A&ARv*, **15**, 145
- Helmi, A., & White, S. D. M. 1999, *MNRAS*, **307**, 495
- Hernitschek, N., Cohen, J. G., Rix, H.-W., et al. 2018, *ApJ*, **859**, 31
- Hopkins, P. F., Wetzel, A., Kereš, D., et al. 2018, *MNRAS*, **480**, 800
- Huxor, A. P., Mackey, A. D., Ferguson, A. M. N., et al. 2014, *MNRAS*, **442**, 2165

- Ibata, R., Irwin, M., Lewis, G., Ferguson, A. M. N., & Tanvir, N. 2001, *Natur*, **412**, 49
- Ibata, R., Mouhcine, M., & Rejkuba, M. 2009, *MNRAS*, **395**, 126
- Ibata, R. A., Lewis, G. F., McConnachie, A. W., et al. 2014, *ApJ*, **780**, 128
- Ibata, R. A., McConnachie, A., Cuillandre, J.-C., et al. 2017, *ApJ*, **848**, 129
- Irwin, M., & Lewis, J. 2001, *NewAR*, **45**, 105
- Ivezić, Ž., Sesar, B., Jurić, M., et al. 2008, *ApJ*, **684**, 287
- Johnston, K. V., Bullock, J. S., Sharma, S., et al. 2008, *ApJ*, **689**, 936
- Jurić, M., Ivezić, Ž., Brooks, A., et al. 2008, *ApJ*, **673**, 864
- Kafle, P. R., Sharma, S., Lewis, G. F., Robotham, A. S. G., & Driver, S. P. 2018, *MNRAS*, **475**, 4043
- Kilic, M., Bergeron, P., Dame, K., et al. 2019, *MNRAS*, **482**, 965
- Lewis, G. F., Braun, R., McConnachie, A. W., et al. 2013, *ApJ*, **763**, 4
- Lowing, B., Wang, W., Cooper, A., et al. 2015, *MNRAS*, **446**, 2274
- Mackey, A. D., Huxor, A., Ferguson, A. M. N., et al. 2007, *ApJL*, **655**, L85
- Mackey, A. D., Huxor, A. P., Ferguson, A. M. N., et al. 2010, *ApJ*, **717**, L11
- Mackey, A. D., Huxor, A. P., Martin, N. F., et al. 2013, *ApJL*, **770**, L17
- Malin, D., & Hadley, B. 1997, *PASA*, **14**, 52
- Martin, N. F., Ibata, R. A., McConnachie, A. W., et al. 2013, *ApJ*, **776**, 80
- Martin, N. F., Weisz, D. R., Albers, S. M., et al. 2017, *ApJ*, **850**, 16
- Martínez-Delgado, D., Gabany, R. J., Crawford, K., et al. 2010, *AJ*, **140**, 962
- Martínez-Delgado, D., & Mediavilla, E. 2013, *Local Group Cosmology* (Cambridge: Cambridge Univ. Press)
- McConnachie, A. W. 2012, *AJ*, **144**, 4
- McConnachie, A. W., Ibata, R., Martin, N., et al. 2018, *ApJ*, **868**, 55
- McConnachie, A. W., Irwin, M. J., Ibata, R. A., et al. 2009, *Natur*, **461**, 66
- Merritt, A., Pillepich, A., van Dokkum, P., et al. 2020, *MNRAS*, **495**, 4570
- Merritt, A., van Dokkum, P., Abraham, R., & Zhang, J. 2016, *ApJ*, **830**, 62
- Metz, M., Kroupa, P., & Jerjen, H. 2007, *MNRAS*, **374**, 1125
- Mihos, J. C., Harding, P., Feldmeier, J., & Morrison, H. 2005, *ApJL*, **631**, L41
- Monachesi, A., Bell, E. F., Radburn-Smith, D. J., et al. 2016, *MNRAS*, **457**, 1419
- Monachesi, A., Gómez, F. A., Grand, R. J. J., et al. 2019, *MNRAS*, **485**, 2589
- Moore, B., Ghigna, S., Governato, F., et al. 1999, *ApJL*, **524**, L19
- Oman, K., Starkenburg, E., & Navarro, J. 2017, *Galax*, **5**, 33
- Pakmor, R., Springel, V., Bauer, A., et al. 2016, *MNRAS*, **455**, 1134
- Peñarrubia, J., Gómez, F. A., Besla, G., Erkal, D., & Ma, Y.-Z. 2016, *MNRAS*, **456**, L54
- Pillepich, A., Madau, P., & Mayer, L. 2015, *ApJ*, **799**, 184
- Pillepich, A., Nelson, D., Hernquist, L., et al. 2018b, *MNRAS*, **475**, 648
- Pillepich, A., Springel, V., Nelson, D., et al. 2018a, *MNRAS*, **473**, 4077
- Pillepich, A., Vogelsberger, M., Deason, A., et al. 2014, *MNRAS*, **444**, 237
- Purcell, C. W., Bullock, J. S., & Kazantzidis, S. 2010, *MNRAS*, **404**, 1711
- Purcell, C. W., Bullock, J. S., Tollerud, E. J., Rocha, M., & Chakrabarti, S. 2011, *Natur*, **477**, 301
- Purcell, C. W., Bullock, J. S., & Zentner, A. R. 2007, *ApJ*, **666**, 20
- Radburn-Smith, D. J., de Jong, R. S., Seth, A. C., et al. 2011, *ApJS*, **195**, 18
- Renda, A., Gibson, B. K., Mouhcine, M., et al. 2005, *MNRAS*, **363**, L16
- Richardson, J. C., Ferguson, A. M. N., Mackey, A. D., et al. 2009, *MNRAS*, **396**, 1842
- Robertson, B., Bullock, J. S., Font, A. S., Johnston, K. V., & Hernquist, L. 2005, *ApJ*, **632**, 872
- Sanderson, R. E., Garrison-Kimmel, S., Wetzel, A., et al. 2018, *ApJ*, **869**, 12
- Sawala, T., Frenk, C. S., Fattahi, A., et al. 2016, *MNRAS*, **457**, 1931
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *ApJ*, **500**, 525
- Schönrich, R., Binney, J., & Dehnen, W. 2010, *MNRAS*, **403**, 1829
- Searle, L., & Zinn, R. 1978, *ApJ*, **225**, 357
- Sesar, B., Jurić, M., & Ivezić, Ž. 2011, *ApJ*, **731**, 4
- Sick, J., Courteau, S., Cuillandre, J.-C., et al. 2015, in *IAU Symp. 311, Galaxy Masses as Constraints of Formation Models*, 311, ed. M. Cappellari & S. Courteau (Cambridge: Cambridge Univ. Press), 82
- Simpson, C. M., Grand, R. J. J., Gómez, F. A., et al. 2018, *MNRAS*, **478**, 548
- Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, *AJ*, **131**, 1163
- Slater, C. T., Harding, P., & Mihos, J. C. 2009, *PASP*, **121**, 1267
- Springel, V. 2010, *MNRAS*, **401**, 791
- Springel, V., Pakmor, R., Pillepich, A., et al. 2018, *MNRAS*, **475**, 676
- The EAGLE team 2017, arXiv:1706.09899
- Thomas, G. F., Annau, N., McConnachie, A., et al. 2019, *ApJ*, **886**, 10
- Thomas, G. F., McConnachie, A. W., Ibata, R. A., et al. 2018, *MNRAS*, **481**, 5223
- Tollerud, E. J., Beaton, R. L., Geha, M. C., et al. 2012, *ApJ*, **752**, 45
- Trujillo, I., & Fliri, J. 2016, *ApJ*, **823**, 123
- van der Marel, R. P., Fardal, M. A., Sohn, S. T., et al. 2019, *ApJ*, **872**, 24
- Veljanoski, J., Mackey, A. D., Ferguson, A. M. N., et al. 2014, *MNRAS*, **442**, 2929
- Wang, W., Han, J., Sonnenfeld, A., et al. 2019, *MNRAS*, **487**, 1580
- Watkins, L. L., Evans, N. W., & An, J. H. 2010, *MNRAS*, **406**, 264
- Watkins, L. L., Evans, N. W., Belokurov, V., et al. 2009, *MNRAS*, **398**, 1757
- Wetzel, A. R., Hopkins, P. F., Kim, J.-h., et al. 2016, *ApJL*, **827**, L23
- White, S. D. M., & Rees, M. J. 1978, *MNRAS*, **183**, 341
- Xue, X.-X., Rix, H.-W., Ma, Z., et al. 2015, *ApJ*, **809**, 144
- Xue, X.-X., Rix, H.-W., Yanny, B., et al. 2011, *ApJ*, **738**, 79