



HAL
open science

A 2-3 billion year old major merger paradigm for the Andromeda galaxy and its outskirts

F. Hammer, Y. B. Yang, J. L. Wang, R. Ibata, H. Flores, M. Puech

► **To cite this version:**

F. Hammer, Y. B. Yang, J. L. Wang, R. Ibata, H. Flores, et al.. A 2-3 billion year old major merger paradigm for the Andromeda galaxy and its outskirts. *Monthly Notices of the Royal Astronomical Society*, 2018, 475, pp.2754 - 2767. 10.1093/mnras/stx3343 . hal-02296388

HAL Id: hal-02296388

<https://hal.science/hal-02296388>

Submitted on 10 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A 2–3 billion year old major merger paradigm for the Andromeda galaxy and its outskirts

F. Hammer,^{1★} Y. B. Yang,^{1★} J. L. Wang,^{1,2} R. Ibata,^{3★} H. Flores¹ and M. Puech¹

¹GEPI, Observatoire de Paris, PSL Research University, CNRS, Place Jules Janssen, F-92190 Meudon, France

²NAOC, Chinese Academy of Sciences, A20 Datun Road, 100012 Beijing, PR China

³Observatoire Astronomique, Université de Strasbourg, CNRS, 11, rue de l'Université, F-67000 Strasbourg, France

Accepted 2017 December 19. Received 2017 November 6; in original form 2017 September 7

ABSTRACT

Recent observations of our neighbouring galaxy M31 have revealed that its disc was shaped by widespread events. The evidence for this includes the high dispersion ($V/\sigma \leq 3$) of stars older than 2 Gyr, and a global star formation episode, 2–4 Gyr ago. Using the modern hydrodynamical code, GIZMO, we have performed 300 high-resolution simulations to explore the extent to which these observed properties can be explained by a single merger. We find that the observed M31 disc resembles models having experienced a 4:1 merger, in which the nuclei coalesced 1.8–3 Gyr ago, and where the first passage took place 7–10 Gyr ago at a large pericentre distance (32 kpc). We also show that within a family of orbital parameters, the Giant Stream (GS) can be formed with various merger mass ratios, from 2:1 to 300:1. A recent major merger may be the only way to create the very unusual age–dispersion relation in the disc. It reproduces and explains the long-lived 10 kpc ring, the widespread and recent star formation event, the absence of a remnant of the GS progenitor, the apparent complexity of the 3D spatial distribution of the GS, the NE and G Clumps and their formation process, and the observed slope of the halo profile. These modelling successes lead us to propose that the bulk of the substructure in the M31 halo, as well as the complexity of the inner galaxy, may be attributable to a single major interaction with a galaxy that has now fully coalesced with Andromeda.

Key words: galaxies: formation – galaxies: haloes – galaxies: individual: M31 – galaxies: kinematics and dynamics – Local Group.

1 INTRODUCTION

For a very long time, the M31 galaxy has been considered as a Milky Way (MW) twin since both galaxies share a similar morphological classification (Sb and Sbc, respectively), possess a large bar in their centre, and have a prominent galactic disc. However, recent large surveys (Pan-Andromeda Archaeological Survey (PAndAS): Ibata et al. 2014, Panchromatic Hubble Andromeda Treasury (PHAT): Gilbert et al. 2009, Spectroscopic and Photometric Landscape of Andromeda's Stellar Halo survey: Dalcanton et al. 2012) have provided an enormous wealth of data probing the M31 halo, disc, and bulge, respectively. The ensuing findings emphasize how different the two galaxies that dominate the Local Group are:

(i) the classical bulge of M31 is much more prominent than the pseudo-bulge of the MW (Courteau et al. 2011);

(ii) the M31 galaxy possesses a massive ring of star formation at 10 kpc, which is especially clear in the mid-infrared (Barmby et al. 2006; Gordon et al. 2006);

(iii) while only ~ 10 per cent of the MW's stars are in a thick disc (Jurić et al. 2008), almost the whole M31 disc is thick and presents a far steeper stellar age–velocity dispersion relation (Dorman et al. 2015);

(iv) the M31 disc within 20 kpc shows a global, prominent star formation episode, 2–4 Gyr ago (Bernard et al. 2015a; Williams et al. 2015);

(v) the M31 halo is simply exceptional by the variety of streams (Ibata et al. 2014), and includes the Giant Stream (GS, Ibata et al. 2001) which dominates the star counts at projected distances between ~ 30 and 120 kpc.

The MW is known to be a relatively quiet galaxy without any recent, major merger event, while M31 has properties quite close to an average, representative local spiral galaxy (Hammer et al. 2007). Calculations from the Λ cold dark matter (Λ CDM) theory predict approximately one major merger in the past history of galaxies in the MW–M31 range of mass (Stewart et al. 2008), which is confirmed by observations (see Rodrigues et al. 2017 for a recent update). What then could be the cause of such considerable differences in

* E-mail: francois.hammer@obspm.fr (FH); yanbin.yang@obspm.fr (YBY); rodrigo.ibata@astro.unistra.fr (RI)

the past histories of these two giant galaxies which share the same environment?

Many previous studies have suggested an ancient major merger event in the past history of M31, based on its classical bulge (Kormendy 2013; Courteau et al. 2011) and its halo rich in streams of various metallicities (van den Bergh 2005). However, they are not predictive about the epoch of the last merger, or more precisely of the final coalescence of the two nuclei. There has been only a single attempt (Hammer et al. 2010) to study a 5–8 Gyr old major merger reproducing the disc (stars and H I), bulge/total (B/T) luminosity ratio, and the 10 kpc ring, though it does not convincingly reproduce the GS that is likely a dynamically young structure given its sharpness and spatial contrast. A recent overview of all the known properties of M31 is perhaps missing, and we refer to the Davidge et al. (2012) study for a more comprehensive overview of the origins of M31 and past modelling efforts (see also Ferguson & Mackey 2016 for the properties of the outskirts). Indeed, to date, most efforts have been devoted to reproducing individual features of M31 assuming a single minor merger event. Such studies have found that:

(i) Assuming that the GS is a trailing tidal tail of a minor merger (few $10^9 M_{\odot}$), N -body simulations are able to reproduce well the structure and kinematics of the GS and the NE and W shells together (Fardal et al. 2006, 2007, 2008, 2013; Mori & Rich 2008; Kirihara, Miki & Mori 2014; Kirihara et al. 2017; Sadoun, Mohayaee & Colin 2014). Such models favour a small spiral galaxy structure for the progenitor (Fardal et al. 2013; Kirihara et al. 2017) and an event duration of ~ 1 Gyr. Finding evidence for the (still undiscovered) existence of the remnant of the GS progenitor dwarf galaxy would constitute a major success for this minor merger model.

(ii) The inner and 10 kpc rings have been proposed as having been caused by an interaction with M32 (Block et al. 2006; Dierickx, Blecha & Loeb 2014); however, such a passage predicts an expanding ring while the 10 kpc ring appears to have been stable for at least 500 Myr (Lewis et al. 2015).

(iii) Bernard et al. (2012) suggested an interaction with M33 on the basis of the coincidence of a shared star formation episode 2–4 Gyr, while McConnachie et al. (2009) came to a similar conclusion based on the presence of tidal disturbances that resemble tidal tails around M33. However, the discovery that the whole of the M31 disc up to 20 kpc experienced a burst of star formation 2–4 Gyr ago renders it less plausible to be provoked by a single passage of a satellite of ~ 10 per cent of the mass of M31 (Williams et al. 2015).

Tanaka et al. (2010) concluded their study of the numerous streams in the M31 halo by noticing that to explain them all, it would require ~ 15 accretions of subhaloes with masses of typical dwarf galaxies. Formerly, the main argument in favour of minor mergers was to ensure the survival of the M31 disc. However, it has been shown that discs may be rebuilt after a sufficiently gas-rich major collision (Hammer et al. 2005, 2009; Hopkins et al. 2009). Furthermore, the spiral arms in M31 seem not to be triggered by the classical density wave theory (Tenjes et al. 2017), which distinguishes it even more from a quiescent galaxy. It is therefore now necessary to reevaluate the plausibility of a single major merger versus numerous very minor events as a means to reproduce most of the exceptional properties of M31 listed above. Which event(s) is (are) responsible of most of these structures, from the bulge to the halo? Could they be reproduced by a gas-rich major merger? This latter possibility has to be consistent with the ‘dynamically young’ structures in the halo such as the GS and shells, and account for

recent, widespread events in the disc that are suggestive of a recent merger episode.

Fully exploring the major merger paradigm is clearly a daunting task. Its study is far more complex and time consuming than that of minor mergers, since:

(i) models of very minor mergers have only to complete one orbit of the progenitor, and do not affect the overall structure of the galaxy;

(ii) models of major or intermediate mergers have to be evolved over several orbits until the complete destruction of the original discs;

(iii) an extraordinary large space of parameters has to be investigated to reproduce the very numerous properties of the resultant galaxy;

(iv) realizations of the GS at the resolution of PAndAS require ~ 200 k particles in a minor merger (see e.g. Kirihara et al. 2017), while we estimate it requires 30–40M particles for a major merger model.

Notwithstanding the above, the aim of this paper is to verify to which extent the Andromeda (and its outskirts) numerous structures can be reproduced by a unique merger. The major constraints are provided by:

(i) the kinematics of the M31 disc from its rotation curve and its age–velocity dispersion relation;

(ii) the structural parameters of the inner M31 galaxy: B/T , bar, structure of the 10 kpc star-forming ring, bulge, bar and disc size, and H I disc;

(iii) the structures surrounding the thick disc, including the warp (e.g. such as the NE Clump, the Northern Spur, the Warp, and the G1 Clump), the NE and W shells and the GS;

(iv) stellar ages and metal abundances of the above structures.

Ideally such a model, if it exists, should be able to reproduce the many detailed properties of the M31 and outskirts structures. However, we are aware that aiming to reproduce in details the spatial and kinematic structures of a nearby galaxy with a major merger study is perhaps inextricably challenging, because:

(i) the accuracy of the calculations is limited for numerical reasons, especially because of the huge contrast in the required number of particles to explain simultaneously the central galactic regions and the faint streams in the outskirts;

(ii) they are further limited by the first pericentres that cannot be estimated at much better accuracy than ~ 4 per cent (see Section 2);

(iii) important features, e.g. shells (see Cooper et al. 2011) are evolving with time, meaning that they can be found like they are in M31 but perhaps not simultaneously with other features;

(iv) a major merger may also produce additional features that could be a consequence of the – completely unknown – internal light distribution in the progenitors.

In this paper, we propose a first attempt in modelling the huge wealth of information recently obtained about our closest giant neighbour, the Andromeda galaxy. Besides trying to reproduce most of the observed properties together, we aim at proposing a physical interpretation of as many morphological details as we can, in the frame of a single event. In Section 2, we establish the framework of the simulations, including the hydrodynamics, the star formation implementation and initial conditions (ICs), as well as the expected limitations when simulating M31. Section 3 presents the results and compares them to the observations, and then Section 4 discusses

Table 1. Initial and adopted conditions for a major merger model for M31.

Ingredient	Tested range	Comments	Adopted range
Total mass	$8.25 \times 10^{11} M_{\odot}$	20 per cent of baryons	–
Mass ratio	2–5	To reform $B/T \sim 0.3$	4.0 (3.5–4.25)
f_{gas} Gal1	0.4–0.6	Expected at $z = 1.5^a$	0.4–0.6
f_{gas} Gal2	0.6–0.8	Expected at $z = 1.5$	0.6–0.8
Orbit	Near polar	To form the ring	–
Gal1 θ'^b	65–100	GS	35–75
Gal2 θ'^b	–50 to –70	GS	–60 to –70
Gal1 ϕ'^c	115 to 175	GS	165
Gal2 ϕ'^c	75 to 110	GS	95–105
$r_{\text{pericentre}}$	28–40 kpc	See the text	32 kpc (31–33 kpc)
Feedback	$1-5 \times \text{median}^{(d)}$	To preserve gas	$1-2.5 \times \text{median}^{(d)}$

^aRodrigues et al. (2012) found $f_{\text{gas}} = 0.5-0.65$ in galaxies with $M_{\text{baryon}} = 0.8-2.2 \times 10^{11} M_{\odot}$ at $z = 1.5$.

^bRotation along the y' -axis.

^cRotation along the z' -axis.

^dIn few simulations, the feedback is assumed to be high before coalescence of the nuclei and later on, assumed to drop to the medium or low feedback values of Cox et al. (2006).

whether M31 could be or not the result of a single major event instead of numerous minor mergers.

2 SIMULATIONS

The faintness of several features (e.g. the streams) requires a very large number of particles to model correctly. This limits us to only being able to reproduce the brightest halo features including the GS and perhaps suggesting further some possible mechanisms to explain, e.g. Streams A–D reported by the PAndAS team. We start simulations using similar orbits than those used in Hammer et al. (2010), most orbital parameters requiring only fine-tuning at the level of 10 per cent variations (see initial and adopted parameter ranges in Table 1). However the 2–4 Gyr star formation history in the whole disc (Williams et al. 2015) and the steepness of the age–dispersion relation found by Dorman et al. (2015) cannot be reproduced by an ancient merger (coalescence of the nuclei 5–6 Gyr ago) as suggested in Hammer et al. (2010). We indeed verify that such event is unambiguously followed by a star formation episode 5–6 Gyr ago and velocity dispersions far below the observed values.

So next we investigate whether a more recent, 2–3 Gyr old coalescence of the nuclei, could explain most of the widespread activity in the disc. Since observations are very constraining, and hence very demanding in terms of controlling the physics, we considerably improve our former, GADGET2 model by performing simulations with the hydrodynamical resolver GIZMO. A significant part of the study has been devoted to exploring a limited number of parameters, first the mass ratio (mr) and the pericentre. Most features depend considerably on these parameters, since the former defines the number of passages (the smaller secondary mass, the larger the number of passages) and the latter provides the time elapsed between two passages. Then, we have had to search for optimizing resonances that are revealed by the presence of the bar, and a considerable amount of work was devoted to finding parameters that amplify them mostly by varying the progenitor sizes. Besides this we also reproduce the bulge, disc scale length and rotation curve (see Appendix A) as in Hammer et al. (2010).

2.1 Hydrodynamical resolver and star formation and feedback implementation

2.1.1 GIZMO: an optimized hydrodynamical solver

GIZMO is a recently published code for N -body hydrodynamical computation that has been developed by Hopkins (2013, 2014), to whom we refer for a full description of the algorithm. This code is developed and optimized from GADGET3, which is an worldwide-used N -body/smoothed particle hydrodynamics code by Springel (2005). In addition to all the advantages of GADGET3, Hopkins (2013) introduces accurate (to second-order) hydrodynamical solvers that rely on Lagrangian numerical methods, i.e. the meshless finite-mass method (MFM) and the meshless finite-volume method. This makes GIZMO able to perform massive parallel computations, and provides very accurate solutions of gas hydrodynamics, as well as the best conservation of mass, momentum, and energy. These advances motivated us to change our simulation software from 2GADGET2 to GIZMO. Simulations presented in this paper use the MFM hydrodynamical solver and a fixed softening 0.16 kpc for a 2-million particle simulation, and 0.08 kpc for 20-million particle simulation, respectively.

2.1.2 Star formation and feedback implementation

Following Cox et al. (2006), we have successfully implemented a simplified star formation and feedback model into GIZMO, applying the same methodology as done by Wang et al. (2012). This model can describe well the star formation history over large scales in galaxies and has been used intensively in the last decade (see e.g. Hammer et al. 2010; Wang et al. 2012, 2015).

In high gas density regions, gas can form stars. The star formation rate (SFR) of each gas particle is calculated according to the local gas density and the local dynamic time based on the observed Kennicutt–Schmidt law (Kennicutt 1998). A stochastic method (Springel 2000; Cox et al. 2006) is used to convert gas particles into star particles.

During star formation, feedback processes are important to regulate star formation. In this method, the energy released by supernovae during star formation is first stored in a new reservoir of internal energy, which provides additional pressure to support the gas and prevent it from further collapsing to form stars. This feedback energy can be thermalized by a free parameter, which controls the time-scale of this thermalization and then the feedback strength.

The radiative cooling processes are necessary for gas to cool down and collapse to form stars. In this work, the cooling rate is calculated with the method implemented in GIZMO, which treats the gas as a primordial plasma, and the ionization states of H and He for a collisional ionization equilibrium are assumed (Katz, Weinberg & Hernquist 1996).

2.2 Initial conditions, parameter choices, and limitations

We follow the same method used in Hammer et al. (2010) (see also Wang et al. 2012) to create ICs and to test the stability of the progenitors in isolation. Each progenitor is assumed to possess only two components: a dark matter halo and a thin disc that includes stars and gas. We define the gas fraction of the progenitor to be the fraction of gas mass to the total baryonic mass and the mass ratio to be the ratio of the total mass of each galaxy. The density distributions of each component and their N -body realizations follow exactly the

Table 2. Initial parameters of the five fiducial models plus that with 20 million particles (model 336). The second column indicates the number of particles ($M = 10^6$). The third and fourth columns indicate the pericentre and the mass ratio. The fifth column gives the feedback prescription (i.e. versus the median value of Cox et al. 2006), and in models 314 and 336 we have let vary the feedback value before and after coalescence of the nuclei (first and second values, respectively). The next four columns provide the angle (in degrees) of the progenitors against the orbital plane (see the text) and 1st and 2nd indicate primary and secondary progenitors, respectively. The last columns indicate the initial disc scale length (in kpc) of stars and gas in the progenitors, and their initial gas fraction, respectively.

Model	N particles	r_p (kpc)	mr	Feedback \times Cox median	θ' 1st	ϕ' 1st	θ' 2nd	ϕ' 2nd	h_s 1st	h_g 1st	f_{gas} 1st	h_s 2nd	h_g 2nd	f_{gas} 2nd
255	2M	29.4	4.0	1	70	165	−60	85	2.8	8.4	0.4	2.8	8.4	0.6
276	2M	31.9	4.0	1	70	155	−60	85	2.8	16.8	0.6	2.8	16.8	0.6
288	2M	31.5	4.0	1.25	70	165	−60	85	2.8	12.6	0.6	2.8	12.6	0.6
290	2M	32.0	4.0	1	70	165	−60	105	2.8	8.4	0.4	2.8	8.4	0.6
314	2M	34.8	4.0	2.5–1	75	165	−60	105	2.8	8.4	0.4	2.8	8.4	0.6
336	20M	32.8	4.0	2.5–1	70	165	−60	100	2.8	8.4	0.5	2.8	8.4	0.8

method described in Barnes (2002).¹ Both the stellar and the gas discs are assumed to be thin with a scale height equal to 1/10 of the scale length. The scale length of the gas disc is assumed to be three times larger than that of the stellar disc (Cox et al. 2006). The scale of the dark-matter halo core (a_{halo}) is chosen to ensure that the initial disc is consistent with the Tully–Fisher relation. For the simulations analysed in this paper a_{halo} are 11 and 4 kpc, for the large and small progenitors, respectively.

Most of the parameters for ICs used in this work are listed in Table 1. We use the same dark matter fraction (20 per cent) and density profile as in Hammer et al. (2010). Half of our simulations were run with 2 million particles. To test for convergence, we also ran a simulation with 20 million particles. We found the resulting end-point structures were qualitatively unchanged, which indicates that convergence has been achieved (see Appendix B).

Each simulation was launched on a parabolic orbit ~ 1 Gyr before the first passage to allow sufficient relaxation of the progenitors. The parameters in Table 2 are defined in the simulation frame where the orbital plane of the merger is always put in the x' – y' plane with the orbital spin aligned with the z' -axis; the initial pericentre is always on the positive side on the x' -axis. To provide ICs, the progenitor discs are first put in the x' – y' plane with their rotational spin along the z' -axis, then rotated about the axes by angles θ' and ϕ' (for a spherical coordinate system) that are given in Table 2.

2.2.1 Constraints on the mass ratio and pericentre from the HI disc

The Andromeda galaxy is characterized by a prominent classical bulge and a significant bar (Athanasoula & Beaton 2006; Beaton et al. 2007). Based on its polar orbit, the previous model of a 3:1 merger by Hammer et al. (2010) also aimed to reproduce its star-forming, 10 kpc ring, as well as its HI disc (see their fig. 6). Here, we have to also account for the $\gtrsim 2$ Gyr old wide perturbations of the M31 disc. We identify the HI disc to be the feature that is least stable against either a recent merger or a significant amount of feedback. Fig. 1 shows how varying the merger mass ratio and pericentre affects the HI disc size. It suggests a merger mass ratio (mr) from 3 to 4 and a pericentre larger than ~ 32 kpc. Indeed a large pericentre brings more angular momentum to the remnant disc. From our examination of our whole set of models (see also Fig. 1), we found that $\text{mr} \geq 4.5$ as well as $r_p < 30$ kpc are likely excluded values to reproduce the M31 HI disc size.

¹ See also <http://www.ifa.hawaii.edu/~barnes/software.html>.

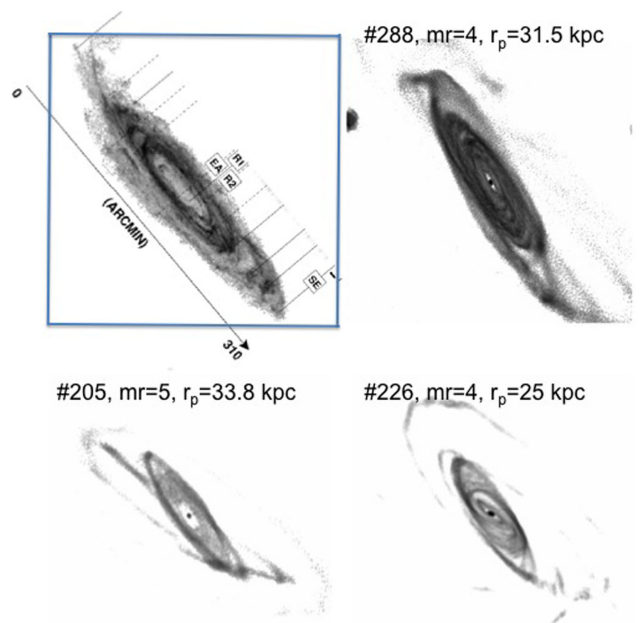


Figure 1. Comparison of the H I discs provided by models with the same orbital parameters but for which the pericentre (right-hand panels) varies from 25 to 32 kpc or the mass ratio from $\text{mr} = 4$ to 5. From top to bottom and left to right: observations from Chemin, Carignan & Foster (2009) and three models for which number, mass ratio, and pericentre are indicated on the top of each panel; the size of the blue box is 60 kpc, and all plots are on the same scale.

2.2.2 Accuracy of the pericentre and its adopted choice

We found that important properties of a merger remnant a few Gyr after the coalescence of the nuclei are well correlated with the pericentre distance at the first passage. Even small changes of the first pericentre, e.g. of 1.2 kpc (i.e. 4 per cent of 32 kpc), will change the elapsed time between first and second passages by ~ 1 Gyr, affecting the final coalescence time, and then the timing of the epochs of star formation in the rebuilt disc. However, the actual pericentre value is not easy to control just from the initial pericentre because it also depends on the different properties of the initial progenitors, such as the gas fraction, size of both stellar and gas discs, and star formation history. Different ICs will lead to different star formation histories and different mass distributions, resulting in changes in, e.g. the mass ratio at later passages, and then to strong variations in dynamical friction. This is demonstrated in

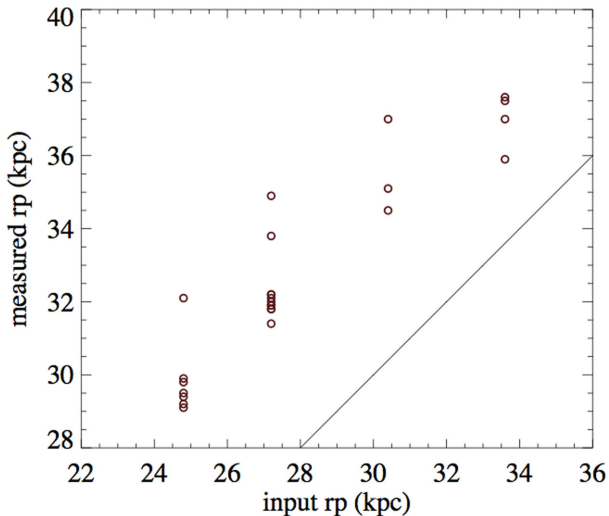


Figure 2. Input versus actual pericentre in the simulations. The dispersion in measured pericentre is due to different ICs that affect the mix between stars and gas.

Fig. 2 in which we show various simulations for which several values of feedback, initial gas fraction, and initial disc sizes have been implemented. We find a strong trend that almost follows the behaviour of a point mass (shown as a straight line in Fig. 2), and differences with point mass can be easily corrected for. Besides this, we find a scatter of ~ 10 percent solely due to the physical differences in the progenitors (see points in Fig. 2). After realizing this, we analysed each simulation and relaunched it with slight changes of the initial position and velocity at the beginning of the simulation, in order to have to the simulation run within 4 percent of the desired pericentre. The pericentre value is found to be a key parameter in our simulations for reproducing as accurately as possible the properties of M31.

We found that adopting a value of $r_p \sim 32$ kpc for the first pericentre distance can give rise to final simulation structures in good qualitative agreement with the observations, as we show below. However the large M31 disc size (especially the H_I disc) may favour even higher values. Increasing the pericentre significantly would have been problematic, as it would increase the absolute uncertainty (in kpc) on the pericentre, and then on the elapsed time between, e.g. first and second passages. In such conditions, it would then have been difficult to compare different simulations for exploring the parameter space, and even to reproduce the same result with two simulations at different resolutions (e.g. we need to increase resolution to reproduce the faint features). Future work will explore further the consequences other choices of the r_p parameter.

2.2.3 Optimizing the parameters to reproduce main features: the bar and the 10 kpc ring, discs, bulge, and rotation curve

For each set of orbital parameters (see Table 1), we changed the scale lengths of the initial stellar and gas discs in order to optimize the bar and the 10 kpc ring. The presence of the bar is also supported by the kinematics of the central regions (Opitsch et al. 2017). Fig. 3 shows that we succeed in reproducing a strong bar accompanied by a young star distribution similar to the observed one. It also demonstrates that the orientation of the bar is distinct from that of the major axis of the disc, as is observed (see e.g. Athanassoula & Beaton 2006 and compare the inserts in the top panels of Fig. 3).

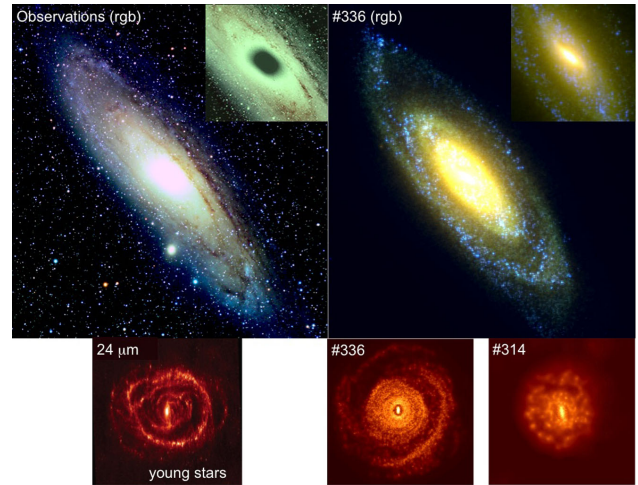


Figure 3. Top row: comparison of red–green–blue (rgb) composite-colour images of the M31 disc (left) with our highest resolution simulation (right, model 336, see Table 2). The spectral energy distribution of each particle was calculated using the Bruzual & Charlot (2003) population synthesis model, assuming a single stellar population and seven metallicity tracks. It does not account for dust, explaining part of the differences between images (see e.g. dust lanes in the observed disc). On the top right of the observed image, an insert (a 26.6 kpc wide box) shows *BVRz* observations and the superposed central bar (*K* band, in black) by Beaton et al. (2007), which can be compared to the simulation for which we have improved the contrast in the central region (see insert on the top left of the simulated disc). Bottom row: comparison of the star-forming regions in the M31 disc within a 40 kpc wide box. From left to right, the 24 μm (Gordon et al. 2006) reconstructed face-on view of the bar and of the 10 kpc ring, and simulations of < 0.2 Gyr old stars for models 314 and 336 (see Table 2).

However our bar realizations often lead to thinner bars than that observed. The 10 kpc ring is often found off-centred (see bottom left panel of Fig. 3) as in the observations (Lewis et al. 2015), and our modelling may be consistent with the (observed) presence of an outer ring (or a part of it, see the bottom middle panel of Fig. 3).

Besides the need to reproduce the actual properties of M31, an important motivation for favouring models that produce bars is that M31 possesses a relatively concentrated density profile, with the bulge representing 30 percent or more of the total stellar mass (see e.g. Blańa Díaz et al. 2017). As shown in Wang et al. (2012), the properties/spatial distributions of the streams strongly depend on the mass concentration of the remnant galaxy. We have also used intensively the Amorisco (2015) study of the dynamics of tidal streams to optimize the reproduction of the properties of the GS (see Section 3.2). We have selected five fiducial models, whose parameters (see Table 2) show a particularly narrow range. They were selected after a considerable search effort, including ~ 300 simulations, half with 500k particles and half with 2M particles.

As in Hammer et al. (2010), we reproduce the H_I disc (compare Fig. 1 with their fig. 6), the stellar density profile of the disc, the bulge to disc ratio as well as the rotation curve (see e.g. their fig. 5 and appendix A of this paper). This is because for a gas-rich merger, the disc settles within ~ 1 Gyr after the coalescence of the nuclei, then similar orbits with similar mass ratios provide similar disc scales and rotation. Our simulations produce a relatively large scale-length final disc (3.3, 5.8, 3.5, 3.9, and 4.8 kpc for models 255, 276, 288, 290, and 314, respectively) although often smaller than the M31 disc value (5.8 kpc, see discussion in Hammer et al. 2007 and also Blańa Díaz et al. 2017). The disc scale length of M31 (as well as the H_I disc size) is particularly large when compared

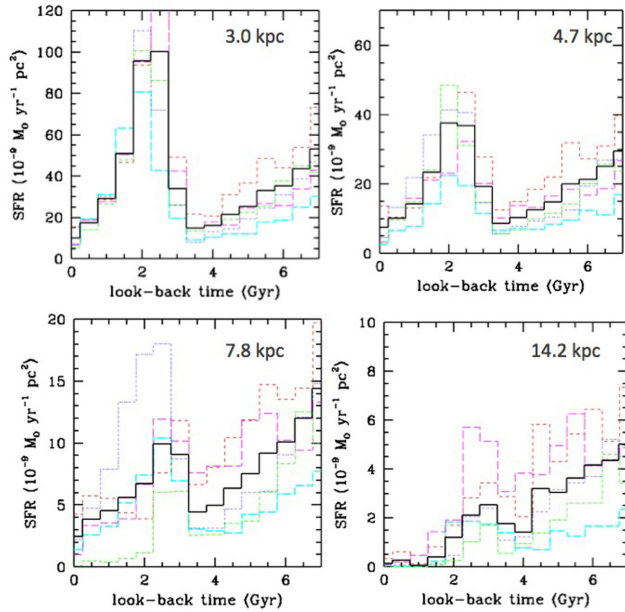


Figure 4. Star formation history from the five fiducial models of Table 2, as measured at 3, 4.7, 7.8, and 14.2 kpc along the disc, i.e. following fields studied by Williams et al. (2015). Green (dot-short dash), magenta (long dash), red (short dash), blue (dot), and cyan (dot-long dash) lines represent models 255, 276, 288, 290, and 314, respectively. The coalescence epoch is 3.0, 2.1, 2.1, 2.1, and 1.8 Gyr ago for the later models, respectively. The black bold (solid) line represents the average of the later models, displaying the presence of a ~ 2 –4 Gyr old star formation episode over the whole disc.

to those of similar spiral galaxies (see e.g. Hammer et al. 2007), and if it results from a merger, a large pericentre would naturally form a large remnant disc. In the previous section, we discussed the difficulty in studying large pericentres: this is likely the main limitation of our study.

3 RESULTS

3.1 The unusual properties of the M31 disc: stellar ages and kinematics

3.1.1 A 2–4 Gyr strong star-forming event

In M31, the outer disc stars are moderately old or young and Bernard et al. (2015b) found that most of them formed ≤ 8 Gyr ago, while a quarter of them formed in the last 5 Gyr. Indeed, such an apparent rejuvenation of the disc is also observed in many spirals with similar mass and type as M31 (see e.g. González Delgado et al. 2017). Williams et al. (2015) have undertaken a systematic study of the age of the stellar disc by sampling PHAT fields for which dust extinction is very small. They however limit their study to the sole period of the last 5 Gyr, which contains the 2–4 Gyr old strong burst of star formation (see also Bernard et al. 2015a).

Fig. 4 presents the result of the 4:1 merger models for which the coalescence of the nuclei occurred from 1.8 to 3 Gyr ago. All models show a significant burst of recent to moderately recent star formation, qualitatively similar to the observations. To compare in a more quantitative way, Fig. 5 reproduces fig. 4 of Williams et al. (2015) to which we have added the average value of the five models of Fig. 4. Our simulations match quite closely the observations in the three first age panels, when accounting for the observational uncertainties (shown in grey), and for the model differences (see the

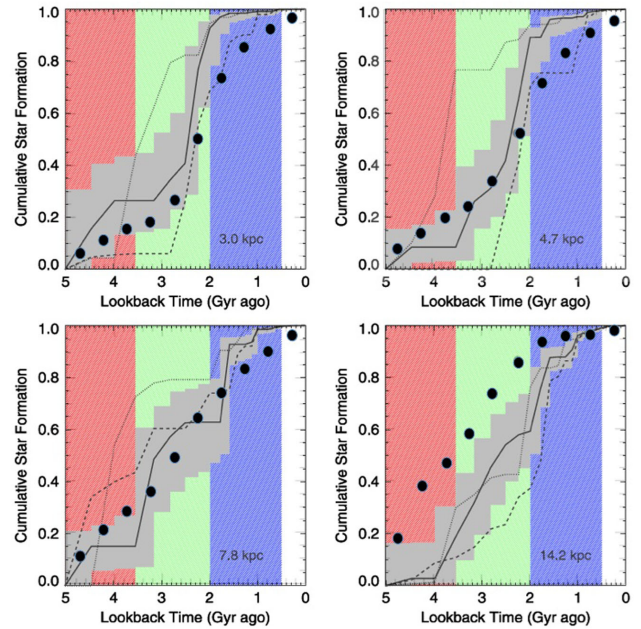


Figure 5. Same as Fig. 4 (solid dots are averaged values of the five models), but for cumulative star formation to enable a direct comparison to fig. 4 of Williams et al. (2015).

dashed and dotted lines, and Williams et al. 2015 for more details). Besides this, the burst of star formation seems too diluted in the outer disc field to reproduce quantitatively the observations, except for model 276 (see magenta, long-dashed line in the fourth panel of Fig. 4), which shows the largest gas-disc scale length. We interpret this as possibly due to our limitations in performing simulations with very large pericentres, i.e. those producing the largest disc scale lengths.

3.1.2 All stars older than 2 Gyr are in a thick disc

The PHAT survey has enabled to establish an extinction map of the red giant branch (RGB) stars in the M31 disc (Dalcanton et al. 2015), revealing a scale height (0.87 kpc) comparable to that of the MW thick disc. Only 10–15 percent of the MW disc stars (Jurić et al. 2008) are in the thick disc, which is assumed to be formed 10–11 Gyr ago, and these stars are much more metal poor than those in the thin disc. This contrasts with the discovery of Dorman et al. (2015) that the RGB stars in the M31 disc have large velocity dispersions (averaging to ~ 90 km s $^{-1}$). It is a definitive proof of an overwhelming dominant thick ($V/\sigma \leq 3$) disc in M31. The M31 disc is much hotter than that of the MW. Dorman et al. (2015) also found a very steep age–velocity dispersion correlation. Fig. 6 shows that such behaviour is expected for a merger remnant after a recent (2–3 Gyr ago) coalescence of the nuclei, and the open points (and bold dashed lines) show that elapsed times since coalescence larger than 3–4 Gyr would be too large to keep the disc as kinematically hot as is observed. In this figure, we have assumed an aperture (1 kpc) that corresponds to the smoothing radius used by Dorman et al. (2015) to create their 2D dispersion maps. Notice that the high σ values cannot be caused by the initial velocity dispersion in the progenitors, which was set to 10 km s $^{-1}$. It is however closely associated with the highly retrograde motion of the primary in the orbital plane (ϕ' slightly below 180°, see Table 2).

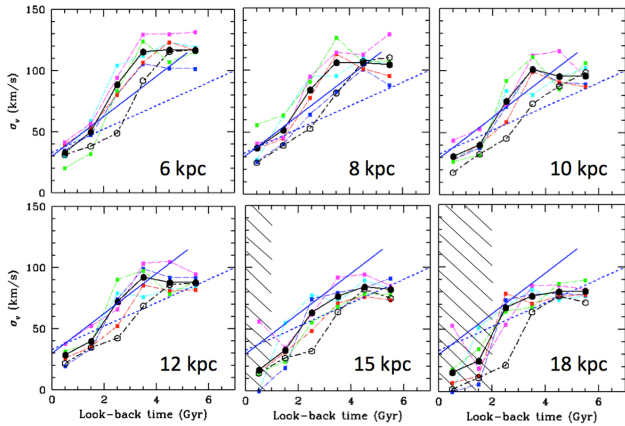


Figure 6. From top left to bottom right: kinematic dispersion of the disc stars at 6, 8, 10, 12, 15, and 18 kpc. The (blue) solid and dashed straight lines are the best fit to the data points at $R \geq 10$ kpc with a constant SFR and decreasing SFR assumed by Dorman et al. (2015), respectively. Cyan, blue, red, magenta, and green represent models as in Fig. 4. The black bold (solid) line and solid dots represent the average of the models, while the bold dashed line and open dots represent the average value of the same models, but taken 1 Gyr later. The hatched areas in the two last panels indicate stellar ages for which the number of (young) stars are not sufficient to derive a reliable velocity dispersion. Note that the panels at 6 and at 8 kpc illustrate the Dorman et al. result that the kinematic dispersion of stars within 10 kpc can reach 120 km s^{-1} .

A slight difference between our model estimates and the observations in Fig. 6 comes from the fact that we lack a sufficient number of particles outside the projected major axis of the M31 disc. To improve the statistics, we run a 20 M particles simulation, model 336, for which the parameters are exactly those of model 314, except for a 5° offset on the angular position of the secondary and an increase of the progenitor gas fractions (see Table 2). Fig. 7 shows the 2D dispersion maps of the modelled M31 assuming four age ranges, which correspond to main sequence, young and old asymptotic giant branch and RGB stars, from the best expectations based on figs 9 and 12 of Dorman et al. (2015). It reveals a remarkably similar σ_v map to the observed one (see fig. 7 of Dorman et al. 2015), both quantitatively and qualitatively. The age–velocity dispersion relation and its steepness is widespread over all the disc, and the velocity dispersion increases significantly at low radii. Figs 6 and 7 show that our modelling also reproduces the dispersion versus radius relationship shown in fig. 16 of Dorman et al. (2015), i.e. a strong increase of the dispersion below $R = 10$ kpc. As with the observations, our modelled 2D maps show local variations in the velocity dispersion. It is beyond the scope of the present contribution to acquire sufficiently detailed maps for, e.g. retrieving the influence of the bar as proposed and tested by Dorman et al. (2015).

3.2 The unusual properties of the M31 outskirts: giant stream, shells, clumps, and warps

The main success of minor merger models (Fardal et al. 2007, 2008, 2013) is to reproduce simultaneously the NE and W shells together with the GS. The association with a less than 1 Gyr old event is also consistent with the sharpness of the observed shells and GS, which are dynamically ‘young’ features. Shells are also formed in major mergers as shown by observations and by, e.g. the modelling of NGC 7600 by Cooper et al. (2011). However, the major merger model of Hammer et al. (2010) did not reproduce well these features,

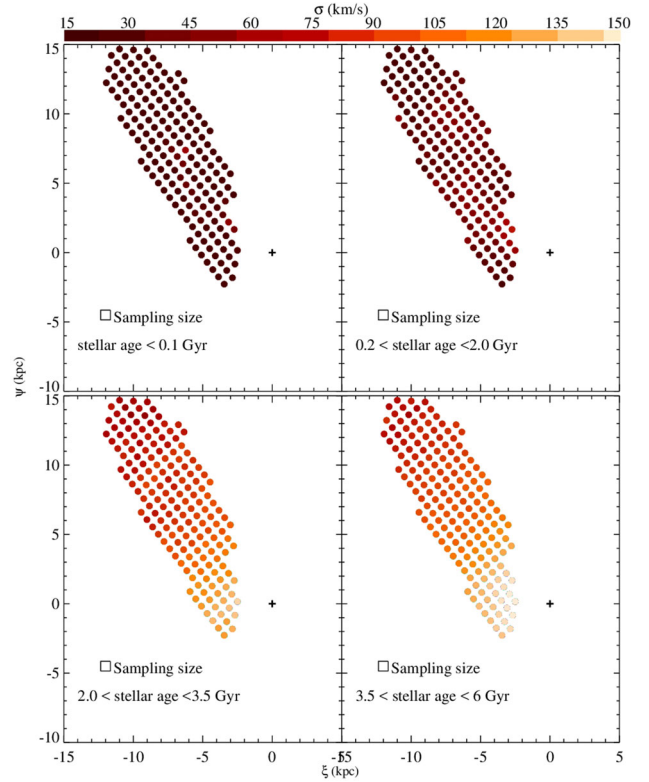


Figure 7. Maps of the velocity dispersions of stars of different ages for a direct comparison to fig. 7 of Dorman et al. (2015). Since 20 M particles are necessary to achieve sufficient particle number per pixel, we show only the results for model 336. The cross represents the M31 centre.

which might be due to the small particle number (200k for stars) used in that study. However, a coalescence of the nuclei occurring 5.5 Gyr ago let few chances to recover the sharpness and contrast of the observed features. We have verified that such an ‘ancient’ merger in M31 would underestimate the amplitude of the velocity dispersion in the M31 disc by a significant factor (~ 2).

Since a 2–3 Gyr old major merger is an excellent contender for explaining the kinematics and star formation of the M31 disc, one may wonder if this event could be sufficiently recent to be also responsible for the GS, the NE, and W shells, including their sharpnesses. Fig. 8 shows that the fiducial models lead to sharp GS-like structures and shells, though the NE shell is often smaller than in the observations except perhaps for model 314 for which it seems even too prominent. The models show a strong warp on both sides of the disc main axis, which share many similarities with the observed Warp and Northern Spur regions. Table 3 gives surface brightness ratios between several features and the GS that range around the observed values within a factor ~ 2 . Moreover, Fig. 8 shows that the NE and G Clumps are more prominent than those observed. We note that changing the progenitor sizes provides different Clump morphologies and a better reproduction of them would require a better knowledge of the initial gas and stellar spatial distribution in both progenitors. The NE and G Clumps have been posited to be tidal debris or undergoing tidal disruption (Zucker et al. 2004; Ibata et al. 2005) from their disorderly morphologies, shared by their gaseous counterparts (Lewis et al. 2013). In our modelling, both clumps correspond to tidal dwarfs associated with the main progenitor that are in a process of disruption, and their kinematics and stellar content are likely similar to those of the M31

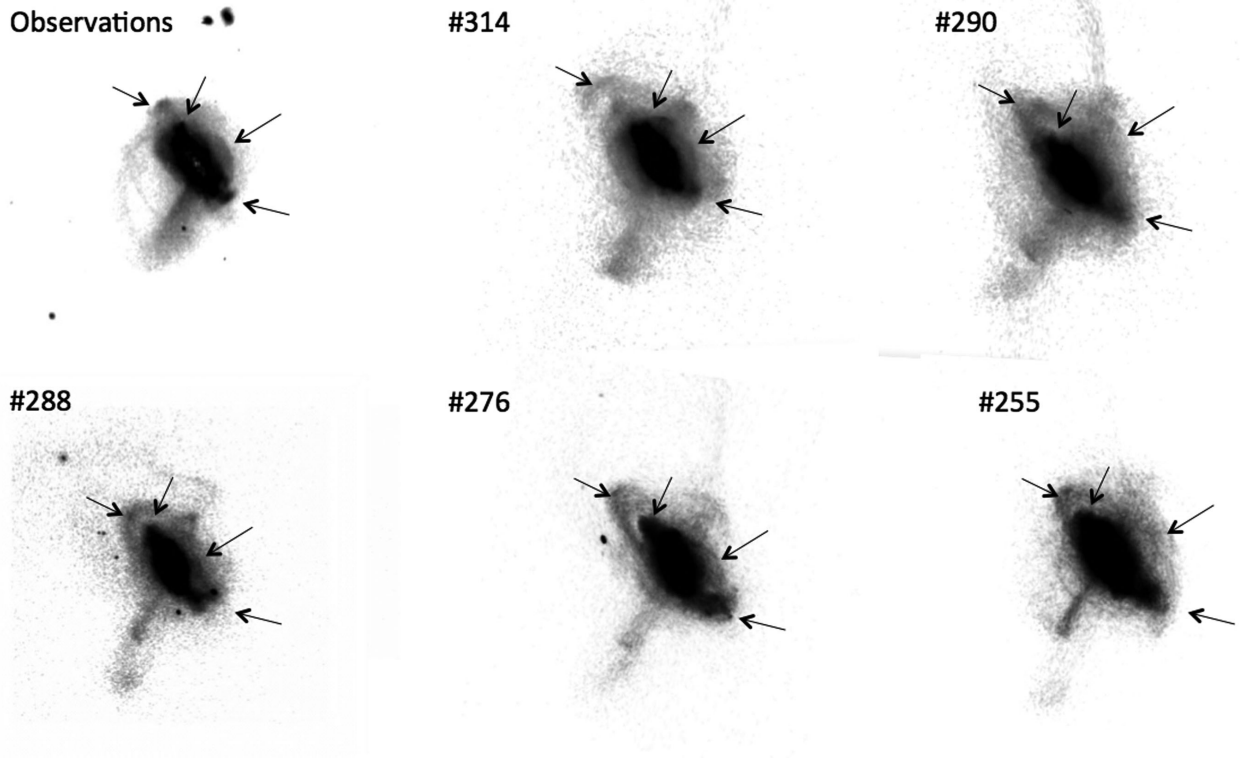


Figure 8. A comparison of the PAndAS imaging (first panel on the top left) with that of five selected models. The height of each panel is 264 kpc. The arrows indicate the most obvious features besides the GS, from left to right, top to bottom: NE Clump, N Spur, W shell, and G1 clump. Notice that the resolution is ~ 16 times better in the observations than in the models.

Table 3. Surface brightness ratios between the following fields: F1 (21 kpc, halo field of Brown et al. 2006), F2 (GS field of Brown et al. 2007), F3 (warp field), F4 (north-east shelf), and the F2 field has been taken as the reference.

	F1	F2	F3	F4
Observations	0.08	1.00	3.23	0.91
Model				
255	0.11	1.00	5.06	0.46
276	0.10	1.00	7.66	0.67
288	0.17	1.00	9.79	3.10
290	0.15	1.00	2.76	0.95
314	0.02	1.00	6.98	1.53

disc, which would reconcile former interpretations of the G1 and NE Clumps.

In Fig. 8, the GS is mostly the superposition of the first and second loops associated with the tidal tail linked to the second passage of the secondary (see Section 4.2). The loop plane is seen edge-on since the first and second tidal tails share the same plane as the disc of satellites, explaining then why it points to the MW (see e.g. Hammer et al. 2013). The second loop is the most recent and the brightest one and could explain the sharpness of the GS structure. Since the brightest (second) loop is further away than M31, this is in qualitative agreement with the distance behaviour of the GS described by Conn et al. (2016, see their fig. 4), and a more detailed description can be found in Section 4.2. Finally, the GS shape strongly depends on the initial orientation of the secondary, though we consider that the contrast and sharpness of the GS have been largely recovered by our modelling.

Fig. 9 shows the stellar age distribution in both the GS and the W shell. It illustrates well how the modelling is able to reproduce

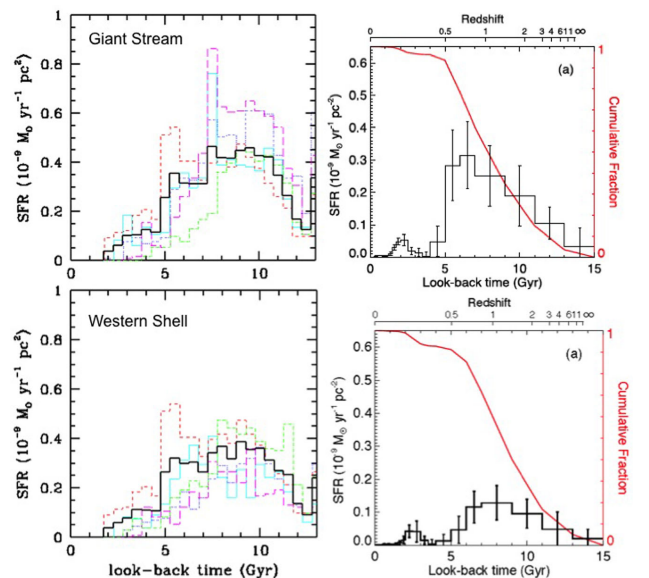


Figure 9. Top row: star formation history derived for the GS, on the left based on the five models shown in Fig. 8 (thick black line indicates the average), on the right as it has been measured by Bernard et al. (2015a). Cyan, blue, red, magenta, and green represent models as in Fig. 4. Bottom row: same as above for the star formation history of the W shell, on the right as it has been measured by Bernard et al. (2015b).

the GS age distribution as well as the SFR density amplitude. As it is observed, the age distribution of the W shell is similar to that of the GS. In fact both structures are linked to the material of the secondary encounter that dominates the GS and the shell. In

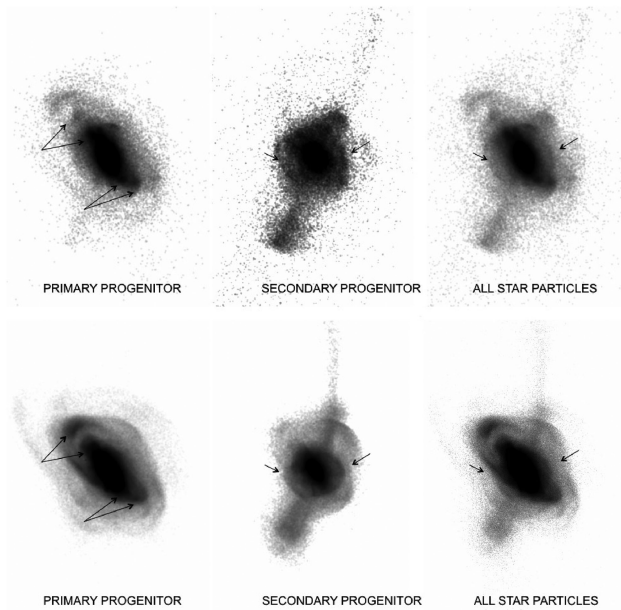


Figure 10. Decomposition of stellar particles from their origin into the two progenitors in model 314 (top row) and in model 336 with 20 M particles (bottom row). The height of each panel is 264 kpc. Left: particles from the main progenitor. The arrows indicate the warps and the clumps. Middle: stellar particles from the secondary. The arrows indicate the shells. Right: all stellar particles. The arrows indicate the shells.

Fig. 10, we decompose the particles of the primary (left-hand panel) and the secondary (middle panel) progenitors. This suggests that the substructures in the outer M31 disc are of tidal origin, with warps and clumps being linked to the primary, while the GS and the shells are mostly the imprints of particles of the secondary progenitor.

Bernard et al. (2015b) have separated outer disc structures into disc-like (warps and clumps) and stream-like (GS and shells) categories, respectively, depending on their star formation history and metallicity properties. Fig. 10 shows that a major merger model can produce precisely these two kinds of structures. Notice, however, that a few particles of the primary have been captured by the secondary during the interaction and can be found in the GS as well as in the shells.

4 DISCUSSION AND CONCLUSION

4.1 Limitations, weaknesses, and successes of a major merger modelling of M31

In this paper, we have investigated the properties of a potential major merger that may have occurred in M31. The main challenges arise from the gigantic number of observational constraints provided by the numerous surveys that have scrutinized our closest neighbour, and the huge contrast between the brightest and faintest features (e.g. a factor of 1000 between the central bulge and the GS). Furthermore, the uncertainties linked to the pericentre determination (at best within 4 per cent) could change by 1 Gyr the elapsed time between the first and second passages, limiting significantly our capability to optimize the modelling. Large pericentre values are however mandatory to rebuild discs as large as that of M31. To mitigate the uncertainties, we have chosen a pericentre radius of $r_p \sim 32$ kpc, which has led us to $mr \sim 4$. We conjecture that a smaller mass ratio (e.g. 3:1) would require a higher pericentre to provide

similar results. On the other hand, larger mass ratios (e.g. $mr \geq 4.5$) are excluded since they cannot produce sufficiently large discs in the remnant galaxy. Our current modelling leads us to assume $r_p \geq 32$ kpc and $mr \leq 4.25$ as robust limits.

Despite these limitations, the modelling of M31 as a recent (major) merger is able to reproduce most internal properties of M31 (bulge, bar, 10 kpc ring, stellar, and H I discs) including its rotation curve and disc dispersion. This has been done after an intensive search to optimize the initial orbital parameters, and those linked to the progenitor structures (see Table 2). Disc-like features (warps and clumps) associated with the outer disc are also produced by the modelling as well as stream-like features such as the GS and shells (see Figs 8 and 10). Figs 4 and 9 confirm the ability of the model to reproduce the star formation histories of these structures.

This apparently contradicts the conjecture made by Hammer et al. (2010) that the ≥ 5 Gyr old star population of the GS can be used as a clock, implying a similarly old coalescence epoch. While this is probably true for their adopted 3:1 mass ratio and 25 kpc pericentre, by increasing mass ratio (to 4:1) and pericentre (to 32 kpc), we have somewhat unsettled the clock. The GS is made of particles returning to the remnant from the tidal tail formed at the second passage, mostly including material from the secondary (see middle panels in Fig. 10). First, as soon as the material is deposited on the tidal tail the star formation ceases, and the time elapsed between the second passage and the coalescence of the nuclei (as well as the number of passages) is increasing with the mass ratio, due to a decreasing dynamical friction. Second, 1 or 2 Gyr before the second passage, the star formation in the secondary is likely smaller for large pericentres, because the system has a much larger time to relax. In this paper, we have verified that a 4:1 merger, 32 kpc pericentre merger reproduces together the 2–4 Gyr old burst of star formation, the GS star formation history, and the age–dispersion relation in the disc.

However, Fig. 8 does not provide an accurate reproduction of the observed outer disc features (see also the right-hand panel of Fig. 10). For example, the simulated NE Clump is often more extended than what is observed. There is also a more prominent counterpart on the opposite side of the GS than in the observations. It appears that the former discrepancy depends on the precise morphology of the progenitors, while the later is reduced by increasing the mass ratio and the initial gas fraction. There is another discrepancy that considerably alters our modelling ability to reproduce features, altogether. This is caused by the occurrence of shells that seem to be highly variable with time, restricting the possibility to accurately model the outskirts of the M31 disc. The W and NE shells of M31 seem very similar to shells seen in simulations of elliptical galaxies (see e.g. Cooper et al. 2011), i.e. shells are arranged along the orbital path of the secondary, to which they are linked, as demonstrated by Fig. 10 (see middle panels). Our modelling shows that shells are as variable as those seen in the Cooper et al. (2011) simulations, and Fig. 11 demonstrates that the variation time-scale may reach values below 100 Myr.

4.2 A single major merger or a multiple minor merger paradigm?

The numerous studies of the GS (Fardal et al. 2006, 2007, 2008, 2013; Mori & Rich 2008; Kirihara, Miki & Mori 2014; Kirihara et al. 2017; Sadoun, Mohayaee & Colin 2014) in a minor merger context have been very successful in reproducing the detailed features of the stream such as, e.g. the asymmetric profile across the GS (Kirihara

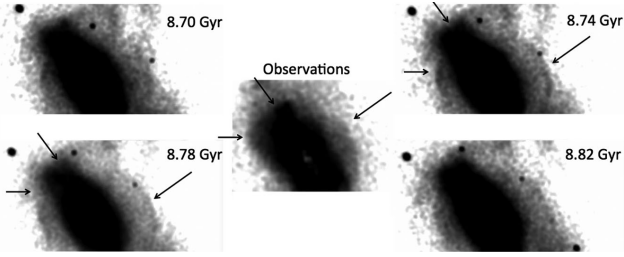


Figure 11. Central image: zoom-on the thick disc outskirts from PAndAS (Ibata et al. 2014). Outer panels from top left to bottom right: evolution of the simulations (model 288) over 120 Myr, showing that many substructures near the thick disc are evolving rapidly, including shells (see arrows).

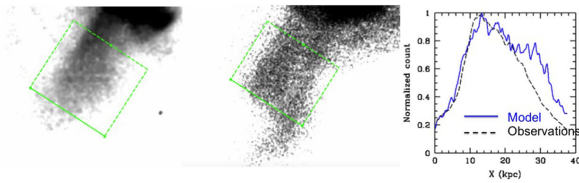


Figure 12. Left: zoom-on the GS from PAndAS (see Ibata et al. 2014), the green box indicating how the profile across the GS has been derived. Middle: same for model 336. Right: profile across the GS.

et al. 2017). Fig. 12 shows that such a property can be retrieved also by our modelling. Interestingly, we have verified that orbits taken for the minor merger (e.g. from Fardal et al. 2006) share similarities with our orbital parameters, especially for the second passage that determines the spatial structure of the GS. This suggests that a major merger (as well as a minor) is able to reproduce the GS together with NE and W shells, those structures being mostly attached to the secondary or satellite (see middle panels of Fig. 10), while other structures such as warps and clumps are mostly attached to the primary progenitor (see left-hand panels of Fig. 10).

In minor merger models, satellite (or secondary) masses range from $3\text{--}4 \times 10^9$ (Fardal et al. 2013; Kirihaara, Miki & Mori 2014) to $4.2 \times 10^{10} M_{\odot}$ (Sadoun, Mohayaee & Colin 2014). It suggests that the success in modelling the GS depends more on orbital parameters than on mass ratio. Sadoun, Mohayaee & Colin (2014) reduced the GS-compatible mass ratio by a factor ~ 15 (from 300:1 to 21:1), and this work allows it to be extended by a supplementary factor 5, to 4:1. We have verified that 2:1 to 5:1 models (see Table 1) may also reproduce the same features.

The main theoretical support for a minor merger explanation for the GS and other streams comes from the simulations of Bullock & Johnston (2005), which predicted a considerable number of minor accretion events. This influential study predicted that mass deposited by hundreds of satellites could explain the stellar content of galactic haloes and their associated streams. However, the prediction of hundreds of subhaloes, e.g. for the MW or M31, is not observed, dubbed as the ‘missing satellites’ problem (see e.g. Bullock & Boylan-Kolchin 2017, or for an alternate viewpoint, Del Popolo & Le Delliou 2017). Perhaps the evidence for multiple minor mergers having occurred in M31 has to be reconsidered. In fact, semi-empirical Λ CDM models have been proposed by Hopkins et al. (2010) and systematically compared to observations. It emerges that at $z = 0$ and for an $M_{\text{stellar}} \geq 10^{11} M_{\odot}$ galaxy (such as M31), the occurrences in the past 10 Gyr of a 4:1, 21:1, and 300:1 merger are 70, 108, and 160 per cent, respectively. In other words, a major merger scenario for M31 would have a similar occurrence as

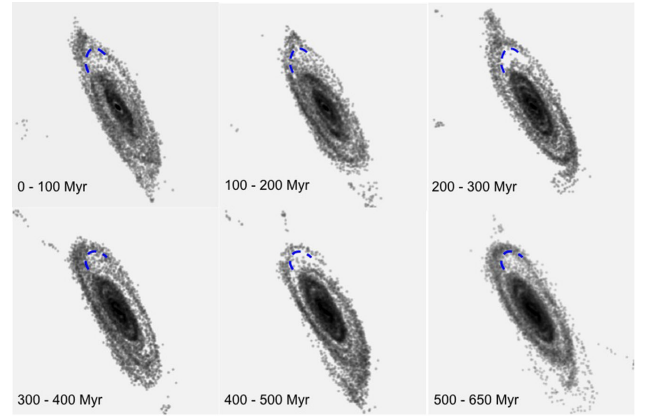


Figure 13. Maps of the SFH density ($M_{\odot} \text{yr}^{-1} \text{kpc}^{-2}$) of model 336 using 650–0 Myr old stars. The time range covered is shown in the lower left of each plot, and the figure can be directly compared to the observations provided in fig. 5 of Lewis et al. (2015), including the blue dashed curve that aids the eye in recognizing structural change between time bins. The size of each panel is 40 kpc.

a scenario with ~ 2 satellites, instead of 15 if each stream is linked to a single event (Tanaka et al. 2010).

A major merger paradigm is supported by the widespread age–dispersion relation in the disc of M31. In fact, a 21:1 merger cannot provide the velocity dispersion amplitude observed in stars lying in the overall M31 disc: for example, at $R = 10$ kpc, fig. 20 of Sadoun, Mohayaee & Colin (2014) predicts $\sim 30 \text{ km s}^{-1}$ that compares to $80\text{--}100 \text{ km s}^{-1}$ observed for the bulk of stars by Dorman et al. (2015). It is recovered by our models (see Figs 6 and 7), due to the highly retrograde orbit of the primary adopted in this paper (see Table 2), which is particularly efficient in stressing the primary disc. We verified that $\geq 5:1$ mass ratios are not sufficient to reproduce the observed velocity dispersion amplitudes. A recent major merger is also consistent with important properties of M31, which are not successfully explained for the moment:

(i) The 10 kpc ring that is a long-lived feature that has survived for at least 400 Myr (Davidge et al. 2012; Dalcanton et al. 2012; Lewis et al. 2015); Fig. 13 shows that in a 4:1 post-merger, the ring is naturally stable for at least 500 Myr and our modelling also captures the presence of the outer ring at 15 kpc (see Fig. 3).

(ii) The 2–4 Gyr star formation event that is widespread over the whole M31 disc (see e.g. discussion in Williams et al. 2015) and consistent with a recent 4:1 merger (see Figs 4 and 5).

(iii) The NE and G Clumps are predicted together with a warp in a major merger; Ferguson & Mackey (2016) recalled that these structures have been supposedly linked to the disruption of a dwarf, which is the case in our modelling (more precisely a tidal dwarf), together with the fact that they have disc-like properties such as the warps.

(iv) The halo profile predicted by a 3–4:1 major merger (see Fig. 14) is consistent with the observations (Gilbert et al. 2012; Ibata et al. 2014), with a constant (projected) slope of $\gamma = -2.1$ to -2.3 , and without a drop until 120 kpc. This property is shared by modern cosmological simulations including major mergers (see e.g. fig. 4 of Font et al. 2011); in contrast and as noticed by Gilbert et al. (2012), 10 among 11 simulations of the dwarf dominated simulations by Bullock & Johnston (2005) show a break to a steeper profile before or at 100 kpc.

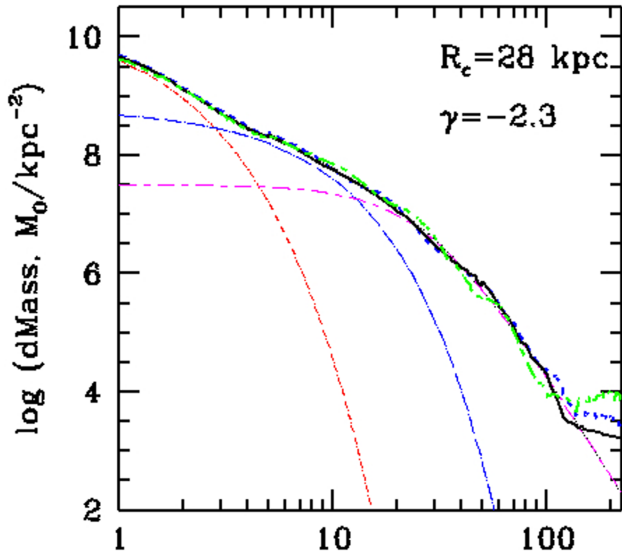


Figure 14. Stellar halo profile for $mr = 4$ (thick black line, average of the five models of Table 2), $mr = 3$ (thick dot, blue line) and $mr = 5$ (thick, short dash–long dash, green line). Models of the bulge, the disc and the halo are provided by thin lines, dot–short dash (red), dot–long dash (blue), and short dash–long dash (magenta) lines, respectively. The adopted halo profile follows a cored power law $(1 + (R/R_c)^2)^\gamma$, as in Ibata et al. (2014). The fit of the halo with $\gamma = -2.1$ to -2.3 is valid until 100, 120, and 130 kpc for models with $mr = 5, 4,$ and $3,$ respectively. Note that the progenitors in our simulations have no initial stellar haloes.

(v) The absence of evidence for a residual core of the satellite responsible of the GS formation is expected since all the secondary material has been wiped out into the different structures of the remnant (see Fig. 10).

The 3D structure of the GS may provide a further, and possibly definitive test of the GS origin. Conn et al. (2016) measure the distance from the MW of stars along the GS. It reveals some complexity from the mid-part to the end of the GS, for which they found probability distributions that are often double peaked. Fig. 15 suggests that these multiple peaks are real and are caused by a complex 3D distribution as is expected if the GS is the superposition of two or even three loops (see left-hand panels). If correct, this causes an increasing number of different density peaks with distance from M31, in qualitative agreement with what is observed. Such complexity may be difficult to explain if the GS is a single tidal tail caused by the passage of a satellite (see e.g. fig. 8 of Kirihaara et al. 2017). Notice that contamination by other components is unlikely in fields this far from the M31 thick disc.

4.3 Concluding remarks

Galaxies are made by mergers of smaller entities according to the hierarchical scenario, for which M31 is perhaps one of the best illustrations. The relative success of a major merger for explaining most of the large-scale properties of M31 does not mean that we have securely identified its past history and fully determined the parameter set of the former interaction. We reach robust, though not precise, predictions for the mass ratio ($<4.5:1$) and for the pericentre (≥ 30 kpc), in order to build sufficiently large stellar and H I discs. This leads to relatively precise predictions for the coalescence of the nuclei epoch (1.8–3 Gyr ago) to reproduce the age–dispersion relation and the 2–4 Gyr old star formation event, and for the angular orbital parameters, which lie within a small range of values to reproduce the 10 kpc ring and the GS (see Table 2). This does not preclude of any other, additional minor merger or interaction, such as that with M33 (see e.g. McConnachie et al. 2009, 2010) creating a faint and shallow bridge between the

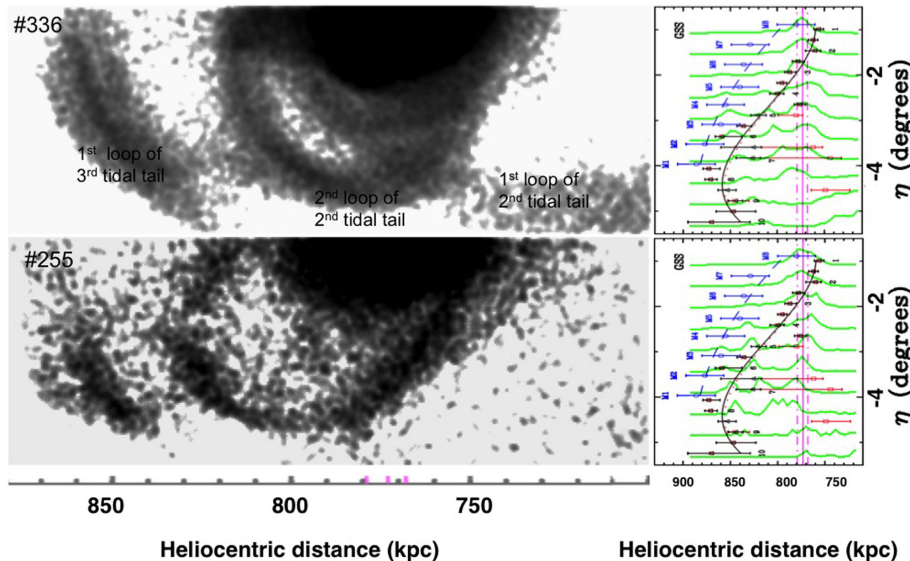


Figure 15. 3D structure of the GS for models 336 (top, 20M particles) and 255 (bottom, 2M particles). The centre of M31 is taken at 773 kpc from the MW (see magenta tick marks and vertical lines in left- and right-hand panels, respectively) as in Conn et al. (2016). Left: zoom-on the GS components (or loops) for model 336 in a (Z, Y) plane. Z corresponds to distance from the MW (from left to right: 873–673 kpc), and Y is the distance in kpc from the M31 centre in the direction from north to south (from top to down: 20–100 kpc). Right: reproduction of fig. 4 of Conn et al. (2016) for which the solid line represents the best estimate of the GS distance distribution, and red squares correspond to alternatives due to multiple peaks. Modelled mass profiles (green lines, extracted from horizontal cuts of the left-hand image) have been superposed on the observations. Although the simulations do not reproduce accurately the observations, they do show a complexity that is quite similar. As in the observations they show a single peak at the top, then multiple peaks near -3° and below. Models also show that from the top to bottom, distances are increasing and then slightly decrease, as in the observations.

two galaxies. If correct, the two events would have happened at a similar epoch (few Gyr ago), and a fly-by with M33 is likely less influential than a major merger at the time when the two nuclei have coalesced.

An important conclusion of this paper is that the GS and shell modelling depend sensitively on angular orbital parameters, but is almost (or fully) mass-ratio independent, i.e. they can be produced either by a single minor or major event.

Furthermore, a recent major merger appears to be a rather compelling scenario to explain the disc age–dispersion relation, and to reproduce a realistic 10 kpc ring, the 2–4 Gyr old star formation event, the NE and G clump properties, the halo profile, and perhaps the 3D structure of the GS.

The outer disc structures of M31 have bimodal age and metallicity properties: GS and shells on one side, disc-like structures such as warp and clumps on the other side. This is precisely mirrors the material that originated in the secondary and the primary, respectively (see Fig. 10). Fig. 16 compares an rgb decomposition of the metal abundances for model 314 with what is observed in M31 (Ibata et al. 2014). They both show an increase of star number density with metallicity (compare b and d panels and then f and h panels), and a quasi absence of low-metallicity stars at the location of the GS. M31 has been found to have a more spherical halo at low metallicity (Ibata et al. 2014), consistent with these simulations.

Intermediate- and high-metallicity stars of model 314 show a hint for a possible bifurcation at the GS end, which might suggest streams B and C. However, the simulations would require ~ 16 times more particles to match the observations. To overcome this difficulty would require us to compute a significant number of 20M particle simulations, which is far beyond the scope of this study, since it would need a significant increase in calculation speed. In this paper, we are left to speculate that stream B (and perhaps stream C) might have originated in the secondary progenitor, while stream D might belong to the very outer tidal tail of the primary (see and compare low-metallicity panels d and h in Fig. 16).

Predictions or tests are often useful to support (or to dismiss) a scenario. If M31 has been structured after a recent, major merger responsible for the formation of the GS, we find that returning stars from the second tidal tail should be at least, barely observable at the depth of the PAndAS survey. Our modelling then predicts a wide and ultra faint overdensity of stars to the North of the M31 halo, in the region including And XXVII, XXV, XXVI, and perhaps also NGC 147 and 185 (see e.g. fig. 1 of Richardson et al. 2011). In Fig. 14 one may notice, at large radii (> 120 kpc), a flattening of the profile instead of a steepening. Such a flattening is due to the recent coalescence of the nuclei since it vanishes at later times. We then predict an overabundance of (especially intermediate and metal rich) stars at 130–200 kpc projected distance from M31, on both Northern and Southern sides, while we would expect an underabundance on the Western side.

Finally, one may notice that tidal tails attached to the secondary lie in a single thin plane, which is the orbital plane. In all models shown in Table 2, it includes the MW, i.e. matching very well the plane of satellites found by Ibata et al. (2013). This result seems robust and linked to our viewpoint at the MW location, which requires us to see the GS as the superposition of loops, i.e. almost edge-on. Such a link may lead to clues about the origin of vast planar structures of satellites, which are not fully understood yet (Pawlowski et al. 2014). Could their origin be linked to an interaction between two satellite systems during the merging, and under which conditions they can

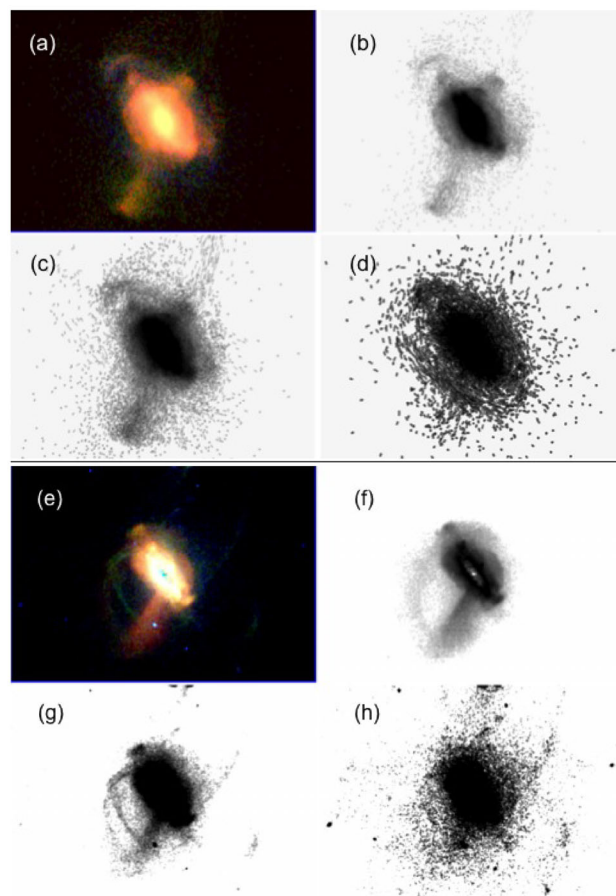


Figure 16. Metal distribution from model 314 (top four panels, a–d) and that observed in M31 (bottom four panels, e–h), for which the resolution is 16 times higher than the former. Top left: RGB colour combination of stars with $\log(Z/Z_{\odot})$ in the range of -2.5 to -1.8 (blue), -1.8 to -1 (green), and 1 to 0 (red), respectively (panel a: modelling and panel e: observations). In the modelling, panels b–d (observations: panels f–h) represent the corresponding distribution of low (blue), intermediate (green), and high (red) metallicity (coded stars), respectively. In the progenitors, we have estimated metallicity assuming it decreases radially by -0.1 dex kpc^{-1} , similar to expectations in the MW, 7 Gyr ago (see e.g. Prantzos 2007). Initial average values for each progenitor have been taken from Rodrigues et al. (2012), assuming the observed slope of the mass–metallicity relation at $z \sim 1.5$ (9 Gyr ago).

lead to such a thin structure as that surrounding M31 (Smith et al. 2016)? If the thinness of the plane of satellites is linked to the thinness of the merger orbital plane, it could even lead one to question the nature of M31 dwarfs, as suggested by Hammer et al. (2013). In our modelling, the G1 and NE clumps are disrupted tidal dwarfs (or fluctuations in a tidal tail), and they would be linked with tidal tails attached to the primary. The latter formed at the same epoch as the (second) tidal tail in the orbital plane, i.e. ~ 1 Gyr before the coalescence of the nuclei, which would not give much predictive power to comparisons between dwarfs in and out the plane (Collins et al. 2015, 2017). Future simulations with large particle numbers have to be done, including, e.g. interactions between pre-existing satellites and/or tidal dwarf systems, to verify under which conditions the former may provide thin planes of satellites and/or the later may provide dwarfs with stars as old as that in the M31 dwarfs, knowing that part of their tidal material is common with that of the GS (see Fig. 9).

ACKNOWLEDGEMENTS

This work was granted access to the HPC resources of TGCC/CINES/IDRIS under the allocation 2016-(i2016047633) made by GENCI, and to MesoPSL financed by the ‘Region Ile de France’ and the project Equip@Meso (reference ANR-10-EQPX-29-01) of the program ‘Investissements d’Avenir’ supervised by the ‘Agence Nationale de la Recherche’. This work has been supported by the China-France International Associated Laboratory ‘Origins’. JLW thanks the China Scholarship Council (No. 201604910336) for the financial support. We are grateful to Phil Hopkins who kindly shared with us the access to the GIZMO code. FH warmly thank Lia Athanassoula for her precious and wise advices on modelling substructures of an individual galaxy.

REFERENCES

- Amorisco N. C., 2015, *MNRAS*, 450, 575
 Athanassoula E., Beaton R. L., 2006, *MNRAS*, 370, 1499
 Barmby P. et al., 2006, *ApJ*, 650, L45
 Barnes J. É., 2002, *MNRAS*, 333, 481
 Beaton R. L. et al., 2007, *ApJ*, 658, L91
 Bernard E. J. et al., 2012, *MNRAS*, 420, 2625
 Bernard E. J., Ferguson A. M. N., Chapman S. C., Ibata R. A., Irwin M. J., Lewis G. F., McConnachie A. W., 2015, *MNRAS*, 453, L113
 Bernard E. J. et al., 2015, *MNRAS*, 446, 2789
 Blaña Díaz M., Wegg C., Gerhard O., Erwin P., Portail M., Opitsch M., Saglia R., Bender R., 2017, *MNRAS*, 466, 4279
 Block D. L. et al., 2006, *Nature*, 443, 832
 Brown T. M., Smith E., Ferguson H. C., Rich R. M., Guhathakurta P., Renzini A., Sweigart A. V., Kimble R. A., 2006, *ApJ*, 652, 323
 Brown T. M. et al., 2007, *ApJ*, 658, L95
 Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
 Bullock J. S., Boylan-Kolchin M., 2017, *ARA&A*, 55, 343
 Bullock J. S., Johnston K. V., 2005, *ApJ*, 635, 931
 Chemin L., Carignan C., Foster T., 2009, *ApJ*, 705, 1395
 Collins M. L. M. et al., 2015, *ApJ*, 799, L13
 Collins M. L. M., Rich R. M., Ibata R. P., Martin N., Preston J., PAndAS Collaboration, 2017, in *Proc. IAU Symp. 321, Formation and Evolution of Galaxy Outskirts*. Kluwer, Dordrecht, p. 16
 Conn A. R. et al., 2016, *MNRAS*, 458, 3282
 Cooper A. P. et al., 2011, *ApJ*, 743, L21
 Courteau S. et al., 2011, *ApJ*, 739, 20
 Cox T. J., Jonsson P., Primack J. R., Somerville R. S., 2006, *MNRAS*, 373, 1013
 Dalcanton J. J. et al., 2012, *ApJS*, 200, 18
 Dalcanton J. J. et al., 2015, *ApJ*, 814, 3
 Davidge T. J., McConnachie A. W., Fardal M. A., Fliri J., Valls-Gabaud D., Chapman S. C., Lewis G. F., Rich R. M., 2012, *ApJ*, 751, 74
 Del Popolo A., Le Delliou M., 2017, *Galaxies*, 5, 17
 Dierickx M., Blecha L., Loeb A., 2014, *ApJ*, 788, L38
 Dorman C. E. et al., 2015, *ApJ*, 803, 24
 Fardal M. A., Babul A., Geehan J. J., Guhathakurta P., 2006, *MNRAS*, 366, 1012
 Fardal M. A., Guhathakurta P., Babul A., McConnachie A. W., 2007, *MNRAS*, 380, 15
 Fardal M. A., Babul A., Guhathakurta P., Gilbert K. M., Dodge C., 2008, *ApJ*, 682, L33
 Fardal M. A. et al., 2013, *MNRAS*, 434, 2779
 Ferguson A. M. N., Mackey A. D., 2016, *Astrophysics and Space Science Library*, Vol. 420, *Tidal Streams in the Local Group and Beyond*. Springer International Publishing, Switzerland, p. 191
 Font A. S., McCarthy I. G., Crain R. A., Theuns T., Schaye J., Wiersma R. P. C., Dalla Vecchia C., 2011, *MNRAS*, 416, 2802
 Gilbert K. M. et al., 2009, *ApJ*, 705, 1275
 Gilbert K. M. et al., 2012, *ApJ*, 760, 76
 González Delgado R. M. et al., 2017, *A&A*, 607, A128
 Gordon K. D. et al., 2006, *ApJ*, 638, L87
 Hammer F., Flores H., Elbaz D., Zheng X. Z., Liang Y. C., Cesarsky C., 2005, *A&A*, 430, 115
 Hammer F., Puech M., Chemin L., Flores H., Lehnert M. D., 2007, *ApJ*, 662, 322
 Hammer F., Flores H., Puech M., Yang Y. B., Athanassoula E., Rodrigues M., Delgado R., 2009, *A&A*, 507, 1313
 Hammer F., Yang Y. B., Wang J. L., Puech M., Flores H., Fouquet S., 2010, *ApJ*, 725, 542
 Hammer F., Yang Y., Fouquet S., Pawlowski M. S., Kroupa P., Puech M., Flores H., Wang J., 2013, *MNRAS*, 431, 3543
 Hopkins P. F., 2013, *MNRAS*, 428, 2840
 Hopkins P. F., 2014, *Astrophysics Source Code Library*, record ascl:1410.003
 Hopkins P. F., Cox T. J., Younger J. D., Hernquist L., 2009, *ApJ*, 691, 1168
 Hopkins P. F. et al., 2010, *ApJ*, 715, 202
 Ibata R., Irwin M., Lewis G., Ferguson A. M. N., Tanvir N., 2001, *Nature*, 412, 49
 Ibata R., Chapman S., Ferguson A. M. N., Lewis G., Irwin M., Tanvir N., 2005, *ApJ*, 634, 287
 Ibata R. A. et al., 2013, *Nature*, 493, 62
 Ibata R. A. et al., 2014, *ApJ*, 780, 128
 Jurić M. et al., 2008, *ApJ*, 673, 864
 Katz N., Weinberg D. H., Hernquist L., 1996, *ApJS*, 105, 19
 Kennicutt R. C., Jr, 1998, *ApJ*, 498, 541
 Kirihaara T., Miki Y., Mori M., 2014, *PASJ*, 66, L10
 Kirihaara T., Miki Y., Mori M., Kawaguchi T., Rich R. M., 2017, *MNRAS*, 464, 3509
 Kormendy J., 2013, in *Falcón-Barroso J., Knapen J. H., eds, Secular Evolution of Galaxies, XXIII Canary Islands Winter School of Astrophysics*, Cambridge Univ. Press, Cambridge, p. 1
 Lewis G. F. et al., 2013, *ApJ*, 763, 4
 Lewis A. R. et al., 2015, *ApJ*, 805, 183
 McConnachie A. W. et al., 2009, *Nature*, 461, 66
 McConnachie A. W., Ferguson A. M. N., Irwin M. J., Dubinski J., Widrow L. M., Dotter A., Ibata R., Lewis G. F., 2010, *ApJ*, 723, 1038
 Mori M., Rich R. M., 2008, *ApJ*, 674, L77
 Opitsch M., Fabricius M. H., Saglia R. P., Bender R., Blaña M., Gerhard O., 2017, preprint ([arXiv:1707.06652](https://arxiv.org/abs/1707.06652))
 Pawlowski M. S. et al., 2014, *MNRAS*, 442, 2362
 Prantzos N., 2007, *EAS Publ. Ser.*, 24, 3
 Richardson J. C. et al., 2011, *ApJ*, 732, 76
 Rodrigues M., Puech M., Hammer F., Rothberg B., Flores H., 2012, *MNRAS*, 421, 2888
 Rodrigues M., Hammer F., Flores H., Puech M., Athanassoula E., 2017, *MNRAS*, 465, 1157
 Sadoun R., Mohayaee R., Colin J., 2014, *MNRAS*, 442, 160
 Smith R., Duc P. A., Bournaud F., Yi S. K., 2016, *ApJ*, 818, 11
 Springel V., 2000, *MNRAS*, 312, 859
 Springel V., 2005, *MNRAS*, 364, 1105
 Stewart K. R., Bullock J. S., Wechsler R. H., Maller A. H., Zentner A. R., 2008, *ApJ*, 683, 597
 Tanaka M., Chiba M., Komiyama Y., Guhathakurta P., Kalirai J. S., Iye M., 2010, *ApJ*, 708, 1168
 Tenjes P., Tuvikene T., Tamm A., Kipper R., Tempel E., 2017, *A&A*, 600, A34
 van den Bergh S., 2005, in *Livio M., Brown T.M., eds, The Local Group as an Astrophysical Laboratory*. Cambridge Univ. Press, Cambridge, p. 1
 Wang J., Hammer F., Athanassoula E., Puech M., Yang Y., Flores H., 2012, *A&A*, 538, A121
 Wang J., Hammer F., Puech M., Yang Y., Flores H., 2015, *MNRAS*, 452, 3551
 Williams B. F. et al., 2015, *ApJ*, 806, 48
 Zucker D. B. et al., 2004, *ApJ*, 612, L117

APPENDIX A: MASS PROFILE AND ROTATION CURVES

Fig. A1 shows the stellar mass profiles (top panels) and the rotation curves (bottom panels) for the six models tabulated in Table 2. Based on these fits, B/T ratios are ranging from 0.35 to 0.42, which is consistent with observations after assuming that B accounts for both bar and bulge.

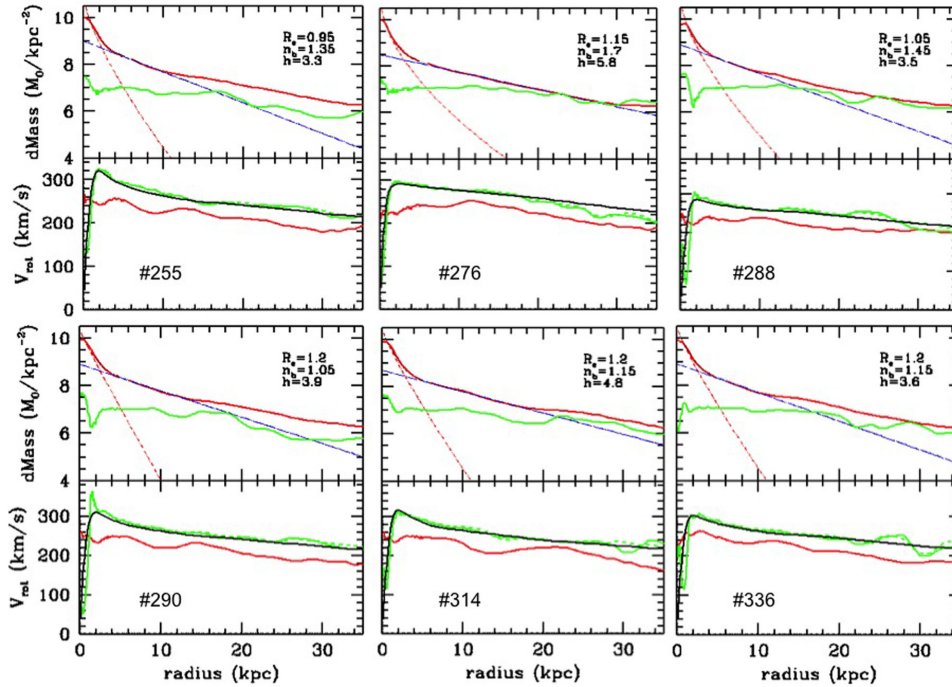


Figure A1. Mass profile and rotation curve for six models (see the model number on each bottom panel) up to 35 kpc from the centre. Top panels: stellar mass profile (red) that is fitted by two components (red dashed line: bulge with a free Sérsic index and blue long-dashed line: exponential disc). Green line indicates the gas profile. On the top right of each panel, the bulge effective radius and Sérsic index, as well as the disc scale length, are given. Bottom panels: rotation curve for each component (red: stars and green: gas) and from the virial theorem (black curve). Notice that the gas is sufficiently relaxed after the 1.8–3 Gyr old merger to follow well the theoretical curve. It also matches quite well the observed rotation curve by Chemin, Carignan & Foster (2009).

APPENDIX B: CONVERGENCE OF SIMULATIONS

Fig. B1 presents the results of the same simulation of a 3:1 merger at three different resolutions. Snapshots have been captured 7.2, 6.2, and 3.5 Gyr after the beginning of the simulation, the first passage and the coalescence of the nuclei, respectively. Notice that most details in the loops and tails are well preserved after passing from 500k to 6 M particles (see Wang et al. in preparation).

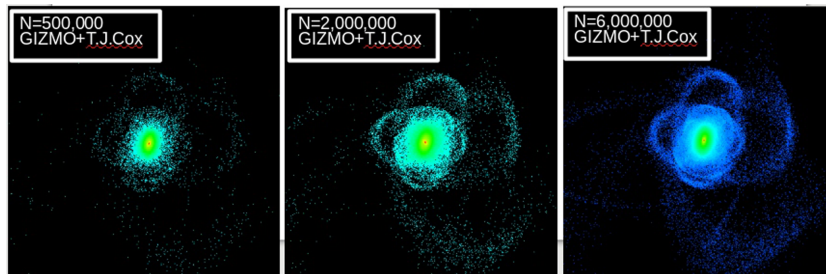


Figure B1. From left to right: result of a simulation of a 3:1 gas-rich merger showing loops and tidal tails in the orbital plane with 500 k, 2M, and 6M particles, respectively. Pericentre is 25 kpc, and initial gas fractions are 50 and 80 per cent in primary and secondary progenitors, respectively.

This paper has been typeset from a \LaTeX file prepared by the author.