



Planes of Satellites around Simulated Disk Galaxies. I. Finding High-quality Planar Configurations from Positional Information and Their Comparison to MW/M31 Data

Isabel Santos-Santos^{1,2}, Rosa Domínguez-Tenreiro^{2,3}, Héctor Artal⁴, Susana E. Pedrosa^{2,5}, Lucas Bignone⁶ , Francisco Martínez-Serrano⁷ , M. Ángeles Gómez-Flechoso⁸ , Patricia B. Tissera^{6,9} , and Arturo Serna¹⁰

¹ Department of Physics and Astronomy, University of Victoria, Victoria, BC, V8P 5C2, Canada

² Departamento de Física Teórica, Universidad Autónoma de Madrid, E-28049 Cantoblanco, Madrid, Spain

³ Centro de Investigación Avanzada en Física Fundamental, Universidad Autónoma de Madrid, E-28049 Cantoblanco, Madrid, Spain

⁴ Intelligent Light, Rutherford, NJ 07070, USA

⁵ Instituto de Astronomía y Física del Espacio, CONICET-UBA, 1428, Buenos Aires, Argentina

⁶ Departamento de Ciencias Físicas, Universidad Andrés Bello, Santiago, Chile

⁷ Dassault Systemes S.E., E-28020 Madrid, Spain

⁸ Departamento de Física de la Tierra y Astrofísica, Univ. Complutense de Madrid, Madrid, Spain

⁹ Millennium Institute of Astronomy, Universidad Andrés Bello, Santiago, Chile

¹⁰ Departamento de Física Aplicada, Univ. Miguel Hernández, Elche, Spain

Received 2019 August 2; revised 2020 February 23; accepted 2020 March 10; published 2020 July 2

Abstract

We address the “plane of satellites problem” by studying planar configurations around two disk galaxies with no late major mergers, formed in zoom-in hydro-simulations. Due to the current lack of good-quality kinematic data for M31 satellites, we use only positional information. So far, positional analyses of simulations are unable to find planes as thin and populated as the observed ones. We follow a novel systematic and detailed plane searching technique to study the properties and quality of planes of satellites, in both simulations or real data. In particular, (i) we extend the four-galaxy-normal density plot method (Pawlowski et al. 2013) in a way designed to efficiently identify high-quality planes (i.e., thin and populated) without imposing extra constraints on their properties, and (ii), we apply it for the first time to simulations. Using zoom-in simulations allows us to mimic Milky Way/M31-like systems regarding the number of satellites involved as well as galactic disk effects. In both simulations, we find satellite planar configurations that are compatible, along given time intervals, with all of the spatial characteristics of observed planes identified using the same methodology. During most of these periods, planes are approximately perpendicular to the galactic disk. However, the fraction of co-orbiting satellites within them is, in general, low, suggesting time-varying satellite membership. We conclude that high-quality positional planes of satellites could be not infrequent in Λ CDM-formed disk galaxies with a quiet assembly history. Detecting kinematically coherent, time-persistent planes demands considering the full six-dimensional phase-space information of satellites.

Unified Astronomy Thesaurus concepts: Local Group (929); Dwarf galaxies (416); Galaxy dynamics (591); Galaxy planes (613); Galaxy kinematics (602)

1. Introduction

The so-called “small-scale problems in Λ CDM” refer to the discrepancies between the predictions for dwarf galaxies in the standard cosmological model as first revealed by dark matter (DM)-only cosmological simulations, and the actual observed properties dwarfs present (see Bullock & Boylan-Kolchin 2017, for a review). Among them, the planar configurations satellite galaxies show around their hosts (“Planes of satellites problem,” see Pawlowski 2018 for a review), observed in the local universe, have long been considered as one of the most challenging.

The high degree of anisotropy of Milky Way (MW) satellite positions, which appear forming a common plane approximately perpendicular to the Galactic disk, was noted several decades ago (Kunkel & Demers 1976; Lynden-Bell 1976) and first quantified by Kroupa et al. (2005) with the then-known 11 “classical”¹¹ satellites. With the addition of globular clusters, streams, and newly discovered fainter satellites (especially with SDSS; York et al. 2000), the anisotropy increased even more so, as these objects contributed further to a “vast polar structure

(VPOS)” around the MW (Pawlowski et al. 2012). Anisotropy among Andromeda’s (M31) satellites was first noted by Koch & Grebel (2006) and Metz et al. (2007), and later confirmed with a larger sample of satellites including those recently discovered with the Pan-Andromeda Archaeological Survey (PAndAS survey; Ibata et al. 2013). It was found that a majority of satellites are lopsided toward the MW’s side (McConnachie & Irwin 2006). In addition, M31 satellites do not define only one main planar structure, like in the MW; instead a thin plane of satellites including approximately half of the satellite sample was singled out (Conn et al. 2013; Ibata et al. 2013; Pawlowski et al. 2013, hereafter “Ibata-Conn-14” plane). Finally, new star clusters and dwarf galaxy candidates have been recently found in other nearby galactic systems in the local universe like CenA or the M101 group of galaxies. Studies suggest as well an anisotropic 3D-spatial distribution (Tully et al. 2015; Müller et al. 2017, 2018).

Additionally, proper motion data has revealed that a high fraction of MW satellites present orbital angular momentum vectors mostly perpendicular to the Galactic disk axis (Metz et al. 2008; Pawlowski & Kroupa 2013; Fritz et al. 2018; Gaia Collaboration et al. 2018). In particular, Fritz et al. (2018) used recent proper motion data from GAIA DR2 to calculate the

¹¹ Fornax, LMC, SMC, Draco, Leo II, Carina, Sculptor, Sextans I, Leo I, Sagittarius dSph, Ursa Minor.

orbital poles of objects orbiting within 420 kpc around the MW. According to their results, approximately $\sim 40\%$ (lower limit) of the confirmed MW satellites present orbital poles within an area of 10% of the sphere around the normal direction to the “VPOS,” which they define as a co-orbitation criteria. Also, claims for co-rotation of satellites in the M31 “Ibata-Conn-14” plane have been made based on the direction of radial (i.e., line of sight) velocities (Metz et al. 2007; Ibata et al. 2013) as no proper motion data is yet available. It has been shown, however, that these are generally not a representative measure of the true 3D-velocities (Buck et al. 2016).

Flattened spatial configurations of satellites in the MW and M31 have been well studied and quantified using information from only the three-dimensional positions of the satellites. In particular, these positional analyses have used sampling techniques like “bootstrapping” (Metz et al. 2007) or the “four-galaxy-normal density plot” method (hereafter 4GND plot; Pawlowski et al. 2013) to statistically show the existence of predominant planar configurations of satellite positions in the MW and M31. These planes are then accurately characterized by their normal vectors, axis ratios, and root-mean-square heights (Δ_{rms}), computed from the eigenvalues of the Tensor of Inertia (ToI) plane-fitting technique (Metz et al. 2007; Pawlowski et al. 2013, 2015). In particular, the 4GND plot method consists in fitting planes to subsamples of four different satellites, and projecting the normal vectors on the sphere, creating a density map. The over-density regions that appear as a consequence of the accumulation of normal vectors broadly point in the normal direction to a predominant planar configuration of satellites. Following this method, Pawlowski et al. (2013) detected the specific subsamples of satellites that mostly contribute to planar configurations in the MW and M31, defining the “VPOS-3” plane of satellites in the MW and the “GPoA” plane in M31.

In a recent paper, Santos-Santos et al. (2019, hereafter Paper I), have extended the 4GND plot method to allow for an identification, systematic cataloging, and more detailed quality study of the planar configurations of satellites in the MW and M31 systems. Rather than deriving a unique plane of satellites per over-density in the 4GND plot found with the previous method, the extension yields a *collection* of planes of satellites, each with an increasing number of members. In this way, it is possible to identify the highest-quality planes in terms of the ToI parameters and the number of satellites considered. New to previous findings, in Paper I, it was shown that *two* distinct planes of satellites are present in M31: the “GPoA” and another plane (labeled “M31-2-18” in Paper I). The two planar structures present very similar characteristics and are, interestingly, oriented perpendicularly to each other.

Since the very advent of these discoveries, theoretical studies have tried to assess the frequency of planar satellite configurations like those observed in cosmological simulations within the Λ CDM paradigm. These studies have mostly made use of large-volume DM-only simulations and pay attention to determining the significance¹² of the observed planes. In particular, Libeskind et al. (2009), Wang et al. (2013), Bahl & Baumgardt (2014), and Cautun et al. (2015) have analyzed different versions of the large-volume Millenium DM-only simulation (Springel et al. 2005; Boylan-Kolchin et al. 2009),

populating subhalos with galaxies following semi-analytic models. Using different methods for plane-identification, they find that planes of subhalos as thin and even thinner than the ones observed in the MW and in M31 are expected in Λ CDM. However, they acknowledge that these are not the mean case found, but only consistent with the tail of the predicted flattening distribution (see Pawlowski et al. 2014, for a different view). On the other hand, Cautun et al. (2015) show that not accounting for the “look-elsewhere effect” results in an important overestimation of the significance of planes in the MW and M31 systems of factors of 30 and 100, respectively. Indeed, according to that work, $\sim 10\%$ of MW-like mass halos in Λ CDM simulations have planes of satellites that are more prominent than those observed in the MW or M31 systems, presenting a large diversity when characterized by their thickness and number of satellites. On the other hand, a different approach has been followed by Buck et al. (2015). Instead of a large volume, they use several DMO zoom-in simulations to show that a thin plane, as the one around M31 with 15 satellites (Ibata et al. 2013), is not a challenge for the Λ CDM paradigm. However, neither the VPOS-3 (with 24 satellites) or VPOSall (27) planes in the MW are recovered in their analysis.

While some insight has been gained from collisionless N -body simulations, these experiments do not allow the formation of galactic disks. Having well-behaved massive MW-like simulated disks, however, may be critical to the planes of satellites issue (see Ahmed et al. 2017; and for a different perspective, Pawlowski et al. 2019). Indeed, the dynamical effect of a live disk potential on satellite planes could change the frequencies alluded to above, due, for example, to the torques that satellites suffer from galaxy disks—except when they are on planar or polar orbits, or far away from the disk plane (Danovich et al. 2015; Welker et al. 2018). Another effect is that galaxy disks preferentially destroy satellites on radial orbits that pass near them (Garrison-Kimmel et al. 2017; Sawala et al. 2017). Riley et al. (2019) show that massive disk potential effects cause the velocity anisotropy parameter, β , to decrease as compared to less massive disks. Including these relevant effects in planes of satellites studies is thus necessary for a fair comparison with results from the MW/M31 disk galaxy systems. Another relevant point is that there are $\gtrsim 30$ confirmed satellites in the MW and M31 (McConnachie 2012): a proper comparison demands as well the analysis of simulated disks surrounded by roughly as many resolved satellites.

Meeting all of the previous requirements is currently a situation not found in large-volume hydrodynamical simulations. For example, in their analysis of the EAGLE simulation, Shao et al. (2019) analyze planes of 11 satellites around central galaxies of any morphology and compare them to the MW “classical” plane—thus, the motivation to analyze zoom-in high-resolution hydrodynamical simulations in order to study planes of satellites.

A few studies exist using zoom-in hydro-simulations. In Gillet et al. 2015 and Ahmed et al. 2017, the method used for optimal plane searching has consisted in fitting a vast number of pre-defined planes with a given constant thickness to the satellite sample. These planes are forced to pass through the center of the main galaxy (see also, Buck et al. 2015). In Libeskind et al. 2007, Maji et al. 2017, and Garaldi et al. 2018, the ToI method has been used, mainly on the subsample of the 10-11 most massive satellites. In this way, planes of satellites

¹² Significance is understood as the inverse of the probability of occurrence of a particular plane of satellites versus an isotropical distribution.

have been found. The best ones, however, are not thin and populated enough so as to reproduce the VPOS-3 and VPOSall planes in the MW (see, e.g., Maji et al. 2017) or even the Ibata-Conn-14 plane in M31 (see, e.g., Gillet et al. 2015).

In this project, we further contribute to the study of zoom-in simulations by introducing a novel systematic and detailed plane searching technique to study the properties and quality of planes of satellites. This methodology is suitable to analyze both observational data and simulations. We aim to gain further insight on the properties of planes of satellites one can find in well-behaved massive disk galaxies (like the MW and M31), thus formed in zoom-in high-resolution cosmological hydrodynamical simulations, where a high enough number of resolved satellites can be identified. This is an important step, whose outcomes are needed prior to attempting any physical interpretation of the origin and/or evolution of planes.

Specifically, in this paper, we develop a detailed analysis of planar configurations of satellites from positional information by applying, for the first time, the 4GND plot method (Pawlowski et al. 2013) and its extension (introduced in Paper I) to simulations. For the reasons explained above, we focus on a set of zoom-in hydrodynamical simulations where well-behaved MW-like disk galaxies and their satellite systems form. Moreover, as the effects of galaxy companions either in binary systems or groups add complexity to this analysis, we consider only isolated galaxies. As mentioned previously, this method allows us to identify the constituent satellites forming planar spatial configurations and analyze their quality in terms of their population N_{sat} (or, equivalently, the fraction of satellites f_{sat}) and the ellipsoid of concentration axes (a, b, c , with $a > b > c$). Planes of high quality are those with a high f_{sat} and a low c/a , meaning they are populated and thin. These analyses are done over the entire galaxy’s evolution after halo virialization. This allows us to glean important insights into the different kinds of planar structures one can find in simulations of disk galaxies, and how they compare to the observed planes of satellites.

We note that in the “planes of satellites problem,” the persistence issue (i.e., is there a same group of satellites that is spatially distributed in a planar-like configuration across time?) is closely related to the kinematical character of planes. Therefore, in a forthcoming paper (I. Santos-Santos et al. 2020, in preparation, Paper III), the full six-dimensional phase-space information on satellite motions will be used to carry out kinematically based analyses as an optimal methodology to address satellite plane persistence.

The paper is organized as follows. In Section 2, we introduce the simulations analyzed, while in Section 3, their corresponding satellite samples and some of their properties are presented. Section 4 describes the method used for positional plane searching and plane quality analysis. Results obtained are reported in Sections 5 and 6. In Section 7, we assess the possible co-orbitation of satellites within the planes they form. In Section 8, we discuss the implications of our results. Finally, Section 9 summarizes the results and exposes the conclusions reached.

2. Simulations

We have chosen to analyze planes of satellites orbiting around isolated, simulated galaxies that resemble the MW system. In particular, we demand that the simulation meets the following requirements:

- a) to contain a central galaxy with a thin gaseous and stellar disk at redshift $z \sim 0$, with a large radial extent ($R = 15\text{--}25$ kpc).

We note that very thin disks are not that common in hydro-simulations yet.

Moreover, as merger events could destabilize the galaxy +satellites system, complicating clean plane detections as well as the possibility of reaching conclusions concerning the origin of these planes, we as well require:

- (b) an overall quiet assembly history, i.e., free of major merger events after virialization.

This is in line with the current understanding of the MW’s disk formation and accretion history (Belokurov et al. 2018; Helmi et al. 2018).

To allow for a proper statistical comparison of the results obtained with those coming from observations of the MW and M31 at $z = 0$, which harbor (at least) around 30 satellites each,¹³ the system must also:

- (c) host a numerous (~ 30) satellite population around the central galaxy.

Finally, in order to accurately compute the center of mass and the orbital angular momentum of the baryonic component of a simulated satellite, and rely on it as physical,

- (d) we demand satellite objects must include more than 50 baryonic particles.

We have pre-analyzed a set of different zoom-in cosmological hydro-simulations, finding among them two that reach the previous prerequisites: Aquarius- C^α and PDEVA-5004. Both simulations follow the “zoom-in” technique but make use of very different initial conditions, codes, and physics prescriptions. This fact will allow us to reach conclusions that are independent of simulation modeling.

2.1. Codes and Host Galaxies

2.1.1. Aquarius- C^α (Aq- C^α)

The initial conditions of this simulation come from the Aquarius Project (Springel et al. 2008), a suite of high-resolution dark matter simulations of Milky Way-sized halos, formed in a $100h^{-1}$ Mpc Λ CDM, cosmological box with parameters: $\Omega_m = 0.25$; $\Omega_b = 0.04$; $\Omega_\Lambda = 0.75$; $\sigma_8 = 0.9$; $n_s = 1$; $H_0 = 73$ km s $^{-1}$ Mpc $^{-1}$. In this project, we analyze a new re-simulation of the so-called “Aquarius-C” halo (hereafter Aq- C^α), including the hydrodynamic and sub-grid models described in Pedrosa & Tissera (2015). These include a multiphase model for the interstellar medium and a supernovae feedback scheme, where energy from both SNe Ia and SNe II is considered (see Scannapieco et al. 2005, 2006, for more details). The initial mass resolution of baryonic and dark matter particles is $m_{\text{bar}} = 4.1 \times 10^5 M_\odot$, and $m_{\text{dm}} = 2.2 \times 10^6 M_\odot$, respectively.

This galaxy presents a long period during which there is no merger, namely from $z \approx 1.5$ to $z \approx 0.15$. Soon after, a massive satellite galaxy collides with the disk, and it further suffers a very close encounter with another massive object at $z = 0$. Therefore,

¹³ Given the resolution we can afford currently in hydro-simulations, this implies that satellite mass functions are biased toward more massive satellites when compared to the MW or M31 mass functions.

the analysis we will describe in the following sections has been carried out up to $z = 0.18$. Properties of this galaxy measured at this z are $M_* = 7.6 \times 10^{10} M_\odot$, $M_{\text{gas}} = 5.6 \times 10^{10} M_\odot$, $M_{\text{vir}} = 1.5 \times 10^{12} M_\odot$, and $R_{\text{vir}} \simeq 219$ kpc. It is roughly more massive and larger than PDEVA-5004, as we show in the next section.

The halo mass-growth history follows a standard two-phase process: first a fast one with high mass-growth rates and then a slower one where this rate is lower. An important timescale for halo evolution is its collapse or virialization time when it gets decoupled from global expansion. This moment can be identified as the time when the radius enclosing the $z = 0$ halo mass stabilizes. Or, almost equivalently, as the time when the time derivative of the virial mass growth reaches a low value, indicating the end of the fast phase of mass assembly. For this halo, this happens between a universe age of $T_{\text{vir,AqC}} \simeq 6\text{--}7$ Gyr. In this case, 25% of the mass is accreted after collapse, with around a 10% in the last merger event near $z = 0$ (not analyzed in this paper).

2.1.2. PDEVA-5004

The PDEVA code is the OpenMP parallel version of the DEVA code, an AP3M+SPH code specially designed so that conservation laws (e.g., momentum, energy, angular momentum and entropy) hold accurately (Serna et al. 2003). It includes the detailed chemical feedback and cooling methods implemented by Martínez-Serrano et al. (2008) as well as inefficient star formation parameters in order to mimic the effects of energy feedback on regulating star formation (Agertz et al. 2011), which are assumed to work on sub-grid scales (Serna et al. 2003; Doménech-Moral et al. 2012). In particular, star formation is implemented following a Kennicutt–Schmidt law, with a $\rho_* = 1 \times 10^{-25} \text{ g cm}^{-3}$ density threshold and $c_* = 0.008$ efficiency. The following Λ CDM, parameters are assumed: $\Omega_\Lambda = 0.723$, $\Omega_m = 0.277$, $\Omega_b = 0.04$, and $h = 0.7$; in a 10 Mpc per side periodic box.

Several simulations have been run with this code, yielding a suite of different galaxies. The one used in this project is PDEVA-5004, previously studied in Martínez-Serrano et al. (2009), Doménech-Moral et al. (2012), Obreja et al. (2013), and Domínguez-Tenreiro et al. (2014, 2015, 2016, 2017), where satisfactory consistency with observational data has been found in all of the comparisons addressed. This galaxy has a remarkably thin gaseous and stellar disk and a relatively quiet history after virialization. At redshift $z = 0$ it has the following properties: $M_* = 3.05 \times 10^{10} M_\odot$, $M_{\text{gas}} = 8.6 \times 10^9 M_\odot$, $M_{\text{vir}} = 3.44 \times 10^{11} M_\odot$, $R_{\text{vir}} \simeq 183$ kpc. The mass resolution of baryonic and dark matter particles is $m_{\text{bar}} = 3.94 \times 10^5 M_\odot$, and $m_{\text{dm}} = 1.98 \times 10^6 M_\odot$, respectively. Particle masses do not change during the simulation.

The halo growth history can be found in Figure 1 of Domínguez-Tenreiro et al. (2017) where, again, a two-phase process can be clearly distinguished with $T_{\text{vir,5004}} \simeq 7.0$ Gyr, even if mergers around this event refrain us from a clean identification. Only 20% of the virial mass is assembled after this time, and no major mergers show up.

It is worth noting that the mass growth histories for both Aq-C $^\alpha$ and PDEVA-5004 are standard for galaxy halos with a quiet history after virialization, as demanded by the selection criteria above.

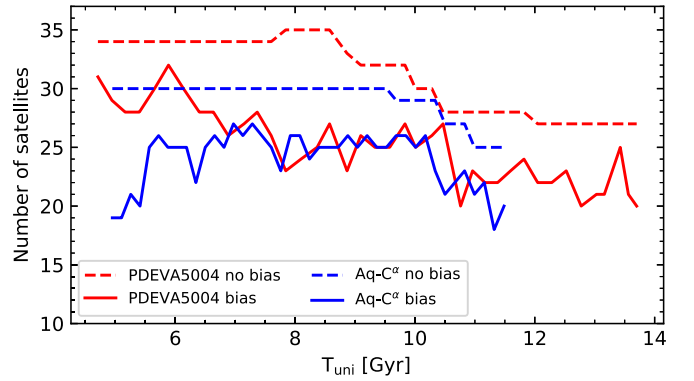


Figure 1. The total number of satellites N_{tot} in the two simulated samples as a function of the universe age T_{uni} . Red: PDEVA-5004; blue: Aq-C $^\alpha$. The dashed line shows results when all satellites are considered (“no bias”); the solid line shows results when the observational Galactic obscuration bias is applied (“bias”), hiding satellites orbiting in the plane of the disk at latitudes $|b| < 12^\circ$.

3. Satellite Samples

3.1. Identification

The samples of satellites in each simulation have been selected following the same criteria, that is, to take all objects with stars ($M_* > 0$) within a distance of 350 kpc from their host that are bound to the host galaxy, with a resolution limit of presenting at least 50 baryonic particles ($M_{\text{bar}} \approx 1 \times 10^7 M_\odot$). This selection has been made at redshift $z \sim 0.5$, to include satellites that may end up accreted by the disk at $z = 0$. To prove if a given object is indeed a satellite (i.e., is bound to its host), we have computed its orbit. This fixed sample of satellites has then been followed back and forth in time.¹⁴

For a proper comparison with MW results, we take into account Galactic obscuration—which prevents us from observing satellites orbiting in the plane of the disk of the MW, by applying an observational bias to the simulated satellite sample at each timestep. Following Pawłowski (2016), we have chosen it to hide objects with projected positions on the sphere at latitudes $|b| < 12^\circ$ (angular distance as measured from the plane of the galaxy disk). This is a first approximation to try to mimic the obscuration effects of the MW’s disk; however, we acknowledge that a more precise model to account for Galactic obscuration should depend on satellite distances and luminosities.

The selection of satellites in Aq-C $^\alpha$ has been done using a Friends-of-Friends algorithm to identify structures and then the SubFind halo finder (Springel et al. 2001) to construct subhalo catalogs at each timestep. Particle IDs have been used to trace back in time the selected satellites. The tool used for the selection of satellites in PDEVA-5004 has been IRHYS (by H. Artal, under development). This visualization and analysis tool permits the selection of objects (i.e., satellites) as sets of particles and enables us to trace them back and forth in time.

A total number of $N_{\text{tot}} = 30$ (35) satellites have been detected in Aq-C $^\alpha$ (PDEVA-5004) at selection time ($z \sim 0.5$), of which 25 (27) survive until the last analyzed timestep, respectively. In Figure 1, such numbers are plotted as a function of the universe age T_{uni} . N_{tot} changes because satellites disappear as they are accreted by the central disk galaxy. Also,

¹⁴ No new satellites, different to these, are accreted after selection time during the periods analyzed.

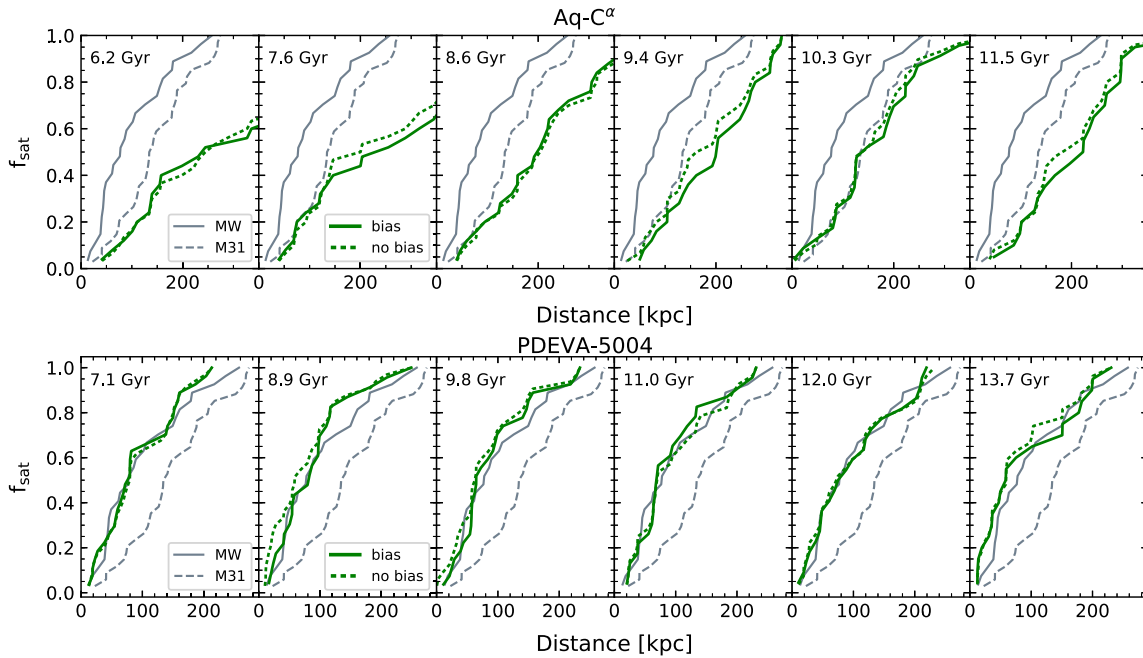


Figure 2. The radial distribution of satellites at different universe ages. Top panels: Aq-C $^{\alpha}$. Bottom panels: PDEVA-5004. The green solid line shows the distribution for all of the satellites present at the given timestep, while the dotted line takes into account the observational bias for Galactic obscuration. The gray lines show the distributions of satellites in the MW (solid) and in M31 (dashed) at $z = 0$.

satellites are not considered during periods where they orbit at distances >450 kpc (this happens with a couple of backsplash galaxy cases). In the case when obscuration in the plane of the disk is considered (“bias”), N_{tot} varies additionally because satellites go into and out of the avoidance volume.

3.2. Satellite Radial Distances and Their Distributions

The suite of Aq-C $^{\alpha}$ and PDEVA-5004 satellites presents different properties and evolutionary histories that reflect in a variety of orbits. We find some satellites that progressively lose angular momentum and are eventually accreted by the disk, some that follow orbits where successive apocenter and pericenter distances do not show important variations, and some backsplash galaxies that have just recently been captured by the halo and orbit at long distances occasionally even outside the virial radius. Interestingly, when all radial distance histories are plotted together, a coincidence of pericenters is clear at certain moments. In this way, as the simulations evolve, there are moments of maximum spatial expansion and moments of a maximum compression of the satellite systems.

These effects can be observed when analyzing the evolution of the radial distribution of satellites with cosmic time. Figure 2 (top panel: Aq-C $^{\alpha}$, bottom panel: PDEVA-5004) shows the fraction of the total number of satellites within a certain distance from the center of the main galaxy, compared to the MW and M31 distributions at $z = 0$.¹⁵ Two colored lines are shown per panel, which represent the results obtained using all of the satellites, or taking into account the observational obscuration bias, at each timestep. The distributions change with time, showing periods where there is a higher concentration of satellites at shorter distances (corresponding precisely to

the moments of maximum collective approach) and others where there is a higher expansion of the system. As an example, the PDEVA-5004 system (in the case where all satellites are considered, i.e., “no bias”) is more compact (higher f_{sat} within 100 kpc) at $T_{\text{uni}} = 8.9$ Gyr, and more expanded at $T_{\text{uni}} = 11$ Gyr. On the other hand, in the Aq-C $^{\alpha}$ system, it is not until $T_{\text{uni}} \sim 9$ Gyr that the complete sample of satellites is within a distance of ~ 350 kpc. A moment of maximum compactness is $T_{\text{uni}} \sim 10.3$ Gyr. Curiously, PDEVA-5004’s satellite radial distribution resembles very well that of the MW for a long period of time, while Aq-C $^{\alpha}$ ’s is similar to that of M31 at $T_{\text{uni}} \sim 10$ Gyr. These resemblances are kept when using N_{sat} instead of f_{sat} in the case that total satellite numbers of simulations and observations are matched (see Sections 8.1 and 8.2). We explore the effect of radial compactness on the quality of planes of satellites in Section 8.3.

3.3. Mass Distribution of Satellites

Satellites in Aq-C $^{\alpha}$ (PDEVA-5004) show baryonic masses ranging between $M_{\text{bar}} = 8.6 \times 10^6 - 8.9 \times 10^8 M_{\odot}$ ($M_{\text{bar}} = 3.9 \times 10^7 - 1.8 \times 10^8 M_{\odot}$). This differs from the mass range of confirmed MW/M31 satellites. Indeed, the objects with lowest stellar masses considered in this work have $M_{*} \sim 8 \times 10^6 M_{\odot}$ (Aq-C $^{\alpha}$) and $M_{*} \sim 1 \times 10^7 M_{\odot}$ (PDEVA-5004) (see requirement (d) above), while observed MW/M31 satellites reach as low as $M_{*} \sim 5 \times 10^2 M_{\odot}$ (e.g., SegueI), with 13 out of 27 galaxies in the MW presenting masses lower than $M_{*} < 5 \times 10^4 M_{\odot}$.¹⁶ In recent years, cosmological hydrodynamical simulations have reached a very high mass and spatial resolution; however, it is still not high enough as to

¹⁵ The sample of MW and M31 satellites used is that described at the beginning of Section 5.1.

¹⁶ Stellar masses for observed galaxies are calculated from the luminosity values in McConnachie (2012), using the mass-to-light ratios from Woo et al. (2008) according to galaxy morphological type.

produce as many resolved bound objects with masses as low as those at the low-mass end of the MW and M31 satellite mass functions.¹⁷

These differences between the simulated and observed mass distributions are not expected to introduce a determinant bias on the formation of planes of satellites as analyzed in this work. Indeed, it was shown in Paper I with the observed MW and M31 satellites (which span a wider mass range than simulated ones) that stellar mass is not a satellite property determining its membership or not to the high-quality positionally detected planes found at $z = 0$ with the 4GND plot method.¹⁸ Additionally, from the empirical side, the fact that objects of different mass scales, like some young globular clusters and stellar streams, seem to be as well within the observed VPOS plane of satellites in the MW (Pawlowski et al. 2012, Riley et al. 2020), supports our findings. This is an important result in view of the rather narrow satellite baryonic mass range we can currently afford in hydrodynamical simulations. Therefore, we are allowed to meaningfully compare planes from the observed MW/M31 satellite samples and those of our simulations, even if the masses of the involved satellites span different mass ranges.

4. Searching for Planes of Satellites from a Positional Analysis

To search for planar positional configurations of satellites in our simulations we have followed the *4-galaxy-normal density plot* (4GND plot) method presented in Pawlowski et al. (2013), extended as explained below (see also Paper I). This method allows us to determine if there is a subsample out of a given sample of N_{tot} satellites that defines a planar arrangement in terms of the outputs of a fitting technique based on the Tensor of Inertia (ToI; Metz et al. 2007; Pawlowski et al. 2013). Satellite planes are searched for through a regression method that minimizes orthogonal distances from the points to the optimal plane solution. Apart from the plane (or equivalently, its normal vector \mathbf{n}), the outputs of the regression can be characterized by the following parameters in terms of the corresponding ellipsoids of concentration (Cramér 1999):

- (i) N_{sat} : the number of satellites in the subset (or, the fraction of satellites it involves $f_{\text{sat}} \equiv N_{\text{sat}}/N_{\text{tot}}$);
- (ii) c/a : the ellipsoid short-to-long axis ratio;
- (iii) b/a : the ellipsoid intermediate-to-long axis ratio;
- (iv) Δ_{rms} : the root-mean-square thickness perpendicular to the best-fitting plane;
- (v) D_{CG} : the distance from the center-of-mass of the central galaxy to the plane.

These outputs are used to quantify the quality of the best-fitting structures to a subsystem of N_{sat} satellites. First of all, assuming $c/a < 1$, b/a indicates whether the distribution is planar

($b/a \sim 1$) or filament-like ($b/a \ll 1$). High-quality planes are those that involve many satellites and are thin, therefore demanding high f_{sat} and low c/a (or equivalently low Δ_{rms} , a quantity that most often is correlated with c/a once the system acquires its stable size). Low D_{CG} planes pass near the disk center, a requirement asked of a physically consistent satellite system when its gravitational potential minimum lies approximately at this center. Finally, \mathbf{n} allows us to visualize the plane orientation with respect to a given reference frame, for example, the host galaxy disk plane. At the end of this section, the quantification of plane quality, as well as how to compare the qualities of two or more planes, is described in more detail (see also Paper I).

4.1. Method: Four-galaxy-normal Density Plots

We have applied the 4GND plot method to each timestep of the two simulations. A thorough description follows (see also Section 2.4 in Pawlowski et al. 2013, and Paper I).

1. A plane is fitted (using the ToI method) to the positions of every combination of four different satellites taken from the total sample of N_{tot} satellites. As three points always define a plane, four is the lowest possible amount to take into consideration under the condition of making the number of combinations high enough to get a good outcome signal.¹⁹
2. The axes sizes a , b , c and normal vector directions (i.e., four-galaxy-normals) of the planes fitted to each combination of four satellites are stored. Then, all of the four-galaxy-normals are plotted in a galactocentric coordinate system such that the central disk's spin points toward the south pole, and a density map (i.e., 2D-histogram) is drawn from their projections on a regularly binned sphere with N_{bin} bins. Spherical projections are shown with Aitoff diagrams in Galactic (longitude l , latitude b) coordinates in a $l = [-90^\circ, +90^\circ]$ projection because opposite normal vectors are equivalent. As in Pawlowski et al. (2013), each normal is weighted by $w = \log\left(\frac{a+b}{c}\right)$ to emphasize planar arrangements of satellites over filament-like or spherical-like ones. In these plots, over-densities (or density peaks, i.e., areas of four-galaxy-normal vector accumulation) are signaling groups of four satellites contributing to the same dominant planar space-configuration. As illustrated in Figures 3 and 4, we show examples of 4GND plots for the Aq-C α and PDEVA-5004 simulations, respectively. In some cases, these over-dense areas are more extended, and in others, more concentrated. (Note that the expectation from a random distribution of satellites is a density map with equal density in each bin.)
3. Density peaks are differentiated and isolated. A set of N_{peaks} density peaks are selected around the highest-density bins of the 4GND plots, with the requirement that they are separated more than 15° from the center of all of the $(N_{\text{peaks}} - 1)$ over-densities. The specific peak location in (l, b) is given by the center of the corresponding high-density bin.²⁰

¹⁷ For example, Buck et al. (2019) (i.e., a higher-resolution re-simulation of NIHAO MW-like galaxies with $m_{\text{gas}} \approx 5 \times 10^4 M_\odot$), produce one case with 20 satellites at $z = 0$, reaching a low-mass end of $M_* = 1.5 \times 10^5 M_\odot$. Wetzel et al. (2016)'s Latte MW simulations have a mass resolution of $m_{\text{gas}} \approx 7 \times 10^3 M_\odot$, and produce a number of 13 satellites at $z = 0$ that reach a lowest mass of $M_* = 8 \times 10^4 M_\odot$. Ahmed et al. (2017)'s sample of MW-like galaxies with $m_{\text{gas}} \approx 3 \times 10^4 M_\odot$ present a large number of satellites at $z = 0$ (~ 30), and reach a lowest mass of $M_* = 2 \times 10^4 M_\odot$.

¹⁸ This was quantitatively confirmed by finding low correlation coefficients between the total stellar mass of a satellite and its number-contribution, $C_{p,s}$ (see definition in Paper I and in Section 4.1) to the main planar configurations of satellites (i.e., the highest-quality planes).

¹⁹ The number of such combinations is given by $\frac{N_{\text{tot}}!}{N_{\text{pl}}!(N_{\text{tot}} - N_{\text{pl}})!}$, where N_{pl} is the number of satellites included in the planes.

²⁰ As we show in Figure 9, changing the size of the bins does not modify our results or conclusions.

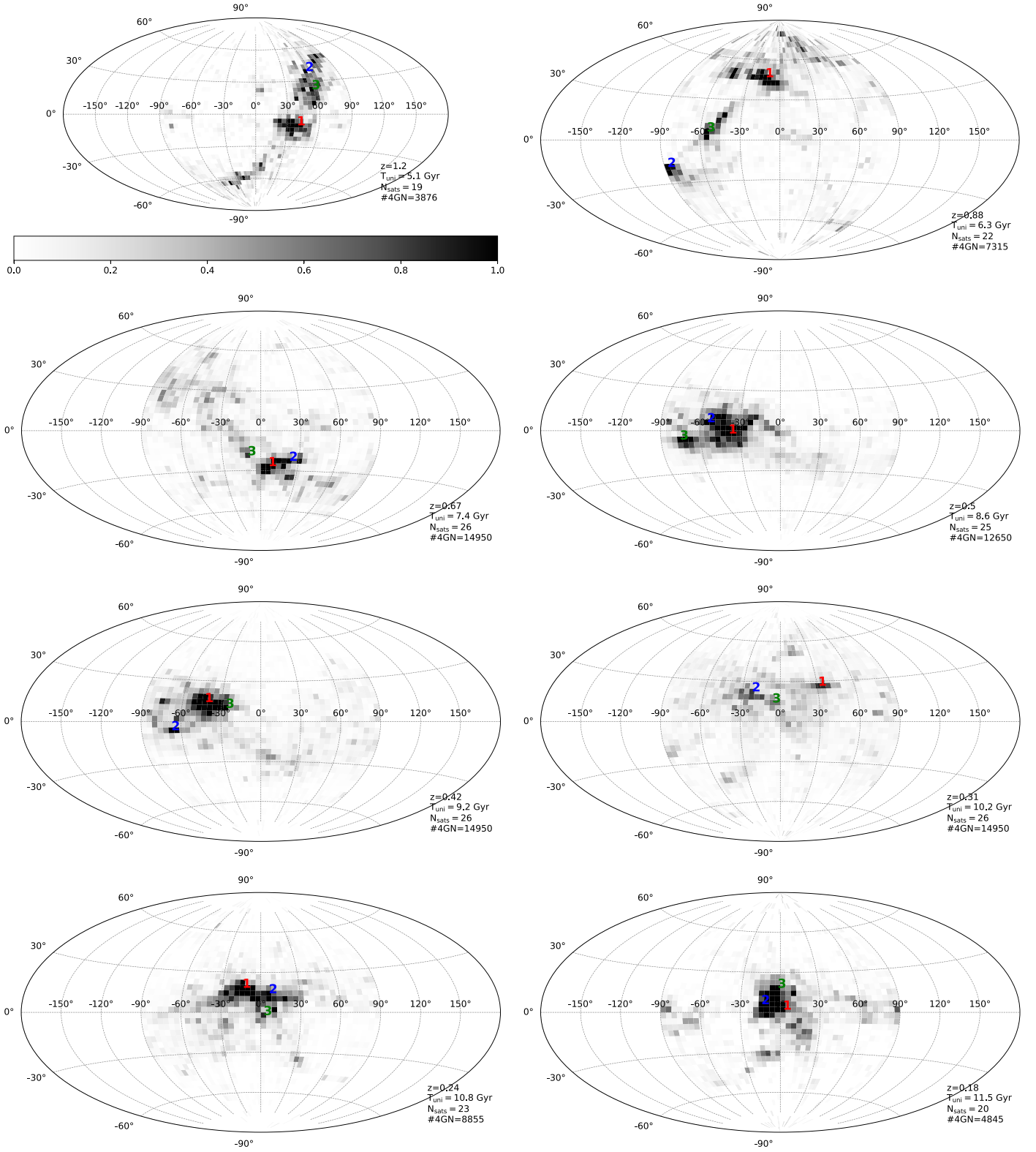


Figure 3. Examples of four-galaxy-normal density plots (4GND plots) for galaxy Aq-C^α (having applied the observational obscuration bias) at different times. The legend shows the redshift z , the cosmic time it corresponds to in Gyr, the *total* number of satellites considered, and the *total* number of four-galaxy-normals at that timestep (#4GN). The main density peaks, used for analyses in this work, are marked with numbers ordered according to the density of their central bin. A color code is also used to identify their contributions in the next figures. The grayscale colorbar is common for all timesteps and its values are proportional to the normalized bin density.

4. We determine how much a given satellite s contributes to the p th specific peak (i.e., its respective contribution-number, $C_{p,s}$, with $p = 1, \dots, N_{\text{peak}}$ and $s = 1, \dots, N_{\text{tot}}$). To this end, we select all of the four-galaxy-normals that fall

within an aperture angle of 15° ²¹ around the p th peak location. Each of the four satellites contributing to these

²¹ We take the same angle as that used in Pawłowski et al. (2013)'s analysis.

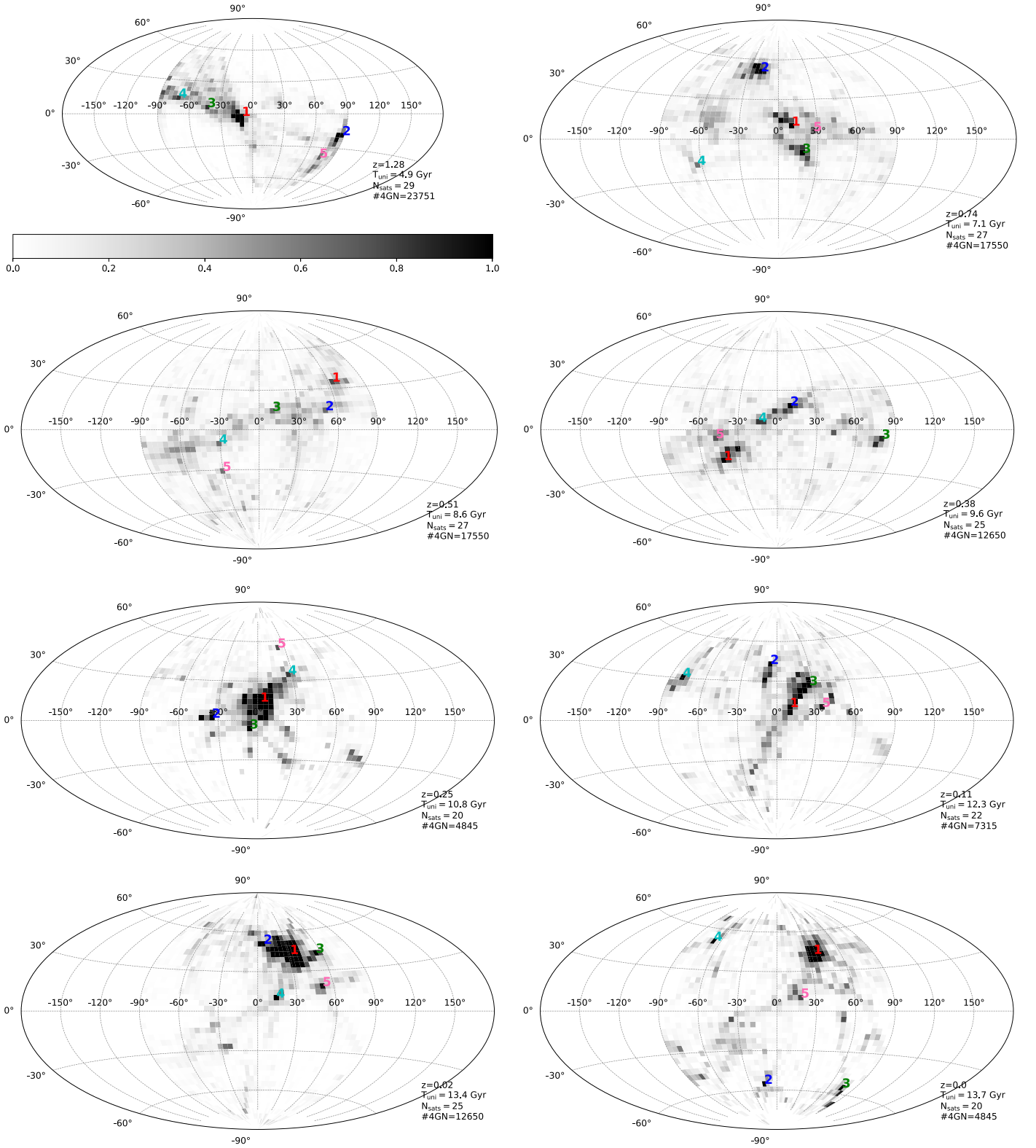


Figure 4. Same as Figure 3 for PDEVA-5004.

four-galaxy-normals is counted as contributing the four-galaxy-normal's *weight* to the peak. This has been normalized using $C_{N,\text{all}}$, the total weighted number of four-galaxy-normals, including those that are not within 15° of some peak center, such that the sum $\sum_{p,s} C_{p,s} = 1$. Such normalization is necessary for a meaningful

comparison of results at different timesteps (where N_{tot} varies) and, also, with observational data. At fixed p , $C_{p,s}$ is high when satellite s contributes to many of the four-galaxy-normals laying within 15° of the peak center. We note that over-densities that are located close to each other on the sphere generally share many of the satellites

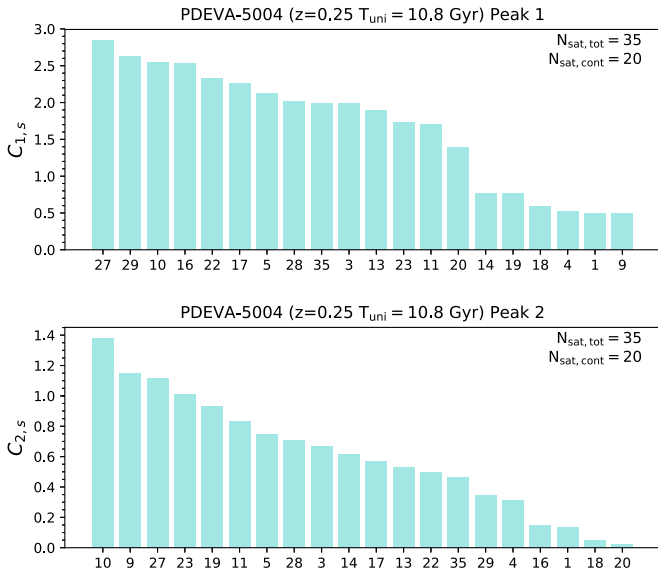


Figure 5. Bar charts showing the contribution ($C_{p,s}$) of satellites to four-galaxy-normals in 15° around the first- and second-most important over-densities in PDEVA-5004’s 4GND plot at $T_{\text{uni}} = 10.8$ Gyr (including the observational obscuration bias). The x-axis shows contributing satellite IDs in decreasing $C_{p,s}$ order. The total number of satellites considered at the given timestep, $N_{\text{sat,tot}}$, and the total number of satellites contributing to four-galaxy-normals to the given peak, $N_{\text{sat,cont}}$, are stated in the right corner of the panels.

that contribute most to four-galaxy-normals within 15° of the peak. However, the peak isolation criterion used in the previous step cures our peaks of such redundancies.

5. For a given peak p , we order all satellites by decreasing $C_{p,s}$ to it. This is done for all of the isolated peaks found in a given 4GND plot, and for each of them, we obtain an ordered list of contributing satellites. Examples of such lists are shown in Figure 5 for two different peaks defined at a given timestep in PDEVA-5004. The x-axis shows satellite IDs in decreasing $C_{p,s}$ order (see y-axis values).

4.2. An Extension to the Method

We have extended the 4GND plot method presented in Pawlowski et al. (2013) to thoroughly evaluate the properties and quality of the planar structure of satellites revealed by each over-density.

4.2.1. Peak Strength Analysis

In order to analyze the number of relevant density peaks at each timestep and how this number evolves with time, to each peak we assign a number, C_p , the peak strength, defined as the normalized number (or%) of four-galaxy-normals within 15° of the respective peak center; that is $C_p \equiv \sum_s C_{p,s}$, where the contribution-number $C_{p,s}$ of the s satellite to the p th peak is defined in step (4) above. For example, in Figure 5, C_1 (C_2) would be obtained by summing up the $C_{1,s}$ ($C_{2,s}$) corresponding to all of the satellites in the upper (lower) panel of the figure.

By reckoning the number of peaks with C_p above given thresholds (i.e., the observational ones, for example), we can compare to observations. We can also calculate how many peaks of a given strength there are at given timesteps in the simulations and how this number evolves with time.

4.2.2. Plane Quality Analysis

To analyze individually each over-density in terms of quality, as explained in Paper I, we start by fitting a plane to the seven satellites that contribute the most to it.²² Then, following the order of $C_{p,s}$ contribution, we iteratively add one more satellite at a time and fit a plane to the new resulting satellite set, storing the ToI fitting outputs described at the beginning of this section. This plane-fitting process is repeated until all of the contributing satellites to the peak under consideration are used. In this way, for each peak found at a given timestep of the simulation, we obtain a collection of planes of satellites, each consisting of an increasing number of members.

As stated above, plane quality is measured through the f_{sat} and c/a values. Being a two-parameter notion, when comparing the quality of two planes, if, in one of them, c/a is lower and f_{sat} is higher than in the second, then the first plane has higher quality. Other cases when the qualities of two planes can be compared are when either f_{sat} is constant (in which case the plane with lowest c/a has a higher quality), or when c/a is constant (or at least it varies slowly with f_{sat}), in which case, the higher f_{sat} , the better the quality.

As a practical implementation of these ideas, in Section 6.2 we study how b/a and c/a vary as the number of satellites included in the plane-fitting increases; see, e.g., Figures 7 and 8, where the collection of planes associated with a given density peak is characterized by a line.

5. Results: Density Peak Analysis

5.1. Four-galaxy-normal Density (4GND) Plots

In Paper I, the extended 4GND plot method has been applied to the same MW and M31 satellite samples used in Pawlowski et al. (2013), consisting of $N_{\text{tot}} = 27$ and 34 satellites for the MW and M31, respectively. These are the confirmed satellites within 300 kpc from their hosts, according to the McConnachie (2012) “Nearby dwarf galaxy database”²³ as of June 2013.²⁴ The MW shows one important peak, while M31 shows two. A detailed analysis of the corresponding planar configurations they point to is presented in Paper I.

Figures 3 and 4 show examples of 4GND plots for Aq-C $^\alpha$ and PDEVA-5004, respectively, where the observational Galactic obscuration bias has been applied. The legend shows the redshift z , universe age T_{uni} , total number of satellites considered N_{tot} , and the total number of four-galaxy-normals (# 4GN) at that timestep. The main peaks, used for analyses in this work, are marked with numbers, ordered according to the their central bin density. Note that peaks are selected and ordered independently at each timestep, and that a peak labeled $p = 1$ is not necessarily related to another labeled the same way at a different timestep. A color code is also used to identify their contributions in the next figures. The number, strength, and location of over-densities changes with time from showing several intermediate/low over-densities at some moments, to one that clearly dominates at others (especially at the last

²² This number $N_{\text{sat}} = 7$ is low enough to allow for an analysis of the ToI parameters behavior as N_{sat} increases and at the same time high enough that we begin with populated planes. Note that taking instead $N_{\text{sat}} = 7 \pm 2$ to begin with does not alter our conclusions.

²³ http://www.astro.uvic.ca/~alan/Nearby_Dwarf_Database_files/Nearby_Galaxies.dat

²⁴ In Paper I, the most up-to-date sample of confirmed MW satellites is studied as well.

Table 1Peak Strengths C_p for the Main Two Peaks Found in the MW and M31 4GND Plots (see Paper I)

	$C_1 \pm \sigma$ (%)	$C_2 \pm \sigma$ (%)
MW	22.92 ± 0.26	14.31 ± 0.20
M31	10.53 ± 0.62	10.52 ± 1.62

Note. Peak strength is computed as $C_p \equiv \Sigma_s C_{s,p}$, where $C_{s,p}$ is normalized to the total weighted number of four-galaxy-normals (see Section 5.3). Results shown are means and 1σ standard deviations calculated from 100 random realizations using the radial distance uncertainties. Peaks #1 are the strongest ones and Peaks #2 follow in strength.

timesteps analyzed). This behavior will be studied in the sections that follow.

Results obtained when the observational obscuration bias is not applied, and therefore all satellites are taken into account, do not differ substantially from those shown in Figures 3 and 4. Only at some timesteps are new features able to be seen around the poles of the Aitoff diagram as compared to its “bias” counterpart figure, contributed to by satellites orbiting in a plane close to that of the disk of the central galaxy.

5.2. Satellite Contribution-numbers to Peaks

As explained at the end of Section 4.1, for each peak p in a density plot, we obtain a list of satellites ordered according to their respective contribution-numbers $C_{p,s}$. This is the order in which satellites are added to the plane-fitting procedure explained in Section 4.2.2, to build the peaks’ corresponding collection of planes. In Figure 5, we illustrate the $C_{p,s}$ histograms corresponding to satellites s contributing to Peak 1 (top panel) and Peak 2 (bottom panel) of PDEVA-5004’s 4GND plot at $T_{\text{uni}} = 10.8$ Gyr (see Figure 4). The x-axis shows the IDs of satellites; only the nonzero contributions have been plotted. Some satellites show a high contribution to one peak while others do not, meaning that they are involved in a low number of four-galaxy-normals close to the respective peaks. In this particular case, we see that those satellites showing a high $C_{p,s}$ relative to the main peak are not among those contributing the most to the second peak.

5.3. Peak Strength Analysis

The C_p peak strengths of Peak 1 and Peak 2 (i.e., C_1 and C_2) for the MW and M31 are given in Table 1. Errors are 1σ deviations over 100 random realizations of their radial distance uncertainties, as explained in Paper I. These are especially large in the case of M31 satellites; as a result, peaks in its 4GND plot are more blurred as compared to the MW ones. This contributes to lower C_1 and C_2 values in M31.

In the upper panels of Figure 6, we present the value of C_1 at each timestep for Aq-C $^\alpha$ and PDEVA-5004. C_1 fluctuates, reaching values that can be even higher than those of the MW or M31 at $z = 0$. In general, the application of the observational obscuration bias enhances the strength value of the main peak.

The universe ages T_{uni} where the respective values of C_1 (in the case where all satellites are considered, “no bias”) have maxima (minima) are marked by green (magenta) vertical lines. These time intervals of local maxima and minima have an average duration of 0.5–1 Gyr (consistent with the values Shao et al. (2019) find in their analyses of the EAGLE simulation).

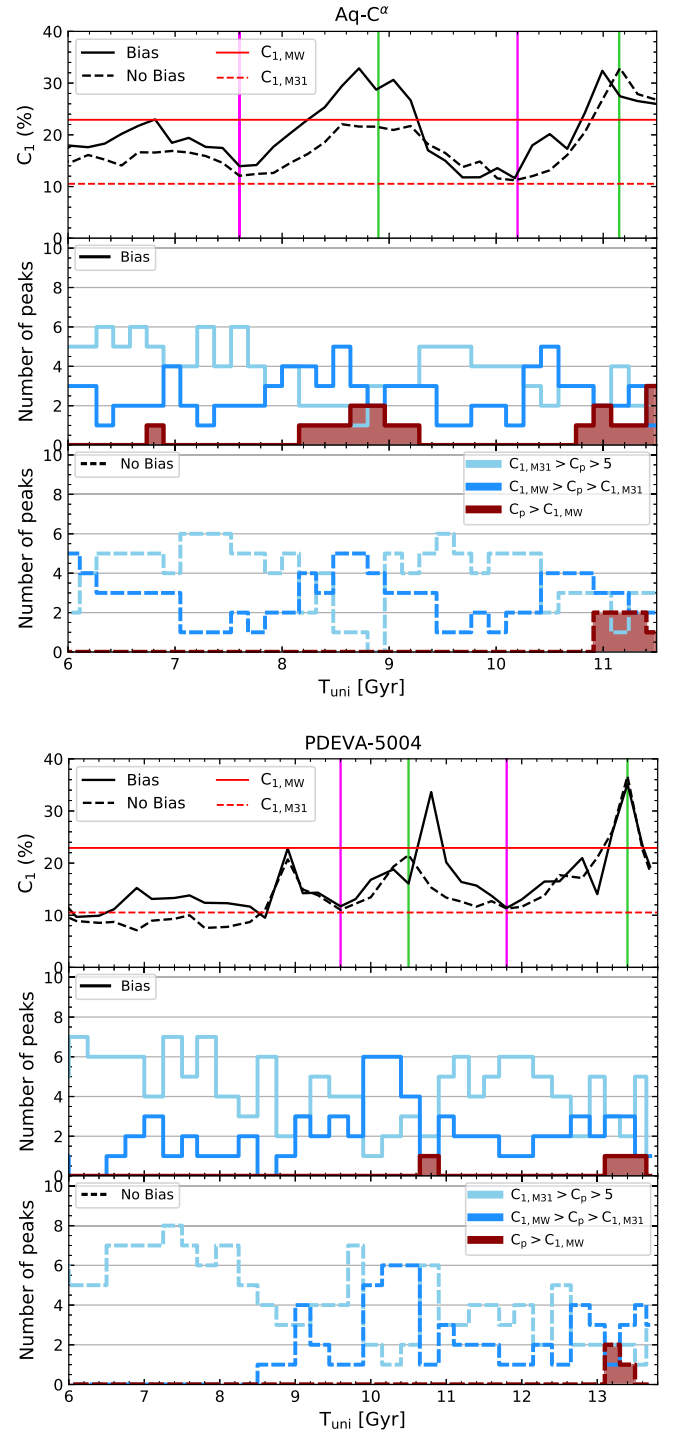


Figure 6. Upper panels: C_1 peak strength as a function of the universe age, for both biased and non-biased satellite samples. The green (magenta) vertical lines mark the moments when non-biased C_1 reaches maximum (minimum) values. Lower panels: number of density peaks at each timestep with C_1 strengths within given intervals. These intervals are defined by the MW and M31 C_1 values, shown as horizontal lines in the upper panels (see also Table 1). Top set of panels: Aq-C $^\alpha$; bottom set of panels: PDEVA-5004.

This has been estimated from their FWHM, where we take the mean C_1 as the floor value. These periods will be related with plane quality in the next sections.

Another interesting possibility that the peak strength C_p allows is to determine the number of peaks with strengths above given thresholds or within given intervals, at different

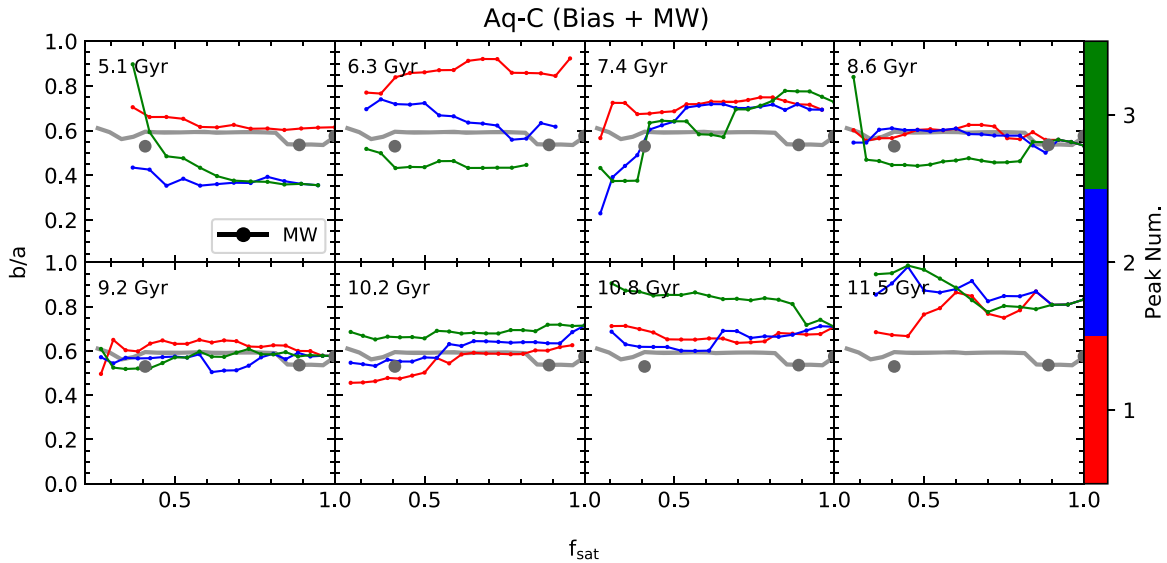


Figure 7. The intermediate-to-long axis ratio b/a as a function of the fraction of satellites f_{sat} included in the plane at different universe ages, T_{uni} , for Aq-C $^{\alpha}$. We compare to the MW Peak 1 result and, hence, have applied the observational Galactic obscuration bias to the simulation. The results for the planar structures defined by the most-prominent density peaks are shown as lines of different colors. The color code and numbering allow us to find the corresponding peak in the 4GND plots (Figure 3). The gray solid line shows the result obtained from the MW’s main density peak, and points show the specific values for MW observed planes of satellites mentioned in the literature (i.e., classical, VPOS-3, VPOSall, Pawlowski et al. 2013).

universe ages. This is shown in the lower panels of Figure 6, with respect to the strengths of the peaks labeled as #1 in the MW and M31, i.e., $C_{1,\text{MW}} = 22.9\%$ and $C_{1,\text{M31}} = 10.5\%$. At given times, there are a few peaks encompassing a high % of four-galaxy-normals (high C_1) that then break into several different peaks with lower strengths. These later collimate into high C_p peaks again. That is, the number of peaks with C_p within given peak strength intervals fluctuates with time. We recall that the background is also accounted for to normalize the peak strengths at fixed timesteps.

Summing up, as measured with C_1 , the peak strengths in observations and simulations are consistent within given time intervals (~ 3 periods with peaks stronger than the MW in the ‘bias’ case, and 1 such period in the ‘no-bias’ case.). Regarding the number of peaks, we see that when C_1 takes high values, the number of weak peaks decreases and that of stronger peaks increases. This happens especially at the last timesteps analyzed. In particular, the number of strong peaks ($C_1 > C_{1,\text{MW}}$) reaches 1–3 at most, (occurring at late times), consistent with observations. Again, we can see that the obscuration bias favors the emergence of strong peaks in both simulations.

6. Results: Plane Quality Analysis

6.1. Comparing to the MW and M31 Satellite Systems

The quality of planar configurations of satellites obtained from the two main density peaks in the MW and M31 has been analyzed in Paper I. The extension to Pawlowski’s 4GND plot method presented in Paper I has revealed a richer and higher-quality plane structure in the MW and M31 than that reported previously in the literature. In particular, in both the MW and M31, the quality of the planar structure of satellites provided by Peak 1 is, at any N_{sat} , better than that corresponding to Peak 2 (although, in the case of M31, the differences between peaks are not that important when the error bands are taken into consideration).

In this paper, we focus on the best-quality planes at constant f_{sat} that can be found in a satellite system at a given moment; therefore, we will compare our simulation results to the MW and M31 *Peak 1* results, while those of Peak 2 will not be used in this paper for comparison purposes.

As said above, it was also shown in Paper I that there is no correlation between the stellar mass of an observed satellite and its $C_{p,s}$ contribution to the main density peaks found. This allows for a fair comparison between simulations and observations despite the different satellite mass ranges involved.

6.2. Quality of Simulated Planes in Terms of the Satellite Fraction Involved

6.2.1. b/a versus f_{sat}

Concerning the application of the method to the simulation data, we first address the planar ($b/a \sim 1$) or filamentary ($b/a \ll 1$) character of the best-fitting structures found with the ToI analysis, where b/a is the intermediate-to-long axis ratio in the ToI scheme. As an illustrative example of our results, in Figure 7, we plot b/a versus f_{sat} for the main density peaks found in Aq-C $^{\alpha}$ at different timesteps. In this figure, and in the following ones that compare the changes of a ToI output with f_{sat} , each panel corresponds to a given timestep. Lines of different colors stand for the collections of planes of satellites associated with the respective peaks numbered with the same color coding in Figure 3. Based on Figure 6 and the 4GND plots shown in Figures 3 and 4, the consideration of a number of $N_{\text{peak}} = 3$ peaks for Aq-C $^{\alpha}$ and $N_{\text{peak}} = 5$ peaks for PDEVA-5004 seems a reasonable choice ensuring the exploration of all possibly relevant planar configurations.

Observational data results are shown as gray lines and points. Points show the specific values for MW/M31 observed planes of satellites at $z = 0$ mentioned in the literature (i.e., MW: classical, VPOS-3, VPOSall; M31: Ibata-Conn-14,

GPoA; see values in Table 1 of Paper I). When comparing to the MW, we show simulated results where the obscuration bias is applied; when comparing to M31, we show results considering all satellites.²⁵

We find no filamentary structures; in fact, at all timesteps, a planar structure exists whose b/a is larger than the observational case, at all f_{sat} . The general behavior of b/a , both for Aq-C $^{\alpha}$ and PDEVA-5004, and in the “bias” and “no bias” cases, is that b/a changes only slightly when new satellites are added to the fit, giving rise to wide f_{sat} intervals where b/a is almost constant. This means that the planar character of the spatial configuration of satellites does not depend very much on the number of satellites involved. This behavior is also found in the MW and M31 (see gray lines in Figure 7).

6.2.2. c/a versus f_{sat}

Having confirmed that the structures found in our simulations are indeed planar, we can proceed with the study of the quality of such planes through c/a . As explained previously, quality is assessed by a two-parameter notion ($c/a, f_{\text{sat}}$) such that at a given f_{sat} , the plane with the lowest c/a presents the highest quality. In particular, we define that a *strong* consistency exists between plane collections from simulations and observations when there is one colored line from simulations with similar or lower c/a values than that of the MW/M31 gray line at all f_{sat} . A *weaker* condition refers to consistency between an observed plane and one detected in simulations with the same *particular* f_{sat} . In this case, the peak assuring consistency between data and simulations can vary from f_{sat} to f_{sat} .

We have carried out the analysis of plane population and thickness (c/a versus f_{sat}) for the main peaks found in Aq-C $^{\alpha}$ and PDEVA-5004, comparing to both the MW and M31. As examples of the results obtained, in Figure 8, we present the results for Aq-C $^{\alpha}$ versus MW (top panel; the observational obscuration bias is applied) and PDEVA-5004 versus M31 (bottom figure; all satellites are considered).

Independently of applying the observational obscuration bias or not, both Aq-C $^{\alpha}$ and PDEVA-5004 present thin and highly populated planes at all timesteps. In general, c/a is low ($\lesssim 0.3$) for all peaks at all timesteps when including up to $\sim 80\%$ of satellites. This is already proving the oblate spatial distribution of the entire satellite population in both simulations. In particular, at *all* timesteps and in both simulations, there are planar structures compatible (in terms of c/a and f_{sat}) with the M31 Ibata-Conn-14 plane and the MW classical plane. The GPoA value is also recovered in almost all timesteps. The strong consistency condition is met in many cases. For example, we can find very similar or higher-quality planar structures than that of the MW in Aq-C $^{\alpha}$ (“bias” case) at $T_{\text{uni}} = 8.6, 9.2$ and 10.8 Gyr; and we find similar or higher-quality planar structures than that of M31 in PDEVA-5004 at $T_{\text{uni}} = 4.9, 9.6, 10.8, 13.4$ and 13.7 Gyr.

6.2.3. Evolution of Plane Quality with Cosmic Time

A more compact and clearer way of presenting the results on plane quality analysis showed in Figure 8 is looking at the “best” plane found at each timestep including a fixed fraction of

the total number of satellites. This best plane is selected as the one with the lowest c/a at fixed f_{sat} , which can be easily read from Figure 8 (note that this best-quality plane does not necessarily correspond to the same peak as f_{sat} changes).

In Figure 9, we show the properties of the “best” planes of satellites found at each timestep. In particular, we focus on c/a , Δ_{rms} height, and the inclination of the plane relative to the disk (latitude angle). Different shades of blue stand for planes with different $f_{\text{sat}} = 30\%, 50\%, 70\%, 90\%$. We show results with and without applying the observational obscuration bias in the left and right panels, respectively. In the right panels, green (magenta) lines mark the values of the universe ages where C_1 maxima (minima) appear in Figure 6. For comparison, the results for the best planes in the MW and M31 at $z = 0$ with the same f_{sat} are also shown as horizontal lines. These values, together with those of simulated results averaged over the last 1 Gyr analyzed, are given in Table 2. Note how at high f_{sat} (70% and 90% lines), M31 presents very large c/a and Δ_{rms} , due to the system configuration in two almost perpendicular planes (see Paper I).

In terms of c/a and f_{sat} , both Aq-C $^{\alpha}$ and PDEVA-5004 simulations present high-quality planes. The best planes of satellites take c/a values that change with time, reaching at some timesteps, and particularly near the respective last periods analyzed, values compatible with those in the MW and M31 at $z = 0$ involving the same fraction of satellites.

We note that the fact that quality in our simulations increases toward low redshift is in contrast to Shao et al. (2019) findings, who report thinner planes of satellites at early times in EAGLE. On the other hand, Figure 9 shows that biased c/a results are systematically somewhat lower than non-biased ones in both simulations. This occurs because when applying the bias, we are removing a fraction of the volume where satellites can be considered, which prevents plane-thickening. This result may indicate that the quality of the MW’s planar structure of satellites can appear artificially enhanced because of Galactic obscuration.

In terms of Δ_{rms} , PDEVA-5004 reflects similar results and the same fluctuation patterns seen with c/a . Especially at low redshifts, very low Δ_{rms} heights are found. In Aq-C $^{\alpha}$, despite the low c/a values, we find larger Δ_{rms} values. This is because Δ_{rms} is a dimensional quantity that therefore depends on the overall size of the system at issue. This parameter very clearly decreases as the system evolves: Aq-C $^{\alpha}$ is still settling its size until $T_{\text{uni}} \sim 9$ Gyr. At the last moment of our analysis ($T_{\text{uni}} = 11.5$ Gyr), the Δ_{rms} heights of planes are generally compatible with their observed counterparts at $z = 0$ (except for the “biased” (versus MW) results involving f_{sat} 90% and 70%, and the “non-biased” ones (versus M31) involving a 50%). However, we note that Aq-C $^{\alpha}$ has still 2 Gyr to reach $z = 0$, and the system could still evolve toward a lower Δ_{rms} value.

In the third rows of each panel of Figure 9, we plot the angle formed at each timestep by the normal vector to the plane of satellites and the galaxy’s disk plane (we use Galactic latitude angle, with $|b| \in [0^\circ, 90^\circ]$). We can see that while low f_{sat} curves show a fluctuating behavior with T_{uni} , the fluctuation level decreases as f_{sat} increases, and, finally, almost no fluctuations show up at $f_{\text{sat}} = 90\%$. An important variation in this angle is an indication that the identities of the satellite members of planes with given f_{sat} have changed. Therefore, these results are indicating that the satellite members of the

²⁵ We acknowledge that for an even more accurate comparison, other observational biases could be applied to M31, such as that of the mask of the PAndAS survey, which discovered most of its satellite galaxies (see, e.g., Gillet et al. 2015). Nonetheless, as PAndAS presents a very homogeneous panoramic coverage (see Figure 2 in Conn et al. 2013), we neglect any bias in this work.

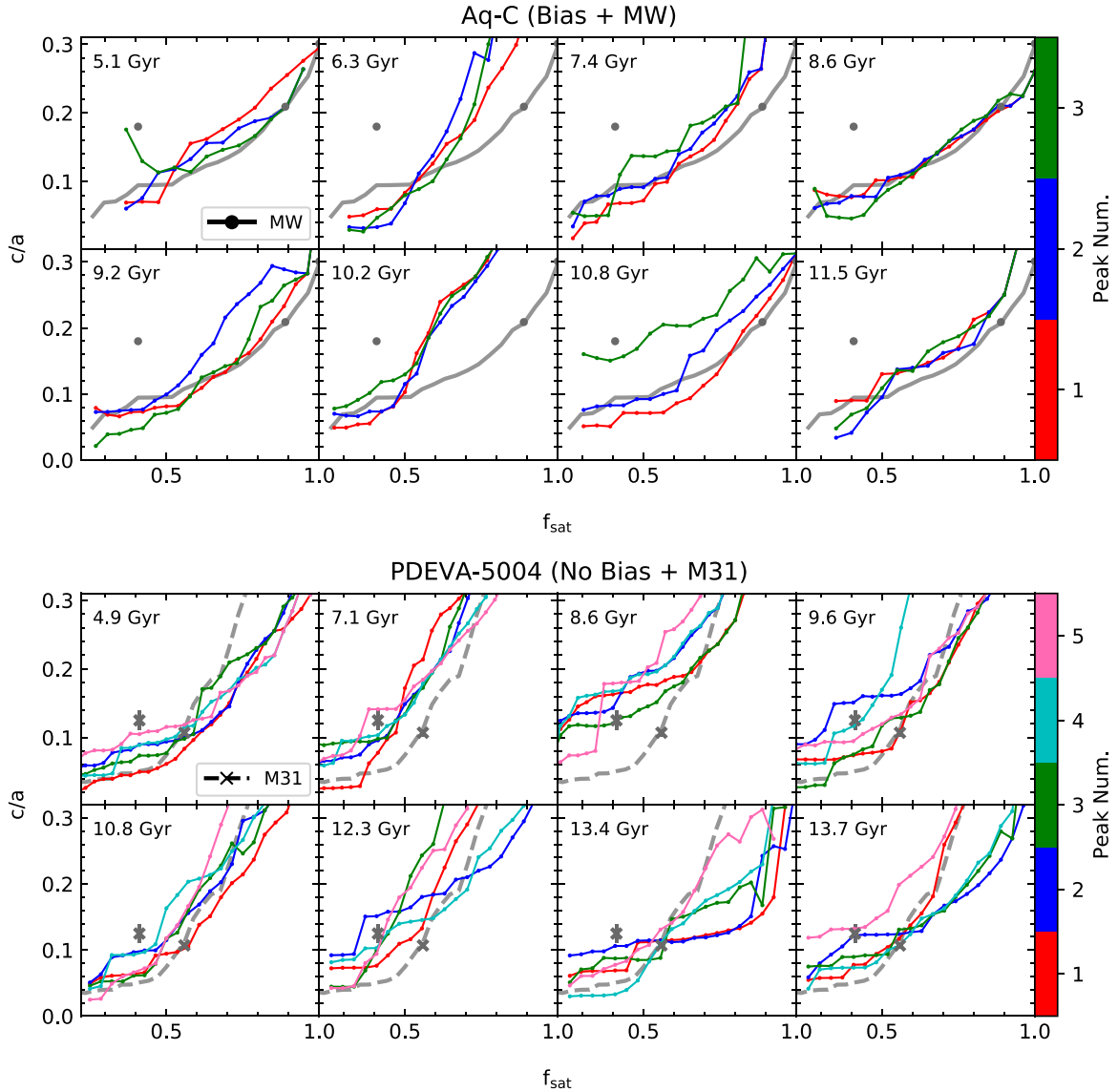


Figure 8. Examples of quality analysis of the planar structures found: the short-to-long axis ratio, c/a , as a function of the fraction of satellites f_{sat} included in the plane at different universe ages. Top set of panels: Aq-C $^\alpha$ vs. the MW. Simulated results include the observational obscuration bias. The gray solid line shows the result for the MW’s main density peak, and the points show the specific values for MW observed planes of satellites mentioned in the literature (i.e., classical, VPOS-3, VPOSall). Bottom set of panels: PDEVA-5004 vs. M31. The gray dashed line shows the result for M31’s main density peak, and the crosses show the specific values for M31 observed planes of satellites mentioned in the literature (i.e., Ibata-Conn-14 and GPoA).

best-quality planes change quite a lot at low f_{sat} , while at $f_{\text{sat}} = 70\%$ or even 50% , these identities are kept to an important extent.

Moreover, at times when c/a reach their minima (and the main peak strength C_1 reach their maxima) the latitude angle in both simulations is small and sometimes close to 0° (except for PDEVA-5004 at $z \approx 0$); that is, satellite planes are nearly perpendicular to the galaxy’s disk. This result suggests that the best quality of satellite planes is, in many cases, reached at time intervals when no (or rather low) galaxy disk torques act upon the satellites belonging to the plane that best fits the satellite set (see, e.g., Danovich et al. 2015).²⁶

Finally, we focus on the distances D_{CG} , or offsets, from the center of the galaxy to the previously presented best planes found in Aq-C $^\alpha$ and PDEVA-5004. Contrary to other plane-identification methods used in simulation studies (see, e.g., Buck et al. 2015; Gillet et al. 2015; Ahmed et al. 2017), in the 4GND plot method, planes are not constrained to pass through the center of the main galaxy. Table 2 shows the offsets to the best planes averaged over the last 1 Gyr with $f_{\text{sat}} = 30\%, 50\%, 70\%, 90\%$. For comparison, the MW VPOS-3 ($f_{\text{sat}} = 24/27 = 88\%$) and M31 GPoA ($f_{\text{sat}} = 19/34 = 56\%$) planes present an offset of 10.4 and 1.3 kpc from the center of the MW and M31, respectively (see Table 3 in Pawlowski et al. 2013). Also, the collection of planes defined by the second peak in the 4GND plot of M31 presents distances D_{CG} between ~ 15 and 35 kpc (see Paper I). Compared to these, the plane offsets measured in both simulations have reasonable values, passing close to the center of the main galaxy.

²⁶ We notice the robustness of our results against the bin size in the 4GND plots. The very thin blue-series lines in Figure 9 show results calculated with half the bin size used to calculate the thicker lines therein. Differences are small and unimportant.

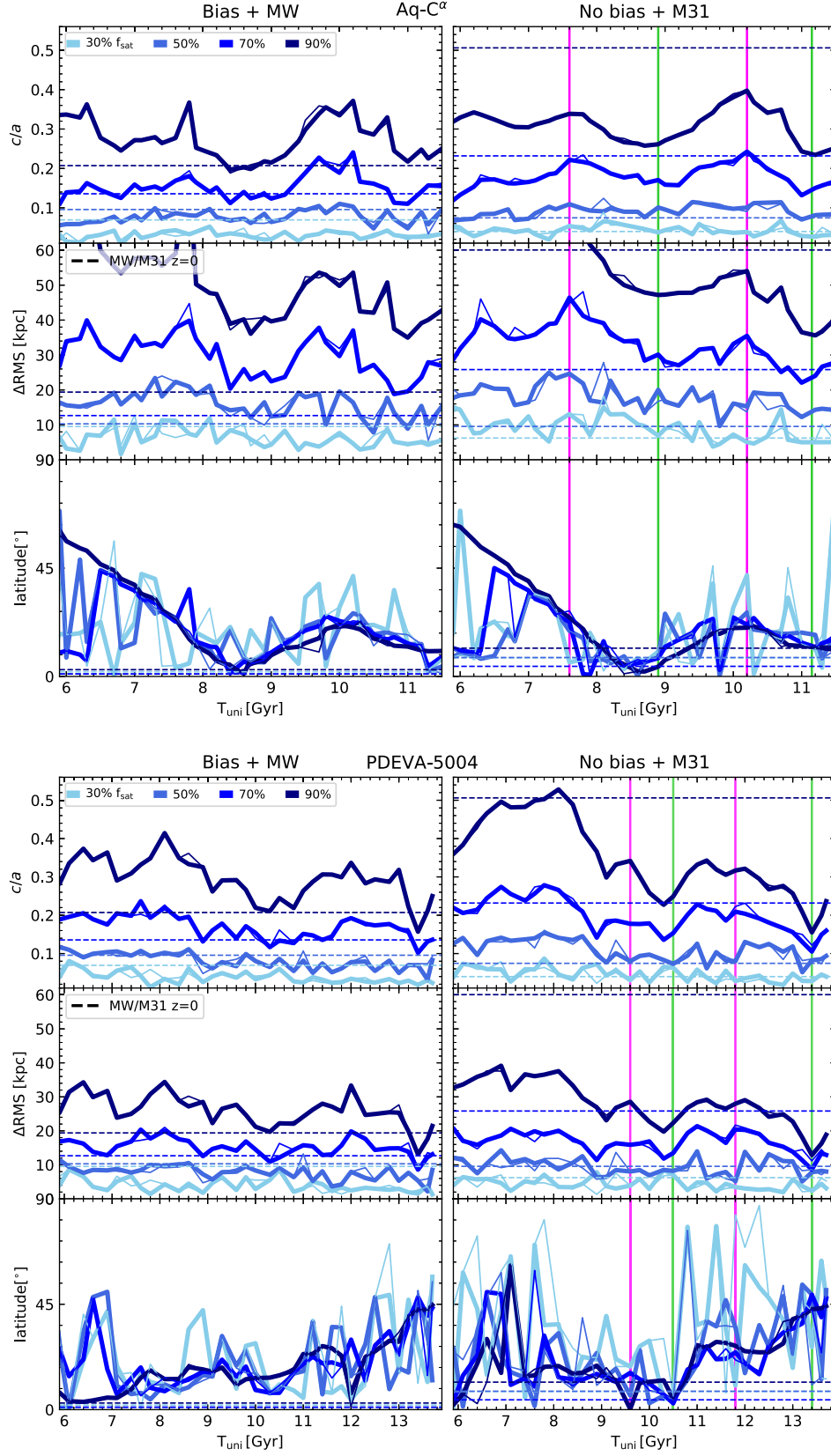


Figure 9. Short-to-long axis ratio c/a , Δrms height and plane of satellites inclination relative to the disk, for the best planes found at each timestep including a fraction $f_{\text{sat}} = 30\%$, 50% , 70% , and 90% of the total number of satellites. Top set of panels: Aq-C α . Bottom set of panels: PDEVA-5004. Left panels: results having applied the observational obscuration bias compared to the MW $z = 0$ values. Right panels: results considering all satellites compared to the M31 $z = 0$ values. Observational values are shown as horizontal dashed lines with the same color code. The very thin lines show the results obtained if we use a two-times-smaller bin size in the 4GND plot method. Table 2 provides the specific parameter values for observations and simulations as averaged in the last 1 Gyr analyzed.

Table 2
Plane Parameters of the Best (i.e., with Lowest c/a) Planes Found in Aq-C $^{\alpha}$ and PDEVA-5004 Including a Fixed Fraction (X%) of Satellites

		Aq-C $^{\alpha}$ (last 1 Gyr)			PDEVA-5004 (last 1 Gyr)			MW ($z = 0$)			M31 ($z = 0$)		
		c/a	Δ_{rms} (kpc)	D_{CG} (kpc)	c/a	Δ_{rms} (kpc)	D_{CG} (kpc)	c/a	Δ_{rms} (kpc)	D_{CG} (kpc)	c/a	Δ_{rms} (kpc)	D_{CG} (kpc)
Bias	30%	0.03	5.30	23.70	0.03	2.46	12.43	0.07 ± 0.001	9.54 ± 0.15	13.23 ± 0.15			
	50%	0.07	12.70	23.77	0.06	6.27	11.96	0.10 ± 0.001	10.31 ± 0.12	14.61 ± 0.11			
	70%	0.14	23.82	23.07	0.15	13.12	6.70	0.14 ± 0.002	12.67 ± 0.18	15.72 ± 0.15			
	90%	0.26	41.05	12.60	0.25	21.11	4.13	0.21 ± 0.002	19.39 ± 0.19	10.46 ± 0.15			
No bias	30%	0.04	6.72	27.92	0.03	3.45	7.76				6.24 ± 0.003	1.21 ± 0.35	6.17 ± 0.57
	50%	0.09	15.46	16.53	0.09	9.22	5.30				9.56 ± 0.002	2.05 ± 0.20	6.29 ± 0.58
	70%	0.17	26.02	17.73	0.15	13.14	5.83				25.83 ± 0.008	3.30 ± 0.83	2.62 ± 2.04
	90%	0.28	40.98	14.42	0.23	19.68	3.92				60.06 ± 0.009	8.02 ± 1.18	10.17 ± 6.27

Note. We show c/a , Δ_{rms} , and D_{CG} . Values are averaged over the last Gyr of the corresponding analysis period (see Figure 9). Fractions for the MW and M31 have been calculated relative to a sample size of $N_{\text{tot}} = 27$ and 34 satellites, respectively. The results shown are the mean values and standard deviations resultant of 100 random realizations of radial distances within the observational uncertainties.

This section reveals that there are indeed preferential planar configurations of satellites at given moments in both Aq-C $^{\alpha}$ and PDEVA-5004 simulations. These planes are thin and highly populated, compatible on average with all characteristics of the observed planar structures found in the MW and in M31 at $z = 0$, and even defining higher-quality planes at particular given times. A rough comparison to the planes of satellites reported in previous studies with zoom-in hydro-simulations that consider a moderately large sample of satellites (e.g., Gillet et al. 2015; Ahmed et al. 2017; Maji et al. 2017) focusing on the c/a , Δ_{rms} , and N_{sat} parameters when provided by the references, strongly suggests that the planes found in Aq-C $^{\alpha}$ and PDEVA-5004 have, within some time intervals, a higher quality and reveal a higher degree of spatial ordering in the satellite distribution. However, this comparison is not completely unbiased due to the different types of simulations and methods for plane-identification used in our study and in others. First, only comparisons between zoom-in simulations that meet the conditions listed in Section 2 of this paper make sense: the dynamical effect of a massive, MW-like disk potential on satellite planes could be an important piece of the puzzle. Second, in the 4GND plot method, no priors are assumed: we do not choose the 11 most massive (or most luminous) satellites among the simulated satellite sample, planes are not required to pass through the center of the host galaxy, or to be thinner than a given Δ_{rms} thickness, etc.

7. Co-orbitation?

One relevant feature of the main plane of satellites observed in the MW is that it presents a relatively high degree of coherent rotation within the plane (see Section 1). This means that the orbital angular momentum vectors (i.e., orbital poles, \mathbf{J}_{orb}) of the constituent satellites are aligned with the normal to the plane. Orbital angular momentum is defined as $\mathbf{J}_{\text{orb}} = \mathbf{r} \times m\mathbf{v}$, where \mathbf{r} and \mathbf{v} are the position and velocity of the center-of-mass of the satellite relative to the center-of-mass of the host disk galaxy.

To study if the satellites included in the high-quality planes detected in the simulations with the extended 4GND plot method are co-orbiting within the plane, we first compute the \mathbf{J}_{orb} vectors of the satellites at each timestep and project them onto the sphere. Then, we quantify the clustering of \mathbf{J}_{orb} vectors around a given \mathbf{n} direction (where \mathbf{n} is the normal vector to a given plane), and we evaluate the fraction of satellites that are kinematically coherent (i.e., co-orbit) within the plane. To this end, we take this direction \mathbf{n} as a reference axis and measure the angular distance DA²⁷ to each individual satellite orbital pole. In order to do this systematically at each timestep, we take, as the reference axis, the normal \mathbf{n} to the best plane (i.e., with lowest c/a) including 50% of the total number of satellites at the respective timestep (see Figure 9).

This is exemplified in Figure 10, for the last timestep analyzed in each simulation. The x-axis shows $1 - \cos(\text{DA})$, therefore ranging from 0 to 1, and the y-axis shows the fraction of the total number of satellites with \mathbf{J}_{orb} enclosed by a certain angular distance DA from the reference axis \mathbf{n} . Since it is only possible to compare to MW data (no proper motion data is available for M31 satellites), the results shown include the observational obscuration bias. Indeed, the dashed line shows the MW case, where we use the latest available data for the

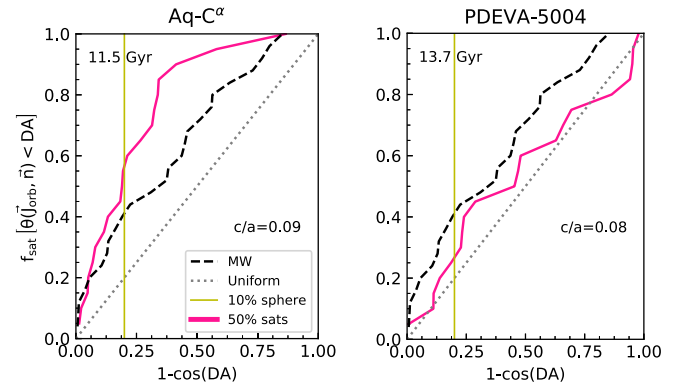


Figure 10. Fraction of satellites with orbital poles enclosed within an angular distance “DA” measured from the normal to the best (i.e., lowest c/a) plane including a fraction $f_{\text{sat}} = 50\%$ of the total number of satellites (see Figure 9). Left panel: Aq-C $^{\alpha}$; right panel: PDEVA-5004; at their last analyzed timesteps. Results include the observational obscuration bias. The dotted line shows the result for a uniform distribution of orbital poles on the sphere, and the dashed line shows the result for the confirmed satellites in the MW, using data from Pawlowski & Kroupa (2013) and Fritz et al. (2018). The yellow vertical line marks an angle of DA = $36^{\circ}87'$, MW satellites with orbital poles enclosed by this angle as measured from the VPOS are considered to co-orbit in Fritz et al. (2018).

confirmed MW satellites (Table 4 in Fritz et al. 2018 calculated from *Gaia*-DR2 data, or alternatively, for the satellites missing there, Table 4 in Pawlowski & Kroupa 2013).²⁸ For comparison, the dotted line illustrates the expectation from a uniform distribution of orbital poles, and a yellow vertical line is depicted at DA = $36^{\circ}87'$, which encloses 10%²⁹ of the sphere surface: this is the angle around the VPOS within which it is considered in Fritz et al. (2018) that MW satellites co-orbit ($\gtrsim 40\%$ of MW satellites co-orbit; see Paper I). We see, at these example timesteps, that while for Aq-C $^{\alpha}$ the fraction of co-orbiting satellites is $\sim 45\%$, for PDEVA-5004 it is much lower and below the MW fraction.

We use the previous analysis to show in Figure 11 the fraction of co-orbiting satellites in the best-quality planes involving $f_{\text{sat}} = 50\%$ and 70% of satellites at each timestep. Figure 11 indicates that while in some cases there is consistency with the MW or even a higher degree of co-orbitation (particularly so in the Aq-C $^{\alpha}$ case), in others, we can see that the fraction of co-orbiting satellites around the \mathbf{n} direction is very low, despite these directions defining the highest-quality spatial planar arrangements found at that moment. Moreover, the abrupt changes in the fraction of co-orbiting satellites from one timestep to another is a consequence of the different identities of satellites constituting the best-quality planes at close times. Important differences in the fraction of co-orbiting satellites are also found between the 50% versus the 70% case at a same timestep: an indication that while one of the planes shows a kinematical coherence, the other (corresponding to a different density peak) does not.

In this respect, it is interesting to compare the times when the C_1 maxima and minima occur (see Figure 6) with a measure of the co-orbitation of the involved satellites. Figure 11 indicates that, in both simulations, vertical green (magenta) bands do not

²⁷ Note that we do not differentiate between co-rotation or counter-rotation with the disk of the central galaxy; therefore, DA can be a maximum of 90° .

²⁸ We consider in this plot a final sample of 25 out of 27 MW satellites as there is no published proper motion for Canis Major or Bootes III.

²⁹ When no distinction is made if satellites co- or counter-rotate with respect to the disk, this angle intercepts a 20% of the sphere surface.

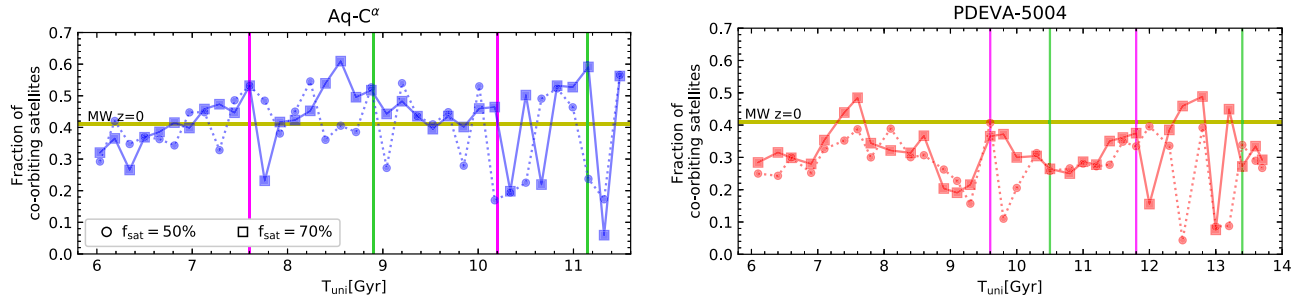


Figure 11. The fraction of co-orbiting satellites in the best-quality planes involving $f_{\text{sat}} = 50\%$ and 70% of satellites at each timestep, vs. the universe age. Left panel: Aq-C $^\alpha$; right panel: PDEVA-5004. The observational obscuration bias has been applied to obtain these results. Co-orbiting satellites are those with orbital angular momentum vectors within $36^\circ 87'$ around the normal to the plane. Green (magenta) vertical lines mark time intervals where the main peak strength C_1 reaches maximum (minimum) values (see Figure 6). An horizontal line marks the corresponding fraction of co-orbiting satellites in the MW at $z = 0$ according to Figure 10.

always correspond to a high (low) fraction of co-orbiting satellites.

These results imply that, in general, the best planes in positions found with the 4GND plot method may just be fortuitous spatial alignments of satellites and, therefore, *transient* structures (see also Gillet et al. 2015; Buck et al. 2016; Ahmed et al. 2017; Fernando et al. 2017; Shao et al. 2019).

In order to efficiently detect kinematically coherent planes of satellites in these simulations, a deeper and more precise analysis of the persistence or not of good-quality planes of satellites across time is needed, which demands using the full six-dimensional phase-space information of satellites for plane searching. In a forthcoming paper (Santos-Santos et al. 2020, in preparation; Paper III), a new method is developed to address the plane of satellites persistence issue.

8. Discussion

8.1. Quality Analysis in Terms of N_{sat}

The use of normalized quantities such as C_p in the peak analysis makes results independent of the total sample sizes N_{tot} . In this line, for the quality analysis of planes, we have used f_{sat} , an N_{tot} independent quantity allowing for a clean comparison of samples of different size.

The analysis has been repeated through c/a versus N_{sat} (i.e., the absolute number of satellites instead of its fraction f_{sat}). To this end, following Riley et al. (2019), the total number of satellites N_{tot} in simulations and observational samples has been matched at each timestep. In particular, for each simulated satellite, we compare its distance to its host with the galactocentric distance of all MW (or M31) satellites, and we select from these the optimal match (without replacement). In this way, a sample of observed galaxies is built with $N_{\text{tot,obs}} = N_{\text{tot,sim}}$. This matching is needed for a proper comparison, because in simulations, N_{tot} depends on time (see Figure 1), while in the $z = 0$ satellite system of the MW and M31, N_{tot} is a fixed number. Results are qualitatively the same as those obtained in terms of f_{sat} . We note that without N_{tot} -matching, results on consistency with observations can be easily obtained from Figure 8 by translating the observational curve rightwards an amount $\frac{N_{\text{tot,obs}}}{N_{\text{tot,sim}}}$. Inconsistency or not would be due to the value of this number, with possible inconsistencies resulting from the different size of simulated and observed satellite samples; which justifies why this exercise is needed.

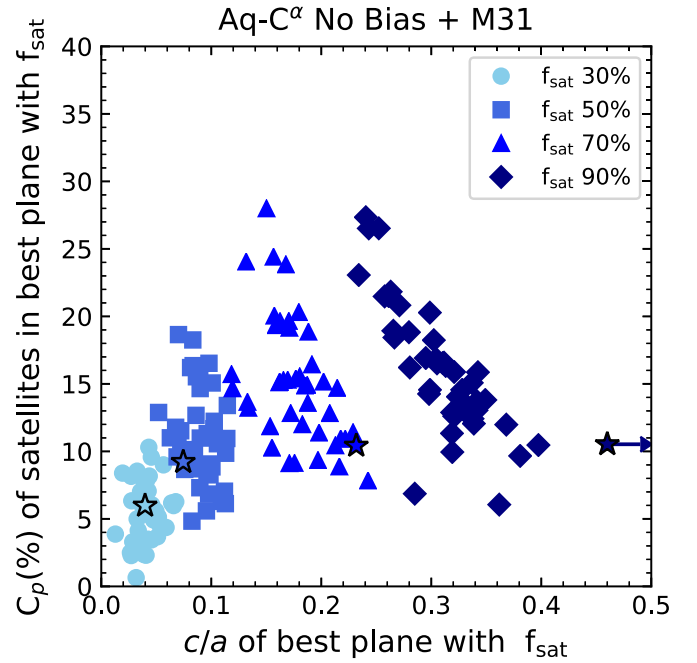


Figure 12. Main peak strength considering different f_{sat} , $C_1(f_{\text{sat}})$, vs. the short-to-long axis ratio c/a of the best plane with f_{sat} satellites. Different blue shades stand for different f_{sat} values, with the same color coding as in Figure 9. Stars show the M31 observational values at $z = 0$.

8.2. Can the Peak Strength Be Used as a Measure of Plane Quality?

The green (magenta) vertical lines in Figure 9 mark the T_{uni} values where the main peak strength C_1 have maxima (minima) in Figure 6. A very relevant result is that maxima occur at T_{uni} values when the c/a value of the $f_{\text{sat}} = 90\%$ curve is at a minimum, that is, when the quality shows a maximum. And conversely, the magenta lines are close to maxima of c/a , that is, bad qualities. To find out whether this behavior keeps at other f_{sat} values and whether we can use peak strength C_p to measure quality at given f_{sat} , we have calculated the main peak strength as a function of f_{sat} , i.e., $C_1(f_{\text{sat}})$,³⁰ and compared it to the c/a of the respective best plane found with same f_{sat} at that timestep.

³⁰ $C_1(f_{\text{sat}})$ is calculated as explained in Section 4.2.1, except that we do not sum over all of the satellites. We stop summing up when a given satellite fraction, f_{sat} , is reached. For example $C_1(50\%)$ in the upper panel of Figure 5, would just involve the 10 first satellites (which are ordered by decreasing $C_{1,s}$).

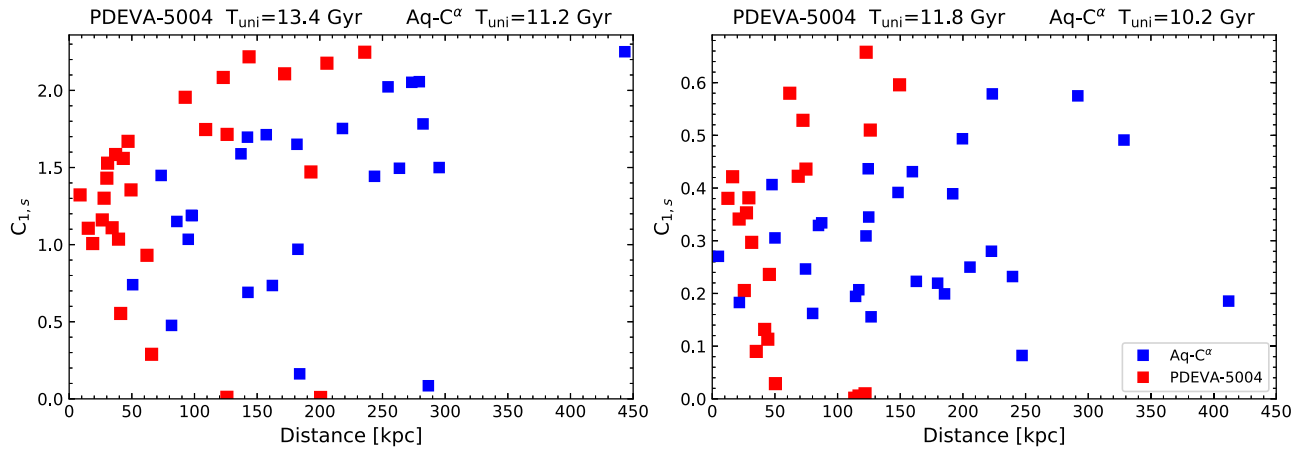


Figure 13. The contribution of each satellite s to the main peak ($C_{1,s}$), vs. the radial distance of the satellite to the center of its host disk galaxy. The left panel corresponds to moments in the evolution of the respective simulations where C_1 is very high ($C_1 \sim 33\%$ for both simulations, see Figure 6). The right panel corresponds to moments when C_1 is relatively lower ($C_1 \sim 10\%$ for both simulations).

As an example, in Figure 12, we see for Aq- C^α (“no bias” case, where all satellites are considered) that a correlation exists at all f_{sat} . There is, however, an important dispersion particularly at low f_{sat} . The same qualitative behavior is found for PDEVA-5004, and in the “bias” case for both simulations. Therefore, we conclude that although indicative, the peak strength C_p is not an accurate enough measure of the quality of its collection of planes.

It is interesting to note that, when absolute numbers of satellites (N_{sat} , without N_{tot} -matching) are used instead of fractions (f_{sat}) to analyze quality, the correlations shown in Figure 12 disappear or are very weak. This is a clear indication that the mutual relationships between quality and peak strength are best manifested when the quality analysis is made in terms of f_{sat} , rather than N_{sat} .

8.3. Does Radial Compactness of the Satellites Affect Quality?

In this subsection, we analyze the possible correspondence between plane quality and the radial compactness of a satellite system.

Figure 13 shows $C_{1,s}$, i.e., the contribution of each satellite s to the main peak, versus the radial distance of the satellite to the center of its host disk galaxy. The left panel corresponds to moments in the evolution of the respective simulations where C_1 is very high ($C_1 \sim 33\%$ for both simulations, see Figure 6), while the right panel corresponds to moments when C_1 is relatively lower ($C_1 \sim 10\%$ for both simulations).

There is a clear correlation for satellites that are farther away to contribute more to a given peak (higher $C_{p,s}$). At fixed C_1 , this effect is slightly intensified in systems where there is an agglomeration of satellites at shorter distances (more compact systems). This is expressed with systematically higher $C_{p,s}$ values but a similar slope in the correlation. See, for example, the left panel, where in PDEVA-5004—which is a more radially compact system than Aq- C^α —satellites show higher $C_{1,s}$ values.

On the other hand, if we compare PDEVA-5004 satellites in the left and right panels (i.e., timesteps where the C_1 is different but there is a similar radial compactness of satellites), we see that the $C_{1,s}$ is dramatically lower for the system with lower C_1 .

Thus, while a spatial configuration with several central satellites and a few farther away ones (i.e., a compact system) increases the probability of four-galaxy-normals pointing in a similar direction, from these results, it is clear that this geometric effect is not the driving reason for having a high clustering of four-galaxy-normals or high C_1 . We conclude that radial compactness does not have a determinant role in setting plane quality.

9. Summary and Conclusions

To address the so-called “Planes of satellites problem” (see discussion and references in Section 1), we have applied, for the first time, the four-galaxy-normal density (4GND) plot method (Pawlowski et al. 2013) to hydrodynamical simulations. An extension of the method, sketched in Santos-Santos et al. (2019, Paper I), is presented in detail and discussed in this paper.

Our choice has been zoom-in simulations because the higher resolution (as compared to large-volume simulations) one can reach by using this method allows MW-type galaxies with massive, extended disks to form, so that the dynamical effects of a live disk potential on satellite planes (torques, destruction of satellites on radial orbits) can be accounted for. Another advantage of zoom-in simulations is that they offer the possibility that these disk galaxies are surrounded by a number of resolved satellites high enough to be comparable to current samples of confirmed MW or M31 satellites (~ 30 , see the McConnachie 2012 database). Well-behaved disks and a large enough satellite number are conditions not recovered, at the moment, by larger-volume cosmological simulations.

The extension to the 4GND method is designed to identify, systematically catalog, and study in detail the quality of the predominant spatial planar configurations of satellites revealed by over-densities in the 4GND plots. It allows us to extract high-quality planes out of the number of combinations we can form with N_{sat} satellites from a sample of size N_{tot} , with a low computational cost. Quality is evaluated through the outputs of a Tensor of Inertia analysis (ToI; Metz et al. 2007) using a normal-regression fitting technique. Quantitatively, planes (i.e., best-fitting solutions with high medium-to-long axes ratio, $b/a \sim 1$), have a good quality if they are populated relative to

the sample size N_{tot} (high $f_{\text{sat}} \equiv N_{\text{sat}}/N_{\text{tot}}$) and thin (low short-to-long axis ratio c/a). Being a two-parameter notion, the quality of two or more planes can be compared if one has lower c/a and higher f_{sat} than another, or if either f_{sat} or c/a are constant.

Density peaks are determined by local, isolated maxima in the 4GND plot. We have defined the peak strength, C_p , as the % of weighted four-galaxy-normals within 15° of the peak center. Peaks can be compared with each other across times through their respective strengths and with those of the MW and M31 satellite systems as well. Different satellites contribute differently to C_p . The satellite s contribution to peak p (i.e., $C_{p,s}$) is defined as the normalized weighted count of s contributions to four-galaxy-normals placed within 15° of the peak center. Satellites are ordered by decreasing $C_{p,s}$ to a peak, and a plane is fitted to groups of increasing N_{sat} satellites following this order. This yields, for each density peak, a *collection* of planes.

In Paper I, we report on the application of the extended 4NDP method to the confirmed MW and M31 satellites. The method extension reveals a richer planar structure (i.e., a *collection* of planes rather than a unique plane per density peak), therefore allowing us to find planes of satellites around the MW and M31 with higher qualities than those previously reported with a given N_{sat} . We find a second populated, high-quality plane around M31. Another important result, in view of the narrow range of satellite mass distributions that can be currently afforded in hydrodynamical simulations, is that satellite mass plays no role in determining a satellite's membership or lack thereof to the respective best-quality planes. This enables us to perform, through the extended 4GND plot method, comparisons between results from simulations and MW/M31 data (where the satellite mass range is wider).

In this paper, we present results of a detailed search for positional planar structures of satellites in two different (initial conditions, sub-grid modeling and numerical approaches) zoom-in cosmological hydrodynamical simulations of isolated disk galaxies that resemble the MW: Aq-C $^\alpha$ and PDEVA-5004. They meet the conditions of having a central host galaxy with an extended disk, an overall quiet merger history after virialization, a numerous (~ 30) satellite population, and more than 50 baryonic particles per satellite. No other particular selection criteria have been applied. In particular, we focus our analysis on the best-quality planes with a given satellite fraction f_{sat} . We compare to the best-quality planes found in the MW and M31 systems at $z = 0$ with given f_{sat} , i.e., those found from their so-called respective #1 Peaks as defined in Paper I.

The analysis goes from the halo virialization time to low redshifts, along the slow phase of host mass assembly. The halo mass-growth histories of Aq-C $^\alpha$ and PDEVA-5004 present, per usual, two phases and no other relevant particularities. Satellite samples have been identified at $z = 0.5$. The number of satellites N_{tot} considered in the analysis at each timestep varies. In particular, satellites are not considered when they are orbiting beyond 450 kpc from the center of the host galaxy, after they have been accreted by the host galaxy, or if they are within the avoidance volume when we apply an observational obscuration bias to compare with the MW. The distributions of satellite radial distances to the center-of-mass of the main galaxy vary with time. Expressed in terms of f_{sat} , Aq-C $^\alpha$'s (PDEVA-5004's) radial distance distribution is very close to that of M31 (the MW), at particular times.

The analysis of density peaks reveals that C_1 varies with T_{uni} , with time intervals where it is comparable or even higher than

the MW value, $C_{1,\text{MW}} = 22.91\% \pm 0.26\%$. The number of strong peaks changes with time too, reaching values of 1–3 at most and only in those time intervals when C_1 is high. In this regard, density peaks found in simulations are similar to those observed in strength and number during given time intervals.

We find *planar* (i.e., high b/a and low c/a) configurations of satellites at all studied timesteps in both simulations. Indeed, no filamentary (i.e., $b/a \ll 1$) configurations have been detected in the period analyzed. The extended 4GND plot method allows us to identify, at many timesteps in both simulations, (i) planes of satellites with qualities that are compatible with the observed ones at $z = 0$ including a specific f_{sat} , and also in some cases, (ii) planar structures that are compatible with the observed ones for *all* f_{sat} .

We study the best-quality planes (i.e., with lowest c/a) including a fixed f_{sat} found at each timestep. In both simulations, their c/a values change with time, independently of the f_{sat} considered. Planes compatible with the observed ones in the MW and M31 at $z = 0$ are found at different timesteps or time intervals. Interestingly, these timesteps turn out to coincide with the time intervals where C_1 shows maxima. More specifically, a correlation has been found between C_1 and c/a at fixed f_{sat} but with important dispersion. Therefore, C_1 can be used as an estimation of plane quality but not to accurately measure it. Another interesting result is that the highest-quality planes are often close to perpendicular to the host disk plane. Ideally, perpendicular planes of satellites would not suffer torques from the disk.

No new information on quality is provided when using Δr_{ms} as a quality indicator once the satellite system size has settled. And no new information on quality (in terms of comparison to observational satellite planes) is obtained either when using absolute satellite numbers N_{sat} instead of fractions $f_{\text{sat}} = \frac{N_{\text{sat}}}{N_{\text{tot}}}$, with $N_{\text{tot,obs}}$ matched to $N_{\text{tot,sim}}$.

Interestingly, when the observational obscuration bias is applied, slightly higher peak strengths are measured, as well as somewhat lower c/a and Δr_{ms} values than in the non-biased case. The is due to solid-angle restrictions in the biased case, where satellites orbiting at low latitudes are neglected.

No clear, conclusive signal on the correspondence between plane quality and the radial compactness of a satellite system has been detected in this work. While satellites at larger distances from the host galaxy provide somewhat higher $C_{1,s}$ than nearby ones, the spatial satellite configurations that show the highest C_1 (and therefore highest overall plane qualities) are not those most compact.

We have further investigated if satellites composing the high-quality planes of satellites found with the extended 4GND plot positional method present a common orbitation within the plane they describe. We find that, in some cases, the fraction of co-orbiting satellites is very low, which we interpret as a sign of these positional planes consisting partly of interlopers: satellites that fall within the plane accidentally. Therefore, planes found with this method based on positional data in general do not constitute a kinematic unit and, in some cases, could be non-persistent in time. Because of this, the search for the physical reasons favoring or destroying positionally detected high-quality planes cannot be meaningfully addressed. In Santos-Santos et al. (2020, in preparation; Paper III), a new methodology is introduced where the full six-dimensional phase-space information of satellites is used, leading to the determination of persistent, kinematically coherent planes of satellites in both simulations.

Summing up, the application of the 4GND plot method (Pawlowski et al. 2013) with its extension presented here, (see also Santos-Santos et al. 2019, Paper I) to two zoom-in hydrodynamical simulations of disk galaxies, leads us to the following conclusions:


1. Satellites are organized in planar configurations (not filamentary, i.e., $b/a \sim 1$) in both the Aq-C $^{\alpha}$ and PDEVA-5004 simulations, at all timesteps analyzed. The plane short-to-long axis ratio c/a , and the plane population (f_{sat}) measure plane quality.
2. The strengths of the strongest peaks in the 4GND plots (C_1) vary with time. Their values are, at all times, higher than that of the strongest peak in M31, and along given periods, also higher than that of the MW.
3. The c/a ratios of the thinnest planes found including a fixed f_{sat} vary with cosmic time, and during some periods, reach a high quality. Along these good-quality time intervals, they are compatible with the observed planar structures found in the MW and in M31 at $z = 0$. These periods coincide with those when C_1 reaches maximum values. The timescale for these plane quality changes is ~ 0.5 –1 Gyr.
4. c/a and C_1 show correlations with increasing dispersion as f_{sat} decreases. C_1 can be used as a quality indicator, but not to accurately measure it.
5. The application of the observational obscuration bias enhances plane quality, either measured by peak strength or plane thickness.
6. The orientations of planes of satellites with respect to the disk of their host galaxy change with time. In most cases, planes are close to perpendicular to the disk during periods of good quality.
7. The compactness of the distribution of satellite-host radial distances does not have a driving role at setting the quality of plane of satellites.
8. In agreement with previous findings, the fraction of co-orbiting satellites found in high-quality positionally detected planes is rather low, suggesting that these planes do not represent a kinematic unit and are not persistent in time.
9. The plane persistence issue in observations and simulations cannot be properly addressed unless the full six-dimensional phase-space information is considered. Such a methodology will be developed in a forthcoming paper.

The general conclusion of this paper is that even if two galaxies do not make a statistical sample, the fact that these two so different MW-type disk galaxies (whose selection method is quite general) do have, at given time intervals, high-quality positional satellite planes, would suggest that these planes can be expected not to be infrequent in Λ CDM L^* disk galaxies in periods when they are free of major merger events in their assembly history.

We thank the anonymous referee for useful comments and suggestions that have helped improve this work. This work was supported through MINECO/FEDER (Spain) AYA2012-31101, AYA2015-63810-P, and MICINN/FEDER (Spain) PGC2018-094975-C21 grants. I.S.S. acknowledges support by the Arthur B. McDonald Canadian Astroparticle Physics Research Institute. This work used the Ragnar cluster funded by Fondecyt 1150334 and Universidad Andrés Bello, and

Geryon cluster (Pontificia Universidad de Chile). We used a version of Aq-C-5 that is part of the CIELO Project run in Marenostrum (Barcelona Supercomputer Centre). This project has received funding from the European Union’s Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement No. 734374 (LACEGAL-RISE). I.S.S. acknowledges funding from the same grant for a secondment at the Astrophysics group of Univ. Andrés Bello (Santiago, Chile), and from the Univ. Autónoma de Madrid for a stay at the Leibniz Institut für Astrophysik Potsdam (Germany). I.S.S. thanks Dr. Patricia Tissera and Dr. Noam Libeskind for kindly hosting her.

ORCID iDs

Lucas Bignone  <https://orcid.org/0000-0002-4746-1627>
 Francisco Martínez-Serrano  <https://orcid.org/0000-0003-1579-7182>
 M. Ángeles Gómez-Flechoso  <https://orcid.org/0000-0002-0692-7187>
 Patricia B. Tissera  <https://orcid.org/0000-0001-5242-2844>
 Arturo Serna  <https://orcid.org/0000-0002-6311-2523>

References

- Agertz, O., Teyssier, R., & Moore, B. 2011, *MNRAS*, **410**, 1391
 Ahmed, S. H., Brooks, A. M., & Christensen, C. R. 2017, *MNRAS*, **466**, 3119
 Bahl, H., & Baumgardt, H. 2014, *MNRAS*, **438**, 2916
 Belokurov, V., Erkal, D., Evans, N. W., Koposov, S. E., & Deason, A. J. 2018, *MNRAS*, **478**, 611
 Boylan-Kolchin, M., Springel, V., White, S. D. M., Jenkins, A., & Lemson, G. 2009, *MNRAS*, **398**, 1150
 Buck, T., Dutton, A. A., & Macciò, A. V. 2016, *MNRAS*, **460**, 4348
 Buck, T., Macciò, A. V., & Dutton, A. A. 2015, *ApJ*, **809**, 49
 Buck, T., Macciò, A. V., Dutton, A. A., Obreja, A., & Frings, J. 2019, *MNRAS*, **483**, 1314
 Bullock, J. S., & Boylan-Kolchin, M. 2017, *ARA&A*, **55**, 343
 Cautun, M., Bose, S., Frenk, C. S., et al. 2015, *MNRAS*, **452**, 3838
 Conn, A. R., Lewis, G. F., Ibata, R. A., et al. 2013, *ApJ*, **766**, 120
 Cramér, H. 1999, *Mathematical Methods of Statistics* (PMS-9) (Princeton, NJ: Princeton Univ. Press), <http://www.jstor.org/stable/j.ctt1bpm9r4>
 Danovich, M., Dekel, A., Hahn, O., Ceverino, D., & Primack, J. 2015, *MNRAS*, **449**, 2087
 Doménech-Moral, M., Martínez-Serrano, F. J., Domínguez-Tenreiro, R., & Serna, A. 2012, *MNRAS*, **421**, 2510
 Domínguez-Tenreiro, R., Obreja, A., Brook, C., Martínez-Serrano, F. J., & Serna, A. 2016, in *The Zeldovich Universe: Genesis and Growth of the*, Vol. 308 ed. R. van de Weygaert et al., **398**
 Domínguez-Tenreiro, R., Obreja, A., Brook, C. B., et al. 2015, *ApJL*, **800**, L30
 Domínguez-Tenreiro, R., Obreja, A., Brook, C. B., Martínez-Serrano, F. J., & Serna, A. 2017, *ApJ*, **846**, 72
 Domínguez-Tenreiro, R., Obreja, A., Granato, G. L., et al. 2014, *MNRAS*, **439**, 3868
 Fernando, N., Arias, V., Guglielmo, M., et al. 2017, *MNRAS*, **465**, 641
 Fritz, T. K., Battaglia, G., Pawlowski, M. S., et al. 2018, *A&A*, **619**, A103
 Gaia Collaboration, Helmi, A., van Leeuwen, F., et al. 2018, *A&A*, **616**, A12
 Garaldi, E., Romano-Díaz, E., Borzyszkowski, M., & Porciani, C. 2018, *MNRAS*, **473**, 2234
 Garrison-Kimmel, S., Wetzel, A., Bullock, J. S., et al. 2017, *MNRAS*, **471**, 1709
 Gillet, N., Ocvirk, P., Aubert, D., et al. 2015, *ApJ*, **800**, 34
 Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, *Natur*, **563**, 85
 Ibata, R. A., Lewis, G. F., Conn, A. R., et al. 2013, *Natur*, **493**, 62
 Koch, A., & Grebel, E. K. 2006, *AJ*, **131**, 1405
 Kroupa, P., Theis, C., & Boily, C. M. 2005, *A&A*, **431**, 517
 Kunkel, W. E., & Demers, S. 1976, *RGOB*, **182**, 241
 Libeskind, N. I., Cole, S., Frenk, C. S., Okamoto, T., & Jenkins, A. 2007, *MNRAS*, **374**, 16
 Libeskind, N. I., Frenk, C. S., Cole, S., Jenkins, A., & Helly, J. C. 2009, *MNRAS*, **399**, 550
 Lynden-Bell, D. 1976, *MNRAS*, **174**, 695
 Maji, M., Zhu, Q., Marinacci, F., & Li, Y. 2017, arXiv:1702.00497

- Martínez-Serrano, F. J., Serna, A., Doménech-Moral, M., & Domínguez-Tenreiro, R. 2009, *ApJ*, **705**, L133
- Martínez-Serrano, F. J., Serna, A., Domínguez-Tenreiro, R., & Mollá, M. 2008, *MNRAS*, **388**, 39
- McConnachie, A. W. 2012, *AJ*, **144**, 4
- McConnachie, A. W., & Irwin, M. J. 2006, *MNRAS*, **365**, 902
- Metz, M., Kroupa, P., & Jerjen, H. 2007, *MNRAS*, **374**, 1125
- Metz, M., Kroupa, P., & Libeskind, N. I. 2008, *ApJ*, **680**, 287
- Müller, O., Pawłowski, M. S., Jerjen, H., & Lelli, F. 2018, *Sci*, **359**, 534
- Müller, O., Scalera, R., Binggeli, B., & Jerjen, H. 2017, *A&A*, **602**, A119
- Obreja, A., Domínguez-Tenreiro, R., Brook, C., et al. 2013, *ApJ*, **763**, 26
- Pawłowski, M. S. 2016, *MNRAS*, **456**, 448
- Pawłowski, M. S. 2018, *MPLA*, **33**, 1830004
- Pawłowski, M. S., Bullock, J. S., Kelley, T., & Famaey, B. 2019, *ApJ*, **875**, 105
- Pawłowski, M. S., Famaey, B., Jerjen, H., et al. 2014, *MNRAS*, **442**, 2362
- Pawłowski, M. S., Famaey, B., Merritt, D., & Kroupa, P. 2015, *ApJ*, **815**, 19
- Pawłowski, M. S., & Kroupa, P. 2013, *MNRAS*, **435**, 2116
- Pawłowski, M. S., Kroupa, P., & Jerjen, H. 2013, *MNRAS*, **435**, 1928
- Pawłowski, M. S., Pflamm-Altenburg, J., & Kroupa, P. 2012, *MNRAS*, **423**, 1109
- Pedrosa, S. E., & Tissera, P. B. 2015, *A&A*, **584**, A43
- Riley, A. H., Fattahi, A., Pace, A. B., et al. 2019, *MNRAS*, **486**, 2679
- Riley, A. H., & Strigari, L. E. 2020, *MNRAS*, **494**, 983
- Santos-Santos, I., Domínguez-Tenreiro, R., & Pawłowski, M. 2019, *MNRAS*, submitted
- Sawala, T., Pihajoki, P., Johansson, P. H., et al. 2017, *MNRAS*, **467**, 4383
- Scannapieco, C., Tissera, P. B., White, S. D. M., & Springel, V. 2005, *MNRAS*, **364**, 552
- Scannapieco, C., Tissera, P. B., White, S. D. M., & Springel, V. 2006, *MNRAS*, **371**, 1125
- Serna, A., Domínguez-Tenreiro, R., & Sáiz, A. 2003, *ApJ*, **597**, 878
- Shao, S., Cautun, M., & Frenk, C. S. 2019, *MNRAS*, **488**, 1166
- Springel, V., Wang, J., Vogelsberger, M., et al. 2008, *MNRAS*, **391**, 1685
- Springel, V., White, S. D. M., Jenkins, A., et al. 2005, *Natur*, **435**, 629
- Springel, V., Yoshida, N., & White, S. D. M. 2001, *NewA*, **6**, 79
- Tully, R. B., Libeskind, N. I., Karachentsev, I. D., et al. 2015, *ApJL*, **802**, L25
- Wang, J., Frenk, C. S., & Cooper, A. P. 2013, *MNRAS*, **429**, 1502
- Welker, C., Dubois, Y., Pichon, C., Devriendt, J., & Chisari, N. E. 2018, *A&A*, **613**, A4
- Wetzel, A. R., Hopkins, P. F., Kim, J.-h., et al. 2016, *ApJL*, **827**, L23
- Woo, J., Courteau, S., & Dekel, A. 2008, *MNRAS*, **390**, 1453
- York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, *AJ*, **120**, 1579