

# 3D hydrodynamic simulations for the formation of the Local Group satellite planes

Indranil Banik<sup>1,2,★</sup>, Ingo Thies<sup>2</sup>, Roy Truelove<sup>1</sup>, Graeme Candlish<sup>3</sup>, Benoit Famaey<sup>4</sup>,  
Marcel S. Pawlowski<sup>5</sup>, Rodrigo Ibata<sup>4</sup> and Pavel Kroupa<sup>5,6</sup>

<sup>1</sup>Scottish Universities Physics Alliance, University of Saint Andrews, North Haugh, Saint Andrews, Fife KY16 9SS, UK

<sup>2</sup>Helmholtz-Institut für Strahlen und Kernphysik (HISKP), University of Bonn, Nussallee 14 – 16, D-53115 Bonn, Germany

<sup>3</sup>Instituto de Física y Astronomía, Universidad de Valparaíso, Gran Bretaña 1111, Valparaíso, Chile

<sup>4</sup>CNRS UMR 7550, Observatoire astronomique de Strasbourg, Université de Strasbourg, 11 rue de l'Université, F-67000 Strasbourg, France

<sup>5</sup>Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany

<sup>6</sup>Astronomical Institute, Faculty of Mathematics and Physics, Charles University, V Holešovičkách 2, CZ-180 00 Praha 8, Czech Republic

Accepted 2022 March 11. Received 2022 February 8; in original form 2021 June 1

## ABSTRACT

The existence of mutually correlated thin and rotating planes of satellite galaxies around both the Milky Way (MW) and Andromeda (M31) calls for an explanation. Previous work in Milgromian dynamics (MOND) indicated that a past MW–M31 encounter might have led to the formation of these satellite planes. We perform the first-ever hydrodynamical MOND simulation of the Local Group using PHANTOM OF RAMSES. We show that an MW–M31 encounter at  $z \approx 1$ , with a perigalactic distance of about 80 kpc, can yield two disc galaxies at  $z = 0$  oriented similarly to the observed galactic discs and separated similarly to the observed M31 distance. Importantly, the tidal debris are distributed in phase space similarly to the observed MW and M31 satellite planes, with the correct preferred orbital pole for both. The MW–M31 orbital geometry is consistent with the presently observed M31 proper motion despite this not being considered as a constraint when exploring the parameter space. The mass of the tidal debris around the MW and M31 at  $z = 0$  compare well with the mass observed in their satellite systems. The remnant discs of the two galaxies have realistic radial scale lengths and velocity dispersions, and the simulation naturally produces a much hotter stellar disc in M31 than in the MW. However, reconciling this scenario with the ages of stellar populations in satellite galaxies would require that a higher fraction of stars previously formed in the outskirts of the progenitors ended up within the tidal debris, or that the MW–M31 interaction occurred at  $z > 1$ .

**Key words:** gravitation – hydrodynamics – galaxies: formation – galaxies: interactions – galaxies: kinematics and dynamics – Local Group.

## 1 INTRODUCTION

It has been known since the early work of Kunkel & Demers (1976) and Lynden-Bell (1976, 1982) that dwarf spheroidal galaxies around the Milky Way (MW) have an anisotropic spatial distribution, as later confirmed by Kroupa, Theis & Boily (2005) and Metz, Kroupa & Jerjen (2007). The orbital poles that can be measured indicate that this Vast Polar Structure (VPOS) is corotating (Metz, Kroupa & Libeskind 2008; Pawlowski, Kroupa & Jerjen 2013; Pawlowski et al. 2017; Pawlowski & Kroupa 2020; Li et al. 2021). This is in stark contrast with expectations based on the standard  $\Lambda$  cold dark matter ( $\Lambda$ CDM) cosmological model (Efstathiou, Sutherland & Maddox 1990; Ostriker & Steinhardt 1995) because the existence of such a corotating plane would *a priori* imply a significant amount of dissipation. The issue is compounded by a similar satellite plane (SP) discovered around Andromeda (M31; Ibata et al. 2013, 2014b) and Centaurus A (Cen A; Müller et al. 2018a, 2021), with radial velocities

(RVs) suggestive of corotation in both cases.<sup>1</sup> Proper motions (PMs) have recently become known for two M31 SP members, and indeed indicate that it too is likely corotating within its plane (Sohn et al. 2020).

Such structures are difficult to explain using dissipationless haloes of CDM, as first pointed out by Kroupa et al. (2005) and more recently by Pawlowski et al. (2014), who considered and excluded a wide range of different proposed explanations within the  $\Lambda$ CDM context. For instance, accreting most satellites as a single group (Metz et al. 2009b) or along a single filament would yield some anisotropy, but not enough to explain the very thin MW SP (Shao et al. 2018). The two available PMs of M31 SP members are also in tension with  $\Lambda$ CDM expectations because they indicate motion nearly within the M31 SP (Pawlowski & Tony Sohn 2021). Dissipationless collapse

<sup>1</sup>Hints of the M31 SP were already evident in Metz et al. (2007) and Metz, Kroupa & Jerjen (2009a). Hints of the Cen A SP were evident in Tully et al. (2015), but the two planes they identified were later revealed to be part of one thicker plane (Müller et al. 2016).

\* E-mail: [indranilbanik1992@gmail.com](mailto:indranilbanik1992@gmail.com)

of CDM haloes is therefore insufficient to account for the observed SPs if they are composed of primordial dwarfs. Including baryonic physics does not change this picture very much (Ahmed, Brooks & Christensen 2017; Pawlowski & Kroupa 2020; Samuel et al. 2021).

Since dissipation in the CDM component would cause a significant Galactic DM disc that is in tension with observations (Buch, Leung & Fan 2019), an obvious possibility is that the necessary dissipation occurred in baryons. Given that the MW SP is almost polar with respect to its disc, this would require a tidal interaction with another galaxy (Pawlowski, Kroupa & de Boer 2011; Pawlowski, Pflamm-Altenburg & Kroupa 2012). In this scenario, the MW SP would consist of tidal dwarf galaxies (TDGs) that condensed out of gas-rich tidal debris expelled during a past interaction. A similar scenario could have occurred around M31. The formation of TDGs has been observed outside the Local Group (LG), for instance around the Antennae (Mirabel, Dottori & Lutz 1992) and the Seashell Galaxy (Bournaud et al. 2004). The LG SPs may have formed analogously. A common origin for both LG SPs is possible as a result of a major merger experienced by M31 (Hammer et al. 2010) expelling tidal debris towards the MW (Hammer et al. 2013), perhaps explaining the high Galactocentric velocities of MW satellites (Hammer et al. 2021). Since the M31 SP is viewed close to edge-on from our vantage point in the MW (Conn et al. 2013; Ibata et al. 2013; Santos-Santos, Domínguez-Tenreiro & Pawlowski 2020), it is possible that if M31 experienced a major merger, then some of the tidal debris expelled from M31 formed its SP while some reached a much larger distance and is now close to the MW.

In a  $\Lambda$ CDM context, TDGs would be free of DM due to its dissipationless nature and its initial distribution in a dispersion-supported near-spherical halo (Barnes & Hernquist 1992; Wetzstein, Naab & Burkert 2007). During a tidal interaction, DM of this form is clearly incapable of forming into a thin dense tidal tail which might undergo Jeans collapse into TDGs. As a result, any TDGs would be purely baryonic and thus have only a very small escape velocity, preventing them from subsequently accreting DM out of the Galactic halo. For this reason, TDGs formed in the  $\Lambda$ CDM framework cannot explain the high observed internal velocity dispersions ( $\sigma_{\text{int}}$ ) of the MW satellites if they are in equilibrium. Non-equilibrium solutions were considered by Kroupa (1997), Klessen & Kroupa (1998), and Casas et al. (2012). While their proposed solutions match many properties of the observed MW satellites, this scenario would not yield elevated  $\sigma_{\text{int}}$  values at Galactocentric distances  $\gtrsim 150$  kpc. Satellites in such a non-equilibrium phase would also be very fragile, requiring us to be observing them at just the right epoch prior to total disruption but after significant tidal perturbation. These issues are undoubtedly a major reason why CDM-free galaxies are very rare in the latest  $\Lambda$ CDM simulations (Haslbauer et al. 2019), making it very unlikely that so many TDGs are in the LG right now around both the MW and M31. Moreover, postulating that most of their satellites are TDGs would mean that they have very few primordial satellites.

This conundrum presents an open invitation to consider a different theoretical framework in which DM is dynamically irrelevant to holding galaxies together against their high  $\sigma_{\text{int}}$ , removing the problem that TDGs are expected to be free of DM. In this case, the effects conventionally attributed to CDM must instead be due to a non-Newtonian gravity law on galactic scales. The best developed such proposal is Milgromian dynamics (MOND; Milgrom 1983). In MOND, galaxies lack CDM but the gravitational field strength  $g$  at distance  $d$  from an isolated point mass  $M$  transitions from the Newtonian  $g_N = GM/d^2$  law at short range to

$$g = \sqrt{g_N a_0} \quad \text{for } g_N \ll a_0. \quad (1)$$

MOND introduces  $a_0$  as a fundamental acceleration scale of nature below which the deviation from Newtonian dynamics becomes significant. Empirically,  $a_0 \approx 1.2 \times 10^{-10} \text{ m s}^{-2}$  to match galaxy rotation curves (RCs; Begeman, Broeils & Sanders 1991; Gentile, Famaey & de Blok 2011). With this value of  $a_0$ , MOND predicts the detailed shape of galaxy RCs very well using only their directly observed baryonic matter (e.g. Kroupa et al. 2018a; Li et al. 2018; Sanders 2019). In particular, observations confirm the prior MOND prediction of very large departures from Newtonian dynamics in low surface brightness galaxies (LSBs; McGaugh 2021, and references therein). More generally, there is a very tight empirical ‘radial acceleration relation’ (RAR) between the gravity inferred from RCs and that expected from the baryons alone in Newtonian dynamics (McGaugh, Lelli & Schombert 2016; Lelli et al. 2017), with RCs asymptotically reaching a flatline level  $\propto \sqrt[4]{M}$  as required by equation (1) (McGaugh 2012; Lelli et al. 2019). The observed phenomenology of galactic RCs confirm all the central predictions of Milgrom (1983), as reviewed in e.g. Famaey & McGaugh (2012).

MOND can also explain the X-ray temperature profile (Milgrom 2012) and internal dynamics of elliptical galaxies (fig. 8 of Lelli et al. 2017), which reveal a similar characteristic acceleration scale to spirals (Chae et al. 2020a; Shelest & Lelli 2020). At the low-mass end, MOND is consistent with the  $\sigma_{\text{int}}$  of pressure-supported systems like the satellites of the MW (McGaugh & Wolf 2010), M31 (McGaugh & Milgrom 2013a,b), non-satellite LG dwarfs (McGaugh et al. 2021), Dragonfly 2 (DF2; Famaey, McGaugh & Milgrom 2018; Kroupa et al. 2018b), and DF4 (Haghi et al. 2019a). These predictions rely on correctly including the external field effect (EFE) arising from the non-linearity of MOND (Milgrom 1986), which we discuss further in Section 2.4.3. For the more isolated galaxy DF44, the MOND prediction without the EFE is consistent with the observed  $\sigma_{\text{int}}$  profile (Bilek, Müller & Famaey 2019; Haghi et al. 2019b). Note that the situation with the EFE is however less clear for ultradiffuse galaxies located deep in the potential well of galaxy clusters (Freundlich et al. 2022). The EFE also plays a role in accounting for the observed weak bar of M33 (Banik et al. 2020), which is quite difficult to understand in the presence of a live DM halo (Sellwood, Shen & Li 2019) due to bar-halo angular momentum exchange (e.g. Debattista & Sellwood 2000; Athanassoula 2002). This problem is related to the bar pattern speeds in galaxies, which seem to be too slow in  $\Lambda$ CDM cosmological simulations due to dynamical friction on the bar exerted by the DM halo (Algorry et al. 2017; Peschken & Łokas 2019; Roshan et al. 2021a,b; though see Fragkoudi et al. 2021). In addition, the morphological properties of MW satellites are more easily understood in MOND as a consequence of differing levels of tidal stability (McGaugh & Wolf 2010). This is because in MOND, the lack of CDM haloes and the EFE render the satellites much more susceptible to Galactic tides, which are not so relevant in  $\Lambda$ CDM (see their fig. 6). Neglecting tides leads to erroneous conclusions regarding the viability of MOND since it predicts that much fewer satellites – especially at the ultrafaint end – are amenable to equilibrium virial analysis (Fattahi et al. 2018).

Further work will however be necessary to make rigorous MOND predictions in a cosmological context, which needs a relativistic theory. A few such theories exist, with one promising proof-of-concept being the relativistic MOND theory of Skordis & Złośnik (2019) in which gravitational waves travel at the speed of light. Such theories allow MOND calculations of weak gravitational lensing by foreground galaxies in stacked analyses (e.g. Brimiouille et al. 2013), which so far seems to agree with expectations (Milgrom 2013; Brouwer et al. 2017, 2021). At larger distances from the central galaxy, the EFE from surrounding structures would cause the

gravity law to become inverse square and to depart from spherical symmetry (Banik & Zhao 2018a), which may explain some recent observations (Schrabback et al. 2021). More detailed calculations would require knowledge of how large-scale structure forms in a Milgromian framework and the resulting EFE on galaxies. This is a critical next step for MOND, though care is required when comparing with observations as these could have a rather different interpretation to what is usually assumed. One possible smoking gun signature of MOND in weak lensing convergence maps would be the discovery that the convergence parameter is negative in some regions (Oria et al. 2021). This cannot arise in GR and is possible only in gravitational theories that are non-linear in the weak-field regime.

At this stage, general statements can however already be made about a MONDian cosmology, in which the key difference with  $\Lambda$ CDM would be faster structure formation (Sanders 1998). This might be relevant for the so-called Hubble tension, i.e. the fact that, to fit the anisotropies in the cosmic microwave background (CMB; Aiola et al. 2020; Planck Collaboration VI 2020),  $\Lambda$ CDM requires a local Hubble constant  $H_0$  below the directly measured value at high significance based on multiple independent techniques (e.g. Riess 2020; Di Valentino 2021; Riess et al. 2021b). This tension could be due to our position within a large local supervoid underdense by  $\approx 30$  per cent out to a radius of  $\approx 300$  Mpc. Such a large and deep underdensity has indeed been observed in multiple surveys and is called the KBC void after its discoverers (Keenan, Barger & Cowie 2013). This is incompatible with  $\Lambda$ CDM at  $6.04\sigma$ , one major reason for the high significance of the tension being that the relevant observations cover 90 per cent of the sky (Haslbauer, Banik & Kroupa 2020). However, a KBC-like void could arise naturally in a MOND cosmology supplemented by light sterile neutrinos playing the role of hot DM (HDM), as proposed by Angus (2009). The main feature in the Haslbauer et al. (2020) void scenario is faster structure formation than in  $\Lambda$ CDM. From an observational point of view, this is also evident in the properties of the high-redshift interacting galaxy cluster El Gordo (Asencio, Banik & Kroupa 2021). The lack of analogues to El Gordo at low redshift could well be due to our location within the KBC void, which might also explain why the MOND simulations of Angus et al. (2013) seemingly overproduced massive clusters when comparing their whole simulation volume with low-redshift data sets. Therefore, MOND with HDM could potentially account for astronomical observations ranging from the kpc scales of galaxies (where HDM would play no role; Angus 2010) all the way to the Gpc scale of the local supervoid, without causing any obvious problems in the early Universe or in galaxy clusters (Angus, Famaey & Diaferio 2010). It is also possible to fit the CMB in MOND without any sterile neutrino component (Skordis & Złotnik 2021), though it is unclear whether this approach can explain the properties of galaxy clusters. For a recent review of MOND that considers evidence from a wide range of scales and discusses the cosmological aspects in some detail, we refer the reader to Banik & Zhao (2022).

Concerning the history of the well-observed LG, MOND implies a very strong mutual attraction between the MW and M31. Acting on their almost radial orbit (van der Marel et al. 2012), this leads to a close encounter  $9 \pm 2$  Gyr ago (Zhao et al. 2013), consistent with the timing of a few other events putatively linked to the flyby (Section 4.2). An  $N$ -body simulation of this interaction showed that it is likely to yield anisotropically distributed tidal debris reminiscent of an SP (Bílek et al. 2018). Around the same time, Banik, O’Ryan & Zhao (2018, hereafter BRZ18) considered a wider range of orbital geometries using a less computationally intensive restricted  $N$ -body approach in which the MW and M31 were treated as point masses

surrounded by test particle discs.<sup>2</sup> Their section 2 demonstrated that the MW–M31 trajectory is consistent with negligible peculiar velocity in the early universe, a constraint known as the timing argument (Kahn & Woltjer 1959). Despite lacking hydrodynamics or disc self-gravity, the initial setup of each galaxy as a rotating disc was sufficient to cause significant clustering of the tidal debris orbital poles. In some models, the preferred directions aligned with the actually observed orbital poles of the LG SPs. Therefore, these structures could well have formed as TDGs that condensed out of tidal debris orbiting in the correct plane. Indeed, Bílek et al. (2021) conclude that all satellite galaxies in the LG SPs are TDGs based on the earlier simulations of Bílek et al. (2018).

TDGs are expected to be more resilient in MOND due to their enhanced self-gravity, as explored with earlier high-resolution hydrodynamical MOND simulations (Renaud, Famaey & Kroupa 2016). Besides helping them to survive, this would also explain the high observed  $\sigma_{\text{int}}$  of satellite galaxies around the MW (McGaugh & Wolf 2010) and M31 (McGaugh & Milgrom 2013a,b). However, the enhancement to the self-gravity would typically be less than provided by a CDM halo, leading to a greater degree of tidal susceptibility in MOND. This is more in line with the observed morphologies of LG satellites (McGaugh & Wolf 2010).

A key distinguishing characteristic of more recently formed TDGs is their high metallicity for their mass (Duc et al. 2014). However, this relies on a long process of metal enrichment in the disc of the progenitor galaxy. When considering a very ancient interaction, there might not have been enough time for such an enrichment (Recchi, Kroupa & Ploekinger 2015), especially as TDGs are expected to form out of material that was initially several disc scale lengths out (BRZ18). This could explain why the M31 satellites within and outside its SP have similar properties, including in terms of metallicity (Collins et al. 2015).

A past MW–M31 flyby also has interesting consequences for the rest of the LG. Due to the high MW–M31 relative velocity around the time of their flyby, they would likely have gravitationally slingshot several LG dwarfs out at high speed. As discussed further in Section 4.2.2, this could lead to the existence of LG dwarfs with an unusually high RV in a  $\Lambda$ CDM context, such as the dwarfs in the NGC 3109 association (Pawlowski & McGaugh 2014; Peebles 2017). These could be backsplash from the MW–M31 flyby, a scenario that was considered in detail by Banik & Zhao (2018c). Backsplash galaxies also exist in  $\Lambda$ CDM, but very rarely have properties resembling NGC 3109 (Banik et al. 2021).

In this contribution, we build on the earlier studies of Bílek et al. (2018) and BRZ18 by conducting 3D hydrodynamical MOND simulations of the flyby using PHANTOM OF RAMSES (POR; Lüghausen, Famaey & Kroupa 2015; Nagesh et al. 2021). Our main objective is to vary the MW–M31 orbital pole and pericentre distance to find if there are models where the tidal debris around each galaxy aligns with its observed SP. Achieving this simultaneously for both the MW and M31 is a highly non-trivial test of the past flyby scenario. We also check if their discs are preserved and end up with realistic properties.

In the following, we describe the initial conditions and setup of our simulations (Section 2). We then present our results and analyses regarding the MW–M31 trajectory (Section 3.1) and proper motion (PM; Section 3.2), the tidal debris (Section 3.3), and the MW and M31 disc remnants (Section 3.4). We discuss our results in Section 4

<sup>2</sup>This leads to a numerically more tractable axisymmetric potential.



and conclude in Section 5. Videos of the LG in our best-fitting simulation with frames every 10 Myr are publicly available.<sup>3</sup>

## 2 METHODS

### 2.1 Poisson equation

The simulations presented in this paper are conducted with POR, which solves the governing equation of QUMOND:

$$\underbrace{\alpha \rho_{\text{eff}}}_{\equiv \rho_{\text{PDM}} + \rho_b} = \nabla \cdot [\mathbf{g}_N \nu(g_N)], \quad (2)$$

where  $\mathbf{g}_N$  is the Newtonian gravity determined from the baryonic density  $\rho_b$  using standard techniques,  $\nu \equiv |\nu|$  for any vector  $\mathbf{v}$ , and  $\mathbf{g}$  is the true gravity. It is often helpful to think of what density distribution would lead to this  $\mathbf{g}$  under Newtonian gravity. The required effective density  $\rho_{\text{eff}} \equiv \rho_{\text{PDM}} + \rho_b$ , where  $\rho_{\text{PDM}}$  is the phantom dark matter (PDM) density which captures the MOND corrections. Equation (2) is derived from a non-relativistic Lagrangian (Milgrom 2010), so QUMOND obeys the usual conservation laws regarding energy and momentum. This is also true of an earlier version (Bekenstein & Milgrom 1984), though we do not consider it here as it is computationally less efficient due to a non-linear grid relaxation stage.

To solve equation (2), we must assume an interpolating function  $\nu$  between the Newtonian and Milgromian regimes. In spherical symmetry, this has the effect that  $\mathbf{g} = \nu \mathbf{g}_N$ , softening the transition between the Newtonian inverse square law and equation (1). In this work, we use the ‘simple’ form of the interpolating function (Famaey & Binney 2005):

$$\nu(g_N) = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{a_0}{g_N}}. \quad (3)$$

This is numerically rather similar to the function used by McGaugh et al. (2016) and Lelli et al. (2017) to fit galaxy RCs, but can be inverted analytically. Other reasons for using this function were discussed in section 7.1 of Banik & Zhao (2018d) in preference to functions with a sharper transition. Equation (3) is quite accurate for  $g_N \approx (0.1 - 10)a_0$  as relevant for the MW–M31 flyby problem, but Solar System constraints imply that a more rapid convergence to the Newtonian result is required for  $g_N \gg a_0$  (Hees et al. 2014, 2016). However, the precise nature of this convergence is not important if the relevant quantity is  $\mathbf{g}$  rather than merely its deviation from  $\mathbf{g}_N$ .

For the computation of  $\mathbf{g}_N$ , we use the standard boundary condition that the Newtonian potential at large distance is

$$\Phi_N = -\frac{GM}{r}, \quad (4)$$

where  $G$  is the Newtonian gravitational constant,  $M$  is the total mass within the simulation volume, and  $\mathbf{r}$  is the position relative to its barycentre. The boundary condition for the true potential  $\Phi$  will be discussed in Section 2.4 based on equations (11) and (14).

### 2.2 Treatment of the MW–M31 orbit

An important aspect of our simulations is choosing an appropriate initial position and velocity for the MW and M31 disc templates to be discussed in Section 2.3. We do this using a semi-analytical backwards integration (hereafter SAM) very similar to that used in

**Table 1.** Our adopted cosmological parameters (Planck Collaboration XVI 2014), with 0 subscripts denoting present values. We assume a standard flat background cosmology and neglect other components as we are not considering the very early universe.

Parameter	Value
$H_0$	$67.3 \text{ km s}^{-1} \text{ Mpc}^{-1}$
$\Omega_{m,0}$	0.315
$\Omega_{\Lambda,0}$	0.685

section 2 of BRZ18, to which we refer the reader for a detailed discussion of the time evolution of the MW–M31 separation  $\mathbf{d}(t)$ . Briefly, their separation  $d$  in physical coordinates (used throughout this paper) is governed by

$$\ddot{d} = \frac{\ddot{a}}{a}d - g + \frac{h^2}{d^3}, \quad (5)$$

where  $a$  is the cosmic scale factor,  $\mathbf{h}$  is the angular momentum with magnitude  $h$  and direction  $\hat{\mathbf{h}}$  (the orbital pole),  $g$  is the radially inward component of the mutual gravity between the MW and M31, and an overdot denotes a time derivative. The term involving  $a$  is present in a homogeneously expanding universe, but  $g$  and  $h$  are zero in this case. As discussed in section 3.1.1 of Haslbauer et al. (2020), we adopt a standard Planck cosmology (Planck Collaboration XVI 2014) at the background level (Table 1). MOND can in principle explain the high locally measured  $H_0$  (e.g. Di Valentino 2021; Riess et al. 2021a) as arising from outwards peculiar velocities induced by the observed KBC void (Keenan et al. 2013), so our choice of cosmological parameters should be consistent with both early and late Universe probes of the expansion rate.

The angular momentum barrier  $h^2/d^3$  is necessary to prevent an unrealistic direct collision between the MW and M31. However, SAM is a timing argument analysis constrained to give zero peculiar velocity when  $a = 0.1$ , which implies  $h = 0$  then. We square this circle by assuming  $h = 0$  prior to first apocentre, after which  $\mathbf{h}$  instantaneously jumps to a particular value that remains fixed until today. The discontinuous behaviour of  $\mathbf{h}$  occurs at a time when the angular momentum barrier is least important to the trajectory, minimizing numerical effects. Physically, it would be reasonable if  $\mathbf{h}$  was mostly gained from tidal torques around the time of apocentre, but the more recent apocentre would be less relevant due to cosmic expansion driving external perturbers much further from the LG (Section 3.2.1). Importantly, our primary objective in this contribution is to conduct POR simulations of the flyby. These are initialized 1 Gyr before the flyby, which is safely after the jump in  $\mathbf{h}$ .

The calculation of  $g$  is rather complicated – we summarize only the main points here, and refer the reader to BRZ18 for a detailed discussion. In the absence of any other bodies and assuming the deep-MOND limit (gravitational fields  $\ll a_0$ ), we get a mutual gravity of

$$g_{\text{iso}} = \frac{Q\sqrt{GMa_0}}{d}, \quad (6)$$

$$Q = \frac{2\left(1 - q_{\text{MW}}^{\frac{3}{2}} - q_{\text{M31}}^{\frac{3}{2}}\right)}{3q_{\text{MW}}q_{\text{M31}}}, \quad (7)$$

where  $M$  is the total mass of the MW and M31, of which a fraction  $q_i$  resides in galaxy  $i$ . The parameter  $Q$  accounts for the finite mass ratio between the MW and M31 (Zhao, Li & Bienaymé 2010). Roughly speaking, it is caused by the PDM halo of one galaxy being reduced

<sup>3</sup><https://seafire.unistra.fr/d/6bb8e94212764324868e/>

in mass due to gravity from the other galaxy, which is a manifestation of the EFE. We assume that  $q_{\text{MW}} = 0.3$  and  $q_{\text{M31}} = 0.7$  (consistently with BRZ18), so  $Q = 0.7937$ .

When the MW and M31 are near pericentre, it is less accurate to assume the deep-MOND limit. However, the higher  $g$  allows us to neglect the relatively much weaker EFE on the whole LG, so we consider it as an isolated two-body problem where we get  $g$  numerically. Such Newtonian corrections are expected to have only a minor impact on our results because only a small portion of the SAM trajectory is subject to them.

When the MW and M31 are near apocentre, we expect the EFE from large-scale structure to be important (Section 2.4.3). In the limit where  $g_{\text{ext}} \gg g_{\text{iso}}$ , we can use the external field (EF)-dominated analytical solution found by Banik & Zhao (2018a) to get that

$$g_{\text{EFE}} = \frac{GMa_0}{d^2 g_{\text{ext}}} \left( \frac{3 + \cos^2 \theta}{4} \right), \quad (8)$$

where  $\theta$  is the angle between  $\mathbf{d}$  and the EF  $\mathbf{g}_{\text{ext}}$ . In reality,  $d$  is never large enough for  $g_{\text{ext}}$  to dominate, requiring an interpolation between the isolated and EF-dominated regimes. We achieve this using equation (14) of BRZ18, which is a fit to numerical results for the case  $\theta = 0$ .

In addition to considering a uniform  $\mathbf{g}_{\text{ext}}$ , we also follow the BRZ18 approach to include the tidal effect of M81, IC 342, and Cen A. Since in each case their gravity on the LG is  $\ll g_{\text{ext}}$  from more distant structures (see their section 2.3.1), we can superpose the gravitational field of each perturber on the LG.<sup>4</sup>

For simplicity, we assume that  $\mathbf{h}$  is constant in SAM except for the above-mentioned discontinuity at first turnaround, which occurs well before our POR simulation starts. Our POR calculations would miss changes in  $\mathbf{h}$  due to tidal torques from perturbers beyond the LG, but we will show later that such effects are quite small (Section 3.2.1). We therefore expect any changes in  $\mathbf{h}$  to arise mostly from the EFE (Section 3.2.2) and from torques on the MW–M31 orbit during the flyby. Neither effect is included in SAM, but both are included in POR. Since the MW–M31 orbit is nearly radial, changes in  $\mathbf{h}$  should have little effect on the timing argument, which is the main purpose of SAM.

In summary, we will hereafter use the present MW–M31 separation, direction, and mutual RV as present-day constraints on the SAM. We vary their orbital pole and mutual two-body angular momentum magnitude to match the observed SP orbital poles. Constraints on the PM of M31 will not be taken into account in this process.

### 2.3 Disc templates

SAM can only tell us the initial position and velocity of the MW and M31 discs. We therefore complement SAM with a system to generate a Milgromian disc template, to which we then apply the appropriate rotation and Galilean transformation. We generate two stable isolated Milgromian discs using the procedures described in Banik et al. (2020), i.e. by using our adapted version of the Newtonian code DISK INITIAL CONDITIONS ENVIRONMENT (DICE; Perret et al. 2014). The MW and M31 are assumed to have exponential surface density profiles, motivated by the fact that disc galaxies usually have an exponential radial profile (Freeman 1970), which arises naturally with about the right mass–size relation when spherical

gas clouds collapse under MOND gravity (Wittenburg, Kroupa & Famaey 2020). We implicitly assume that at the start time of our simulations  $\approx 8$  Gyr ago (redshift  $z \approx 1$ ), the MW and M31 discs had already formed with nearly their present masses. Thin rotationally supported disc galaxies do exist at even higher redshift (Lelli et al. 2018, 2021; Neeleman et al. 2020; Rizzo et al. 2020). The relatively isolated nature of the LG (e.g. Banik et al. 2021) suggests that its major galaxies might well have attained nearly their present mass rather early in cosmic history, especially in a framework with enhanced long-range gravity (Peebles & Nusser 2010) where mergers are less common due to the lack of dynamical friction between extended CDM haloes (Kroupa 2015; Renaud et al. 2016). This is quite plausible given the difficulty faced by the  $\Lambda$ CDM paradigm in explaining the high observed fraction of thin disc galaxies, which could be due to mergers being too frequent in this framework (Peebles 2020; Haslbauer et al. 2022).

The orientation, barycentre position, and velocity of each disc are set to the desired initial values by applying a rotation and Galilean transformation to its particles using an adapted version of the RAMSES patch known as `condinit`. This also assigns the density and velocity of each gas cell (Teyssier, Chapon & Bournaud 2010). To avoid severe thermal effects when the MW and M31 encounter each other, we use the same gas temperature of  $T2_{\text{ISM}} = 4.65 \times 10^5$  K (465 kK) for both galaxies. We set a temperature floor of  $T2_{\text{star}} = 0.8 T2_{\text{ISM}}$  and disable star formation and metallicity-dependent cooling, since at this exploratory stage we are mainly trying to reproduce the observed orientations of the LG SPs. If a suitable encounter geometry can be found, then it would be worthwhile to conduct a more detailed simulation with realistic star formation and stellar feedback prescriptions. However, this is beyond the scope of this work.

Unlike the POR simulations of M33 in Banik et al. (2020), an important aspect of the present POR simulations is that we need to consider the outer parts of the simulated discs in much greater detail because we expect these regions to be the original source material for the SPs. Indeed, the restricted  $N$ -body models of BRZ18 showed that the SPs mainly consist of material at an initial galactocentric distance of  $\approx 50$  kpc (see their figs 6 and 7). Material at such a large distance would be very poorly resolved with a computationally feasible number of equal mass particles. Therefore, we devise a procedure to vary the stellar particle mass in our disc templates so as to maintain a good resolution in the outer parts (see Appendix A). Near the disc centre, the particle mass is approximately constant at  $6 \times 10^5 M_{\odot}$  ( $1.4 \times 10^6 M_{\odot}$ ) for the MW (M31). Each disc template consists of  $5 \times 10^5$  particles. The spatial resolution of the POR gravity solver also needs to be sufficient – this is discussed further in Section 2.6. The resolution is 1.5 kpc in the best-resolved regions, though we show that improving this to 0.75 kpc has little effect on our results (Appendix E).

#### 2.3.1 Initial disc parameters

The rotation curve of each disc template is calculated using MOND gravity with the present value of  $a_0$ , since throughout this work we assume that  $a_0$  remains constant with time. Therefore, the initial MW and M31 discs would lie on the RAR defined by nearby galaxies. Possible consequences of a time-varying  $a_0$  were discussed in Milgrom (2015) and section 5.2.3 of Haslbauer et al. (2020), though we do not consider this here for simplicity. Constraining the evolution is challenging observationally because, amongst other issues, it is not yet possible to use 21-cm observations of neutral

<sup>4</sup>MOND becomes a linear gravity theory if the EFE dominates (Banik & Zhao 2018a).

**Table 2.** Parameters of the MW and M31 discs, each consisting of  $5 \times 10^5$  particles. Both galaxies have an initial gas temperature of 465 kK and a gas fraction of 0.5, which in the MW case is achieved by converting all of its more extended component and part of its less extended component into gas. Thus, the Galactic gas disc has a double exponential profile, while the stellar disc is a single exponential. We adopt an outer limit of 100 kpc for the M31 discs and for the more extended MW component, while 40 kpc is used for its less extended component. This is  $\approx 25$  scale lengths in all cases. The initial spin vector of each disc is given in Galactic coordinates ( $l$ ,  $b$ ).

Galaxy and component	MW inner	MW outer	M31 stars	M31 gas
Total mass	$9.15 \times 10^{10} M_{\odot}$		$2.135 \times 10^{11} M_{\odot}$	
Fraction of mass	0.8236	0.1764	0.5	0.5
Gas fraction	0.3929	1	0	1
Scale length (kpc)	1.29	4.2	4.24	4.24
Aspect ratio	0.15	0.0461	0.15	0.15
Disc spin vector	(55.11°, −84.89°)		(238.38°, −33.71°)	

hydrogen to obtain rotation curves in the distant Universe, so the H $\alpha$  line is typically used instead. The high- $z$  rotation curve data of Genzel et al. (2017) are quite consistent with a time-independent  $a_0$ , which sets some limits on its evolution (Milgrom 2017). Other works also suggest that high- $z$  galaxies are consistent with MOND (Stott et al. 2016; Harrison et al. 2017; Genzel et al. 2020; Sharma et al. 2021). Tighter constraints on this issue would be valuable to better understand the possible theoretical underpinnings of MOND, and more generally to test its prediction that isolated galaxies in dynamical equilibrium at any fixed redshift lie on a tight RAR.

We assume that the present MW and M31 disc scale lengths are as given in table 3 of BRZ18. When setting up their discs at the start of our simulation, we scale their present lengths by 0.8 for M31 and 0.6 for the MW. Starting with smaller discs is required to allow them to expand after the flyby to reach a realistic present-day configuration, and is also in line with the observed expansion of the stellar component of galaxies over cosmic time (Sharma et al. 2021). We use the DICE setting  $Q_{lim} = 1.25$  to ensure that all disc components have an initial Toomre parameter  $Q \geq 1.25$ , with the MOND generalization of the Toomre condition (Toomre 1964) discussed further in section 2.2. of Banik et al. (2020). The initial parameters of each disc are summarized in Table 2.

Gas dissipation in the tidal tails is likely important to obtaining thin SPs. While we cannot include this process as rigorously as we would like due to numerical limitations, it is certainly not appropriate to assume that the pre-flyby MW and M31 had a similar gas fraction to what we observe. Both discs are assigned an initial gas fraction of 0.5 because the flyby was  $\approx 7$ –9 Gyr ago (Zhao et al. 2013), when the gas fractions were likely much higher than today (Stott et al. 2016). Chemical evolution modelling of the MW indicates a gas fraction of  $\approx 0.5$  at that time (fig. 8 of Snaith et al. 2015). For simplicity, we use the same gas fraction for M31.

The stellar and gas discs of M31 are assumed to have the same scale length of 4.24 kpc, so we request only one disc component in the M31 namelist for the hydrodynamical version of DICE. For the MW, a double exponential profile is assumed, so two components are defined at that stage. The outer (more extended) component corresponds to its gas disc today, so we assume this was entirely gas at the start of our simulation. However, it only comprises 17.64 per cent of the total MW mass (Banik & Zhao 2018b), so we also convert a substantial fraction of the inner (less extended) MW component into gas to get a total gas fraction of 0.5.

The aspect ratios of the M31 disc and the inner MW disc are set to 0.15, so the inner MW component has a vertical density profile with a characteristic sech<sup>2</sup> scale height of 193.5 pc. The outer MW disc component is set to an aspect ratio of 0.0461 so that both components have the same scale height in pc. For both galaxies, the radial run of the gas disc scale height is then found by DICE to ensure it is as close to equilibrium as practicable following section 2.3 of Banik et al. (2020). They also described how particles in the DICE template are written out with a reduced mass or not at all, with the removed mass put back in as gas through the `condinit` routine in POR to ensure the correct gas fraction. Therefore, our hydrodynamical MOND version of DICE does not yield an equilibrium disc template by itself – it must be carefully combined with a modified version of POR. All the algorithms we use are publicly available<sup>5</sup> for reproducibility, with an accompanying user manual (Nagesh et al. 2021).

Since it is not possible to guarantee that our disc templates are exactly in equilibrium initially, we start our simulation 1 Gyr before the expected time of the flyby that we found semi-analytically in Section 2.2. The initial thickness profiles of the MW and M31 gas discs (shown in Fig. B1) are similar to that used by Banik et al. (2020) in their 100 kK model of M33.

We expect the MW and M31 discs to precess slightly from their initial orientations (see section 3.1 of BRZ18). Therefore, the observed orientations of the MW and M31 discs differ slightly from our adopted initial orientations, which are given in *final* Galactic coordinates ( $l$ ,  $b$ ) in Table 2. We use this system throughout this article when specifying directions of vectors – it is the system used in the simulation. To iteratively correct for disc precession, we find the rotation matrix between the final simulated and observed orientation of each disc, and then apply the inverse of this rotation to the observed orientation to initialize the next simulation. We will see later that the final simulated orientation of each disc agrees quite closely with observations after just one such iterative correction (Section 3.4).

## 2.4 Adding features to POR

The SAM procedure discussed in Section 2.2 is very accurate for following the overall behaviour of  $d(t)$ , but insufficient to model tidal debris generated by the interaction. This is the main purpose of the POR simulations we will conduct in this paper. There are some slight differences between the physics considered in SAM and in POR, which we try to rectify by adding features to POR and adjusting the initial conditions.

### 2.4.1 An allowance for tides

Tides from objects outside the LG are not directly included in the POR simulations, but are considered in SAM to estimate the flyby time as accurately as possible. To approximately include tides in POR, we estimate the amount of energy gained by the MW–M31 system due to tidal compression in the 1 Gyr preceding the flyby, which we estimate directly from SAM using information on the forces caused by each perturber. This energy is put into the radial component of  $\dot{\mathbf{d}}$  at the start of our POR simulation, which has the effect of slightly increasing how quickly the MW and M31 are approaching each other at that time. Our POR models neglect the impact of tides after the flyby, which is justified as the perturbers are much further apart then due to cosmic expansion.

<sup>5</sup><https://bitbucket.org/SrikanthTN/bonnpor/src/master/>



### 2.4.2 Dark energy

The present MW–M31 separation of 783 kpc (McConnachie 2012) is large enough that our POR model should include the cosmological term in equation (5). This partly consists of a decelerating term due to matter, which is included automatically because the LG mass mainly resides in the MW and M31, which we directly include. At late times, there is also an outwards repulsion from dark energy. We include this while operating RAMSES in non-cosmological mode, since this is required by the POR patch. For some dark energy parameter  $\Omega_{\Lambda,0}$ , the idea is to create an extra repulsive force

$$\Delta \mathbf{g} = H_0^2 \Omega_{\Lambda,0} \mathbf{r}, \quad (9)$$

where  $\mathbf{r}$  is the position relative to the barycentre.

We can reproduce equation (9) with a standard Poisson solver if we adjust the density and boundary condition. Since we want  $\mathbf{g}_N$  to be calculated in a standard way so that it is correctly MONDified, we apply the density increment only to the PDM:

$$\Delta \rho_{\text{PDM}} = -\frac{3H_0^2 \Omega_{\Lambda,0}}{4\pi G}. \quad (10)$$

For consistency, we also adjust the boundary potential for only the MOND stage by

$$\Delta \Phi = -\frac{H_0^2 \Omega_{\Lambda,0}}{2}. \quad (11)$$

Equations (10) and (11) are implemented by appropriate adjustments to the POR algorithm, thereby yielding equation (9) in the interior.

The impact of dark energy on the MW–M31 dynamics is quite small in MOND as their mutual gravity is  $\approx 30\times$  stronger than the cosmological acceleration term in equation (5) (table 10 of BRZ18). We none the less include it for completeness. The fact that the cosmological acceleration is small compared to the internal gravity means that our results are robust with respect to uncertainty regarding how MOND should be applied in a cosmological context (the ‘Hubble field effect’ discussed in section 5.2.3 of Haslbauer et al. 2020). It is however possible to make some plausible assumptions and simulate systems where the average enclosed density differs only slightly from the cosmic mean, as done in that work and in several others (e.g. Katz et al. 2013; Candlish 2016).

### 2.4.3 The external field effect (EFE)

As discussed in section 2 of BRZ18, the EFE from large-scale structure has a significant effect on the LG gravitational field when the MW and M31 are close to apocentre, where they spend a significant amount of time. The EFE is a non-standard phenomenon caused by the non-linearity of MOND (equation 1). It leads to the internal gravitational dynamics of a system being affected by  $\mathbf{g}_{\text{ext}}$  even in the absence of tidal effects (Milgrom 1986). Strong evidence for the EFE in field galaxies was recently reported by Chae et al. (2020b, 2021) by comparing the RCs of galaxies in isolated and more crowded environments (building on similar earlier work; Haghi et al. 2016; Hees et al. 2016).

In the QUMOND approach where we must first get  $\mathbf{g}_N$ , the main change is to add the Newtonian-equivalent external field to the Newtonian gravity sourced by the system under study.

$$\mathbf{g}_N \rightarrow \mathbf{g}_N + \mathbf{g}_{N,\text{ext}}. \quad (12)$$

Assuming that  $\mathbf{g}_{\text{ext}}$  is sourced by a distant point-like object, we get that

$$\mathbf{g}_{N,\text{ext}} \overset{v_{\text{ext}}}{\propto} \left( \mathbf{g}_{N,\text{ext}} \right) = \mathbf{g}_{\text{ext}}. \quad (13)$$

If we know the EF  $\mathbf{g}_{\text{ext}}$ , this can be inverted to obtain  $\mathbf{g}_{N,\text{ext}}$ .

Once we have adjusted  $\mathbf{g}_N$  according to equation (12), we use it to find  $\nabla \cdot \mathbf{g}$  using equation (2) as before. The reason is that this applies to the Newtonian gravity sourced by matter both within and beyond the simulated domain, the extent of which is an arbitrary decision that should have no bearing on the result.

Including the EFE also requires us to change the boundary condition, but only for the MOND stage since the internal dynamics and external field are fully separable in Newtonian gravity. For simplicity, the boundary should be in some asymptotic regime where  $\mathbf{g}$  has a well-understood analytical behaviour. Normally, it is sufficient for the boundary to be distant enough that the simulated system can be approximated as a point mass. If there is also a non-negligible EFE, then the simplest option is to choose a boundary where  $\mathbf{g}_{\text{ext}}$  is much stronger than the internal gravity of the system. Its internal potential then becomes EF-dominated (e.g. Banik & Zhao 2018a):

$$\Phi = -\frac{GMv_{\text{ext}}}{r} \left( 1 + \frac{K_0}{2} \sin^2 \theta \right), \quad (14)$$

$$K_0 \equiv \frac{\partial \ln v_{\text{ext}}}{\partial \ln g_{N,\text{ext}}}, \quad (15)$$

where  $M$  is the total mass within the simulation volume,  $\mathbf{r}$  is the position relative to its barycentre, and  $\theta$  is the angle between  $\mathbf{r}$  and  $\mathbf{g}_{\text{ext}}$ . Due to the  $1/r$  dependence and the fact that potentials from different sources can be superposed in this perturbative framework, the result is reminiscent of standard Newtonian mechanics, so the EF-dominated regime is also known as the quasi-Newtonian regime. Note that equation (14) alone is not our boundary condition because we also include a dark energy adjustment (equation 11).

For the Newtonian stage of solving the QUMOND Poisson equation, we continue to use a boundary potential of  $-GM/r$ , ignoring the constant  $\mathbf{g}_{\text{ext}}$ . This means that our simulations consider the internal dynamics in a freely falling reference frame accelerating at  $\mathbf{g}_{\text{ext}}$ , whose relevance to the internal dynamics is a violation of the strong equivalence principle.

For simplicity, we keep  $\mathbf{g}_{\text{ext}}$  fixed over the course of our POR simulation. However, to maximize the accuracy of the overall MW–M31 trajectory, SAM uses a time-dependent  $\mathbf{g}_{\text{ext}}$  as described in section 2.2 of BRZ18. In both cases, we follow the approach in that paper of assuming that today,  $\mathbf{g}_{\text{ext}} = 0.022 a_0$  directed towards Galactic coordinates  $(276^\circ, -30^\circ)$ , the direction in which the LG presently moves relative to the CMB (Kogut et al. 1993). This direction is assumed fixed throughout cosmic history.

## 2.5 Iterative orbit adjustment

Despite our best attempts to ensure SAM and POR handle the flyby problem as similarly as possible, the two algorithms none the less use very different techniques. Thus, advancing the SAM-generated initial conditions using POR does not quite yield the presently observed  $\hat{\mathbf{d}}$  or its direction  $\hat{\mathbf{d}}$ . In the presence of an EFE, the late-time  $\hat{\mathbf{d}}$  affects the torque exerted by the EFE on the MW–M31 system (Section 3.2.2), which in turn influences the present PM of M31. Moreover, an incorrect final  $\hat{\mathbf{d}}$  suggests that the flyby took place in a different orientation to how it occurred in the simulation, which would influence the SPs.

**Table 3.** Present-day constraints imposed on SAM, which does a backwards integration. The tangential velocity of M31 is varied to best match the observed SP orbital poles.

Initial MW–M31 ...	Value
distance	783 kpc
direction	(121.17°, −21.57°)
RV	−93.4 km s <sup>−1</sup>

To ensure the final  $\hat{\mathbf{d}}$  matches the observed sky position of M31 as accurately as possible, we find the rotation matrix that takes the observed  $\hat{\mathbf{d}}$  to the simulated final value. We then rerun SAM with the inverse of this rotation applied to the presently observed  $\hat{\mathbf{d}}$  and the relative velocity  $\dot{\mathbf{d}}$  used in the previous SAM simulation. The idea is that if  $\hat{\mathbf{d}}$  ends up 10° further north than observed, then we can get approximately the correct final  $\hat{\mathbf{d}}$  by running SAM with a present  $\dot{\mathbf{d}}$  that lies 10° south of the actually observed direction towards M31, because there is some additional physical effect in POR but not in SAM that pushes  $\hat{\mathbf{d}}$  further north by 10°. The initial conditions generated in this way are used to rerun the POR simulation. We find that just one such iteration allows us to match the observed  $\hat{\mathbf{d}}$  to within a few degrees, which we consider sufficient.

## 2.6 Simulation setup and initial conditions

### 2.6.1 SAM

We begin by running SAM with the constraints given in Table 3. The main model parameters that we vary are:

- (i) the MW–M31 orbital pole  $\hat{\mathbf{h}}$ , and
- (ii) the magnitude  $h$  of their mutual two-body angular momentum.

As discussed in section 5.1.3 of BRZ18, we account for reflex motion induced by the Large Magellanic Cloud (LMC) on the MW and by M33 on M31, leading to a present RV of  $\dot{d} = -93.4$  km s<sup>−1</sup>.<sup>6</sup> We keep  $\dot{d}$  fixed, but vary the present tangential velocity  $v_{\text{tan}}$  used in SAM, which does a backwards integration. Changing  $h$  mainly affects the perigalacticon distance, but does not much alter the relative speed then. Varying  $\hat{\mathbf{h}}$  allows us to consider a wide range of possible orbital geometries. While our priority is to match the SP orientations, we subsequently compare our best-fitting model with the observed PM of M31 (Section 3.2). Interestingly, if we neglect both considerations, the timing argument alone sets some constraints on  $\hat{\mathbf{h}}$  because perturbers like Cen A have a different influence on the internal dynamics of the LG depending on  $\hat{\mathbf{d}}$  prior to the flyby. As a result, the timing argument mass of the LG is too high for a large range of  $\hat{\mathbf{h}}$ , which renders the models unlikely as the MOND timing argument mass should correspond to the baryonic mass (section 5.1.3 of BRZ18).

The model parameters which we explore are the same as in BRZ18, except that we do not vary the EF. It was indeed shown that varying the EF within the plausible range has only a small impact on the final results, which is likely due to the well-known LG

velocity in the CMB frame setting some constraint on the history of  $\mathbf{g}_{\text{ext}}$ . We therefore use the best-fitting model of BRZ18 in which  $\mathbf{g}_{\text{ext}} = 0.022 a_0$  directed towards Galactic coordinates (276°, −30°) at the present time. To make the MW–M31 trajectory as realistic as possible, the  $\mathbf{g}_{\text{ext}}$  assumed in SAM varies with time, as discussed in their section 2.2. However, implementing a time-varying  $\mathbf{g}_{\text{ext}}$  in POR would involve significant complications, so this uses the present  $\mathbf{g}_{\text{ext}}$  for the full duration of the simulation. The EFE is not expected to influence details of the MW–M31 interaction, but causes them to reach a larger apocentre at late times by weakening their mutual gravity. As a result, the main role of the EFE is to alter the timing of the MW–M31 flyby, which sets the time available for the tidal debris to settle down after the flyby. We therefore prioritize making the flyby time as accurate as possible in our POR models, which requires a carefully prepared SAM. Due to these slight differences and the more rigorous treatment of the flyby in POR, the MW–M31 trajectory is expected to differ somewhat compared to SAM. We do not run the POR simulations for a different length of time than that indicated by SAM in order to obtain a better agreement with the observed distance to M31, so this is expected to differ somewhat from the observed  $783 \pm 25$  kpc (McConnachie 2012).

It is important to realize that although the final M31 distance in POR need not match observations exactly, requiring even an approximate match places non-trivial constraints on our model. This is because a very close encounter would lead to significant dynamical friction between the baryonic discs, causing a very low apocentre and a subsequent merger within a few Gyr. This possibility was neglected in the models of BRZ18, where the  $d(t)$  returned by SAM was assumed to be exactly correct. In general, a hydrodynamical model of the interaction is obviously a significant advance on the previous restricted  $N$ -body models, even if the much higher computational cost reduces the scope for fine-tuning to match certain observables to high precision. We tried  $\approx 30$  POR models to obtain a good fit to the SP orientations, which we judged visually.<sup>7</sup>

### 2.6.2 POR

Our POR simulations use very similar settings to those described in section 2.4 of Banik et al. (2020), so we briefly mention the main points here. We use the non-cosmological particle-in-cell mode, activate MOND and the EFE, and use the refinement conditions  $m_{\text{refine}} = 10^3 M_{\odot}$  and  $n_{\text{subcycle}} = (1, 1, 2, 2)$ . Since we are simulating the whole LG, we use a much larger box size of 6144 kpc, with 7–12 levels of refinement. The most poorly resolved regions thus have a resolution of  $6144/2^7 = 48$  kpc, which improves to  $6144/2^{12} = 1.5$  kpc for the best resolved regions. Teyssier (2002) provides a more detailed description of RAMSES, including default values of parameters that we do not alter. Perhaps the most important of these is the `gravity_type`, whose default setting of 0 represents self-gravity. We also use the default Poisson convergence parameter  $\epsilon_{\text{ps}} = 10^{-4}$ .

The results of our best-fitting model are discussed next. In this model, the initial position and velocity of the MW and M31 are as given in Table 4 for the centres of two galaxies each consisting of  $5 \times 10^5$  particles and a gas fraction of 50 percent, with the disc orientations given in Table 2. The initial gas temperature is  $T2_{\text{ISM}} = 465$  kK, with the temperature floor  $T2_{\text{star}}$  being

<sup>6</sup>Although the LMC should form out of tidal debris expelled during the flyby, our model does not form individual TDGs. The simulated Galactic disc thus does not experience recoil from an orbiting massive satellite, as it might do in a more advanced model. Since the timing argument is sensitive to the gravity between the MW and M31, any massive satellites should be included in the mass and velocity of each galaxy.

<sup>7</sup>We also ran a few more models to fine-tune the agreement with some observables like the present disc orientations and the MW–M31 direction (Section 2.5).



**Table 4.** The initial position in kpc and velocity in  $\text{km s}^{-1}$  of the MW and M31 disc centres at the start of our best-fitting POR simulation. The combined barycentre is at the origin because the MW:M31 mass ratio is 3:7 (Table 2), consistently with BRZ18. SAM indicates that these initial conditions are valid 8.1 Gyr ago.

Galaxy Direction	MW		M31	
	Position	Velocity	Position	Velocity
$x$	−56.0	5.2	24.0	−2.2
$y$	254.2	−177.8	−109.0	76.2
$z$	44.1	−86.7	−18.9	37.2

20 per cent lower. According to SAM, the initial MW and M31 positions correspond to 1 Gyr before the flyby, so our POR simulations can be considered to start 5.72 Gyr after the big bang (8.1 Gyr ago). Starting the simulations 1 Gyr before the flyby gives the discs some time to settle down first, but avoids the need to simulate very early epochs at which the cosmological term in equation (5) would be significant.

### 3 RESULTS

In this section, we present the results of our best-fitting model and compare it with relevant observations. We use the simulation snapshot after 8.2 Gyr because this corresponds as closely as possible to the lookback time estimated by SAM for the initial conditions of Table 4.

#### 3.1 The MW–M31 orbit

The initial conditions are obtained by running SAM, which gives a timing argument mass of  $3.457 \times 10^{11} M_{\odot}$  for the whole LG. This is somewhat higher than the sum of the disc masses used in POR (Table 2), which we consider acceptable as there would also be some mass in, e.g. a halo of gas around each galaxy, though we do not include a halo explicitly. This issue was discussed further in section 5.1.1 of BRZ18, who argued that although the observed MW and M31 RCs suggest a combined disc mass of  $2.3 \times 10^{11} M_{\odot}$ , a modestly higher timing argument mass for the whole LG is quite feasible due also to satellite galaxies like the LMC and M33. Equation (1) implies that a 50 per cent increase in the mass of a galaxy increases the flatline level of its RC by just 11 per cent. Such a small increase in the MW and M31 RCs at large ( $\gtrsim 100$  kpc) distances is difficult to rule out at present, especially for M31 (Corbelli et al. 2010; Sofue 2015).

Before conducting more detailed analyses, we show the central 500 kpc of the POR simulation as viewed from the direction which would make both discs appear perfectly edge-on if their orientations are as observed – this is very nearly the case (Section 3.4). Fig. 1 shows the stellar particles, while Fig. 2 shows the gas. The discs undergo closest approach  $\approx 1$  Gyr into the simulation. Dynamical friction during the flyby is small, allowing the galaxies to reach a large post-encounter separation. Moreover, the flyby does not disrupt the MW and M31 discs too severely, allowing them to retain a thin disc by the end of our simulation – we will investigate this in more detail in Section 3.4. There is also a small numerical drift of the MW–M31 barycentre, which can be understood using much less computationally intensive methods (see footnote 14 of Banik et al. 2020).

We extract the MW–M31 trajectory from our POR simulation while it is running, without extracting every simulation output. We briefly

describe this technique in Appendix C as it could also be useful for other projects and is part of the publicly available version of POR used here. In this way, we obtain the trajectory shown in Fig. 3, with the top and bottom panel used to show the MW–M31 separation and relative velocity, respectively. The expected result using SAM is shown as a dotted red line in each panel. Both methods give a rather similar overall trajectory, indicating little dynamical friction during the encounter. This is due to the fairly large pericentre distance of 81 kpc. Table 5 summarizes information about the pericentre and apocentre in each trajectory. The higher second apocentre in SAM was discussed in section 5.1.2 of BRZ18, who concluded that it is mainly driven by tides from Cen A due to its relatively high mass and close alignment with the MW–M31 line after but not before the flyby. Tides are not explicitly included in our POR simulation, which moreover starts shortly before the flyby and thus includes only the most recent MW–M31 apocentre, when the perturbers are rather distant.

The simulation snapshot that we analyse (8.2 Gyr after the start) has  $d = 846$  kpc, which slightly exceeds the observed  $783 \pm 25$  kpc distance to M31 (McConnachie 2012). As explained previously, we do not rectify this issue by running the simulation for longer. Doing so should have only a small effect on the SPs, the main focus of this work. Importantly for the overall geometry, the simulated  $\hat{d} = (121.11^{\circ}, -21.99^{\circ})$ , which differs by only  $0.42^{\circ}$  from the observed sky position of M31 ( $121.17^{\circ}, -21.57^{\circ}$ ).<sup>8</sup> Therefore, the overall MW–M31 trajectory in our POR simulation is quite reasonable, and should be consistent with cosmological initial conditions (the timing argument) at much earlier times.

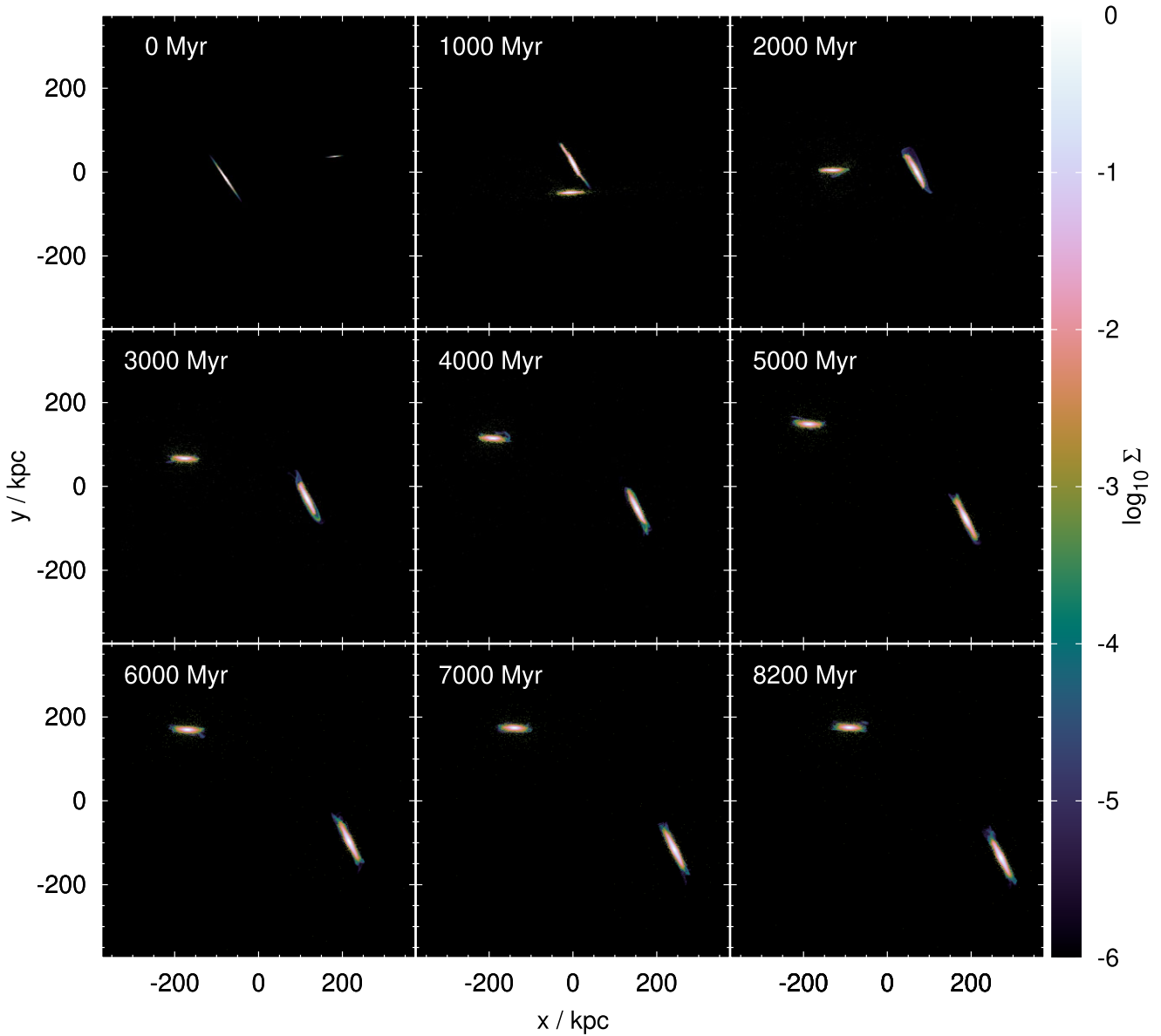
#### 3.2 Final tangential velocity and PM

The PM of M31 provides an important constraint on the orbital geometry of our best-fitting model. When scanning the parameter space (Section 2.6.1), we did not consider this constraint, though we did not consider all possible orbital geometries either as some are highly disfavoured by the timing argument alone (Section 2.2). In the following, we describe how we obtain the predicted PM of M31, and compare this to the latest observational constraints.

To find the separation and relative velocity of the MW and M31, we first have to identify each galaxy’s centre of mass. We obtain an initial guess using the iterative on-the-fly method described in Appendix C based on the particles alone. Much more detailed analyses are possible using a simulation snapshot because we also use our modified version of RDRAMSES (section 3 of Banik et al. 2020) to obtain a list of all gas cells, treating them as particles at the cell centres. This allows the stars and gas to be analysed on an equal footing. We therefore find the barycentre of all material whose position and velocity lies within 250 kpc and  $500 \text{ km s}^{-1}$ , respectively, of our initial guess for the barycentre. These thresholds are deliberately set quite wide to reduce the risk of converging on the wrong density maximum. This process is repeated iteratively until the barycentre position shifts by  $< 1$  kpc and its velocity shifts by  $< 1 \text{ km s}^{-1}$  between successive iterations. The process converges very rapidly because the initial guess for the barycentre based on particles alone (Appendix C) is already highly accurate.

Table 6 shows the position and velocity of the MW and M31 in the simulation reference frame, from which we get the MW–M31 relative separation and velocity. We take the cross product of these vectors

<sup>8</sup>The simulated  $\hat{d}$  is actually found using the method discussed in Section 3.2 to identify the MW and M31 barycentres, but the results are almost identical.

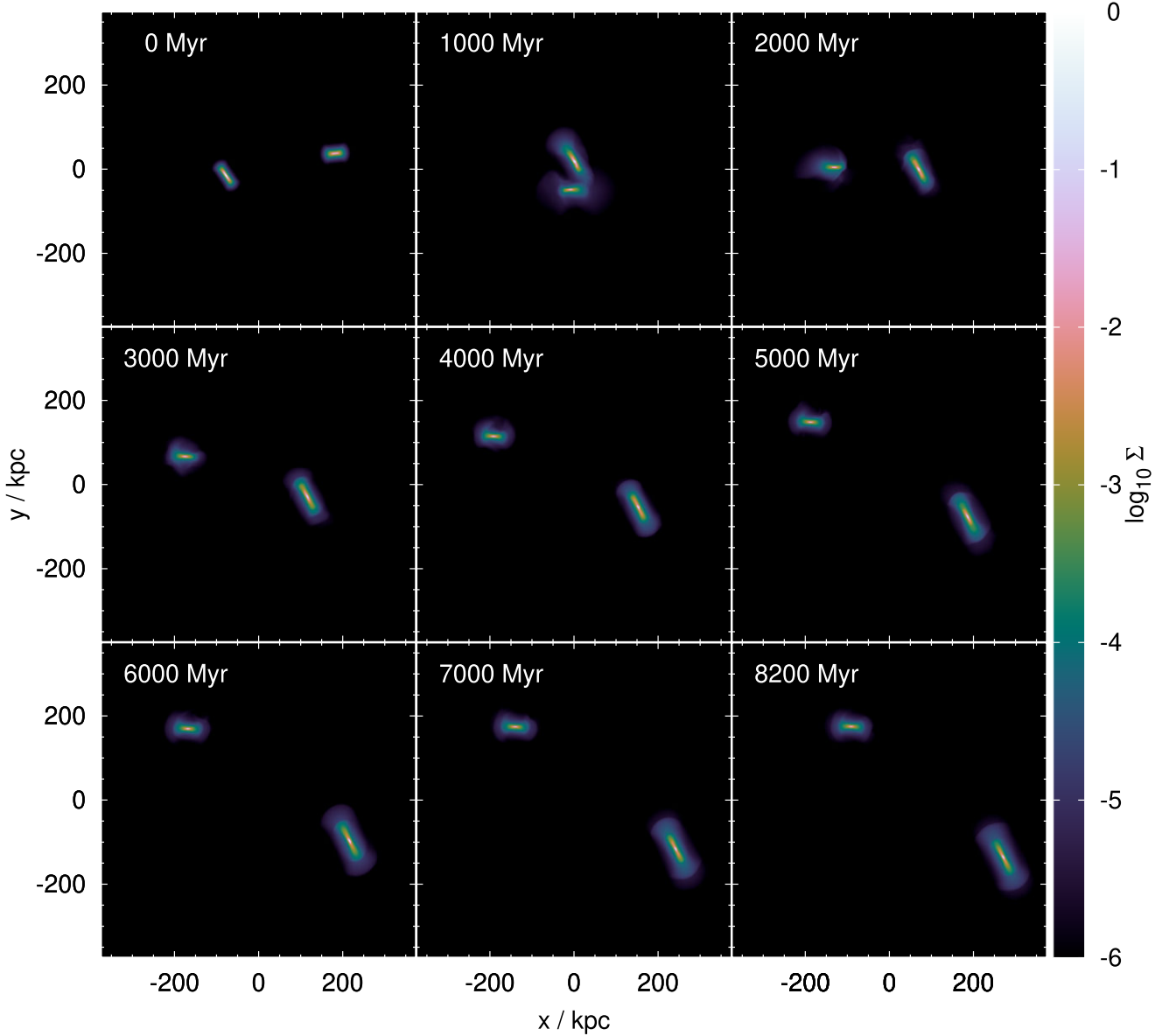


**Figure 1.** The stellar particles in the central 500 kpc square of our POR simulation, viewed along the direction which would make both discs appear edge-on in actual observations at  $z = 0$  (Section 3.4). Since the disc orientations change somewhat during the flyby and we have not perfectly adjusted for this, the discs are not perfectly edge-on in the last snapshots. The projected surface density is shown in units of  $1000 M_{\odot} \text{pc}^{-2}$ . The time since the start of the simulation is indicated at the top left of each panel. The MW disc appears almost horizontal in this view. Notice that the MW and M31 retain thin discs, outside of which there is very little material.

to get their orbital angular momentum  $\mathbf{h}$  and thus the orbital pole  $\hat{\mathbf{h}}$ . We also use  $\mathbf{h}$  to obtain the tangential velocity  $v_{t,783}$  at a distance of 783 kpc, which slightly exceeds the simulated tangential velocity by a factor of (846/783) because the final distance is slightly larger than observed. We then find what PM components of M31 on the night sky would most closely mimic  $\mathbf{d}$  and  $\mathbf{v}_{\text{tan}}$  in the analysed POR simulation output. We do this by adjusting the assumed M31 PM components ( $\mu_{\alpha,*}$ ,  $\mu_{\delta}$ ) in SAM to find the combination which best matches the  $\hat{\mathbf{h}} = (209.77^{\circ}, 3.29^{\circ})$  and  $v_{t,783} = 34.12 \text{ km s}^{-1}$  found from POR, with  $\mu_{\alpha,*}$  and  $\mu_{\delta}$  being the angular velocity in the directions which most quickly increase the right ascension and declination, respectively. In this calculation, the present M31 direction and RV in SAM are fixed to the values in Table 3. The Galactic circular velocity at the Solar circle is assumed to be  $239 \text{ km s}^{-1}$  (McMillan 2011),

while the non-circular motion of the Sun is taken from Francis & Anderson (2014) – uncertainties in these parameters and in the distance to M31 are much smaller than in its PM. We vary ( $\mu_{\alpha,*}$ ,  $\mu_{\delta}$ ) using a gradient descent algorithm (Fletcher & Powell 1963) to minimize the sum of squared errors in  $\hat{\mathbf{h}}$  and  $v_{t,783}$ , with each error scaled to an uncertainty of  $1^{\circ}$  and  $1 \text{ km s}^{-1}$ , respectively. While it is possible to match  $v_{t,783}$  exactly, the best fitting ( $\mu_{\alpha,*}$ ,  $\mu_{\delta}$ ) still gives a small error in  $\hat{\mathbf{h}}$  because of a slight difference in the final  $\hat{\mathbf{d}}$  between POR and SAM. This is reduced by an iterative rerun of the simulation. Since the mismatch is then only  $0.42^{\circ}$ , our approach gives a good idea of what our simulation implies for the present M31 PM.

Our result is shown in Fig. 4. As discussed in section 5.1.3 of BRZ18, this includes a correction for the LMC and M33 altering the barycentric velocity of the MW and M31, respectively. Though the



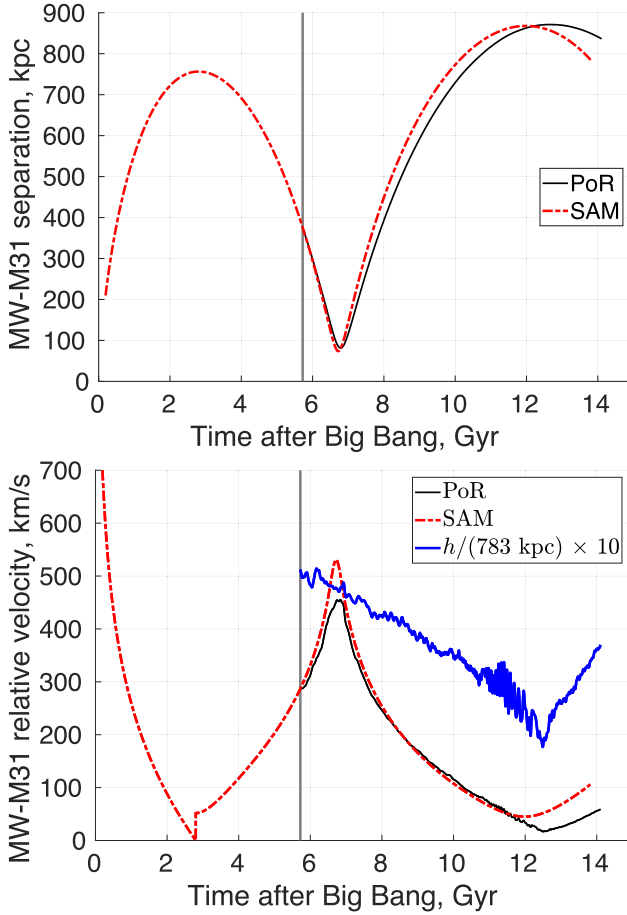
**Figure 2.** Similar to Fig. 1, but now showing the gas. The smallest simulated gas cell has sides of length 1.5 kpc.

LMC and M33 positions and velocities are known fairly well, there is some uncertainty regarding their mass. This affects the results slightly because our model does not form individual TDGs, so it does not properly capture the induced recoil on the MW from the LMC, which ideally should form in a simulation of the flyby. The uncertainties are rather small in a MOND context because all galaxies are purely baryonic and we know the position and velocity of M33 and the LMC rather well. We therefore vary the mass of each satellite by  $\pm 20$  per cent and repeat the above-mentioned PM calculation, which assumes the MW (M31) position and velocity in SAM refers to the barycentre of the MW–LMC (M31–M33) system. As discussed in section 4.4 of Banik & Zhao (2016), the PM correction due to the LMC is not as significant because its Galactocentric velocity is mostly away from M31. This is evident in the very short length of the dark blue line through the model PM in Fig. 4, which indicates the impact of a 20 per cent uncertainty in the LMC mass. The longer pink line shows the same for M33, whose effect is somewhat larger. Even so, the uncertainty on the M31 PM is  $< 1 \mu\text{as yr}^{-1}$  in

both cases, which is much smaller than the difference between its predicted PM and that for a purely radial MW–M31 orbit (open blue circle).

We compare the predicted M31 PM with the measurements in Table 7. For clarity, we only plot the observed values from the *Hubble Space Telescope* (HST) prior to *Gaia* (van der Marel et al. 2012) and the recent result from *Gaia* early data release 3 (eDR3; Salomon et al. 2021). In the latter case, we use their so-called  $B_{\text{pm}}$  sample in their table 4 to include corrections for instrumental effects as estimated from PMs of background quasars, but without including corrections for astrophysical motions beyond the Solar System – these are handled in our analysis (see e.g. section 2.3 of Banik & Zhao 2016). The combination of HST results with *Gaia* data release 2 (DR2; van der Marel et al. 2019) is not shown in Fig. 4 as it gives a similar but somewhat less precise result to the  $B_{\text{pm}}$  sample in *Gaia* eDR3. All three PM estimates are mutually consistent within uncertainties. They also agree very well with the PM of our POR model (Fig. 4). The PM uncertainties are now small enough that





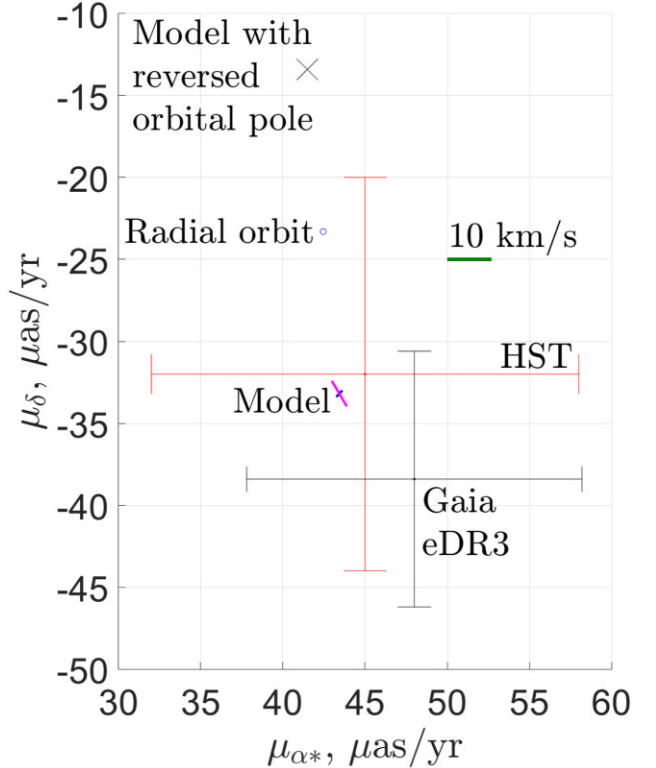
**Figure 3.** *Top:* The MW–M31 separation as a function of time in our PoR (solid black) and SAM (dot–dashed red) models. Although our PoR simulation is not cosmological, its initial conditions are obtained from SAM 1 Gyr before the flyby using a total mass of  $3.457 \times 10^{11} M_{\odot}$ , thus placing the PoR model in the cosmological context of a trajectory consistent with the timing argument (Section 2.2). The PoR model (with a slightly lower total mass as it considers only the discs) is advanced slightly into the future, but our detailed analyses focus on the snapshot 8.2 Gyr after the start of the simulation to most closely match the post-flyby time estimated by SAM. The vertical grey line in each panel shows the start of our PoR simulation. *Bottom:* The MW–M31 relative velocity in SAM (dot–dashed red) and PoR (solid black). The solid blue line with more scatter shows their relative angular momentum as an equivalent tangential velocity at their present separation of 783 kpc (McConnachie 2012), multiplied by 10 for clarity. The discontinuity in the SAM velocity around 3 Gyr arises from assuming no angular momentum prior to first apocentre, to allow the MW–M31 trajectory to reach zero peculiar velocity at very early times (Section 2.2). Notice the lack of a significant change in the angular momentum due to the MW–M31 flyby, indicating that it hardly alters their orbit. The gradual change over many Gyr is due to torque from the EFE (Section 3.2.2) and due to tidal debris falling back on to each galaxy, especially the MW (Section 3.3.3).

**Table 5.** Information about the MW–M31 pericentre and apocentre in our SAM and PoR simulations, based on Fig. 3. Times are relative to the big bang.

Orbital phase Algorithm	Perigalacticon		Apogalacticon	
	SAM	PoR	SAM	PoR
Time (Gyr)	6.72	6.79	12.0	12.6
Distance (kpc)	73.8	81.5	867.8	871.2
Velocity ( $\text{km s}^{-1}$ )	530.9	452.1	45.1	18.6

**Table 6.** The final position in kpc and velocity in  $\text{km s}^{-1}$  of the MW and M31 barycentres 8.2 Gyr into our best-fitting PoR simulation. According to SAM, this is the PoR simulation snapshot temporally closest to the present epoch. The combined barycentre is slightly offset from the origin due to numerical drift, as discussed further in footnote 14 of Banik et al. (2020).

Galaxy Direction	MW		M31	
	Position	Velocity	Position	Velocity
$x$	232.9	−31.0	−172.1	−6.6
$y$	−248.8	73.5	422.6	29.4
$z$	176.4	−0.7	−140.2	−14.4



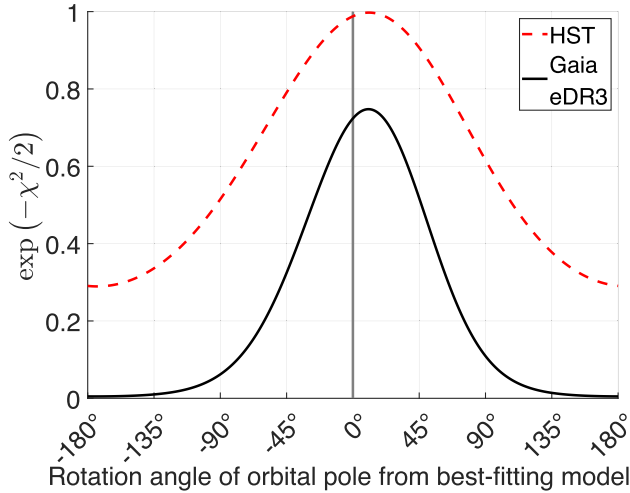
**Figure 4.** The heliocentric PM of M31 in our best-fitting PoR model (intersection of blue and pink lines), found using the method described in Section 3.2. The short blue (pink) line indicates the effect of varying the LMC (M33) mass by  $\pm 20$  per cent. In both cases, a heavier satellite increases the predicted M31  $\mu_{\alpha,*}$ . The open blue circle shows the M31 PM for a purely radial MW–M31 orbit, while the black cross shows the result of reversing the Galactocentric tangential velocity of M31 given by our PoR model. The large red error bars show the observed M31 PM using *HST* (van der Marel et al. 2012), while the black error bars use instead *Gaia* eDR3 (Salomon et al. 2021). The measurement of van der Marel et al. (2019) is not shown here for clarity, but all three PMs are listed in Table 7. The horizontal green line represents  $10 \text{ km s}^{-1}$  at the M31 distance of 783 kpc.

this is not guaranteed. We demonstrate this by showing the expected PM of M31 if the PoR-determined Galactocentric tangential velocity of M31 is reversed, which reverses  $\mathbf{h}$ . The result is the black cross towards the upper left of Fig. 4, which is now inconsistent with the latest determination of the M31 PM (Salomon et al. 2021).

To explore this further, we rotate the PoR-determined  $\hat{\mathbf{h}}$  around the M31 direction  $\hat{\mathbf{d}}$  by all possible angles between  $-180^\circ$  and  $+180^\circ$  in steps of  $1^\circ$ . We then determine the  $\chi^2$  statistic relative to the results of van der Marel et al. (2012) and Salomon et al. (2021). Since this  $\chi^2$

**Table 7.** The PM of M31 in  $\mu\text{as yr}^{-1}$  according to different investigations, which in order are: *HST* alone (van der Marel et al. 2012), *HST* and *Gaia* DR2 (van der Marel et al. 2019), and *Gaia* eDR3 (Salomon et al. 2021) using the  $B_{\text{pm}}$  sample in their table 4. The latter result is in a heliocentric frame, including instrumental corrections for PMs of background quasars and using an outlier rejection system. Astrophysical motions beyond the Solar System are not accounted for in these results, since we do the corrections ourselves. The final column gives the PM in our best-fitting model.

Component	<i>HST</i>	<i>HST</i> + <i>Gaia</i> DR2	<i>Gaia</i> eDR3	Model
$\mu_{\alpha,*}$	$45 \pm 13$	$49 \pm 11$	$48.0 \pm 10.2$	43.4
$\mu_{\delta}$	$-32 \pm 12$	$-38 \pm 11$	$-38.4 \pm 7.8$	-33.2



**Figure 5.** The probability of a higher  $\chi^2$  between the observed PM of M31 and in our POR model if the predicted Galactocentric tangential velocity of M31 is rotated around the simulated direction towards M31 by the angle shown on the x-axis. There is a clear preference for rotation angles close to zero, i.e. for an MW–M31 orbital pole similar to that in the model.

is based on two parameters (the PM components), the probability of a higher  $\chi^2$  arising by chance if the model were correct is  $\exp(-\chi^2/2)$ . Fig. 5 shows this probability for all possible rotation angles, with  $0^\circ$  corresponding to the actual  $\hat{h}$  of our POR model. It is clear that for fixed  $h$ , the direction  $\hat{h}$  preferred by the observed M31 PM (Table 7) corresponds quite closely to that of our best-fitting POR model.

At this point, it is worth emphasizing that we did not consider the PM of M31 when selecting the best-fitting POR simulation – this was based purely on the phase space distribution of the tidal debris. The agreement between the observed M31 PM and in this model constitutes a non-trivial success thereof.

### 3.2.1 Tidal torques

Our POR model neglects the late-time effect of tidal torques from perturbers outside the LG, which would also affect the PM of M31. To estimate the tidal torque from each perturber, we assume the deep-MOND limit and that the EFE dominates. The former assumption is clear because of the large distances to the perturbers, while the latter assumption was justified in section 2.3.1 of BRZ18. As we are only interested in a rough estimate, we approximate that the gravity towards any object of mass  $M$  is  $GMa_0/(r^2g_{\text{ext}})$  at distance  $r$  (equation 8), neglecting the small angular dependence. Using also

**Table 8.** Our estimated tidal torque due to each perturber, defined as the change it causes in the MW–M31 relative tangential velocity (equation 16). The assumed perturber properties are as listed in table 2 of BRZ18. Despite the high mass of Cen A, it has only a small effect because it is almost on the MW–M31 line. The torquing effect of a perturber is maximized if  $|\cos \theta| \approx 1/\sqrt{2}$ , which is almost the case for IC 342. However, even its effect is substantially smaller than the PM uncertainty of  $\approx 50 \text{ km s}^{-1}$  (Fig. 4).

Perturber	$ \cos \theta $	Change in tangential velocity over 5 Gyr ( $\text{km s}^{-1}$ )
Cen A	0.99	2.43
M81	0.32	3.79
IC 342	0.76	8.22

the distant tide approximation and assuming that  $g_{\text{ext}} = 0.022 a_0$ , the relative MW–M31 acceleration in the direction orthogonal to their separation has a magnitude

$$g_{\text{tan}} = \frac{3GMa_0d\cos\theta\sin\theta}{r^3g_{\text{ext}}}, \quad (16)$$

where  $r$  is the distance from the MW–M31 mid-point to the perturber,  $\cos \theta$  is the angle between this direction and  $\hat{d}$ , and  $M$  is the mass of the perturber.  $r$  is measured from the MW–M31 mid-point because the MW and M31 are treated as independent test particles freely falling in the gravitational field of the perturber, so the individual MW and M31 masses are irrelevant. Thus, the accuracy of our calculations would be maximized if using the MW–M31 mid-point. For a perturber with  $r \gg d$ , it does not matter exactly which point we use as the ‘centre’ of the LG.

Assuming the same perturber properties as listed in table 2 of BRZ18 and that  $d = 783 \text{ kpc}$ , we obtain the tidal torque estimates given in Table 8 over a period of 5 Gyr, which is roughly the amount of time for which  $d$  is similar to its maximum value (Fig. 3). Tidal torques would be much less significant around the time of the flyby. As shown by the values in Table 8 and the green line in Fig. 4 representing  $10 \text{ km s}^{-1}$ , it is clear that tidal torques from perturbers hardly affect the present PM of M31. Its measurement accuracy would need to improve another order of magnitude for such subtle effects to be discernible, which would then necessitate more detailed modelling. The small effect of tidal torques arises mainly because of the isolated nature of the LG, but also because one of the most massive external galaxies which could tidally affect it (Cen A) lies almost on the MW–M31 line (Ma et al. 1998), limiting its tidal torque on the LG.

### 3.2.2 The external field torque (EFT)

In MOND, the mutual gravity between two masses does not necessarily align with their separation if there is an external field (e.g. Banik & Zhao 2018a). This creates a non-tidal torque on the MW–M31 system due to objects outside it. Since this is induced by the EFE, we call this the EFT. As described in appendix A of BRZ18, the tangential component of the relative gravity between two masses due to the EFT can be estimated as

$$g_{\text{tan}} = \sin \theta \left[ \cos \theta + \frac{4}{5} \sin^2 \theta \cos(\pi q_{\text{MW}}) \frac{\tilde{r}}{1 + \tilde{r}^2} \right] \times \frac{GMa_0r}{2g_{\text{ext}}(r^3 + r_i^3)}, \quad (17)$$

with the ‘−’ sign changed to ‘+’ in front of the  $4/5$ ... (see section 4.4 of Banik et al. 2020).  $\tilde{r}$  is the separation relative to that at which the

problem becomes EF-dominated:

$$\tilde{r} \equiv \frac{r}{r_t}, \quad r_t = \frac{\sqrt{GMa_0}}{Q g_{\text{ext}}}. \quad (18)$$

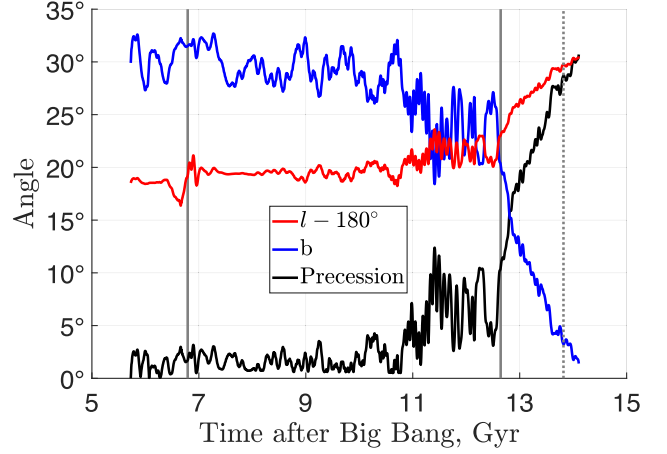
$\theta$  is the angle between  $\mathbf{d}$  and  $\mathbf{g}_{\text{ext}}$ , while  $Q = 0.7937$  for an MW mass fraction of  $q_{\text{MW}} = 0.3$  (equation 7). Note that in appendix A of BRZ18, the two masses considered were an external perturber (e.g. Cen A) and the whole LG, with the latter placed at the position of the MW or M31 depending on the circumstances. Their derivation leads directly to equation (17) if instead we take the two masses to be the MW and M31.

The present work improves significantly upon BRZ18 in that the EFT is directly included in POR (Section 2.4.3). However, this still relies on knowing the appropriate value of  $\mathbf{g}_{\text{ext}}$  and its even more uncertain time dependence. We can use equation (17) to estimate the significance of these factors. Since the torque is small for an isolated system ( $\tilde{r} \ll 1$ ), it is mainly important when the MW and M31 are close to apocentre. If we use  $r = 900$  kpc and integrate equation (17) over a 5 Gyr period (as in Section 3.2.1) assuming constant  $g_{\text{tan}}$ , we get a tangential impulse of  $\Delta v_{t,783} = 33.6 \text{ km s}^{-1}$  for  $g_{\text{ext}} = 0.03 a_0$ , including also a factor of  $900/783$  to account for changes in angular momentum due to the EFT having a greater impact on the present tangential velocity of M31 if it is currently closer to us. If instead we assume that  $g_{\text{ext}} = 0.02 a_0$ , then  $\Delta v_{t,783}$  is only  $25.9 \text{ km s}^{-1}$ . This difference of  $\approx 8 \text{ km s}^{-1}$  is also apparent at other separations, e.g. if we use  $r = 600$  kpc, we get that  $\Delta v_{t,783}$  differs by  $11 \text{ km s}^{-1}$  depending on whether  $g_{\text{ext}}$  is 2 or 3 per cent of  $a_0$ .

In principle, the functional form of the dependency on  $q_{\text{MW}}$  is not known as BRZ18 only had numerical results for the case  $q_{\text{MW}} = 0$ . This means we can replace the factor of  $\cos(\pi q_{\text{MW}})$  by any function with value 1 when  $q_{\text{MW}} = 0$  and which is antisymmetric with respect to  $q_{\text{MW}} \rightarrow 1 - q_{\text{MW}}$ . However, due to the low value of  $\sin^2 \theta = 0.167$ , these results hardly change if instead we use a mass ratio dependency of  $(1 - 2q_{\text{MW}})$  or  $\cos^3(\pi q_{\text{MW}})$  instead of  $\cos(\pi q_{\text{MW}})$ .

Since  $r$  around the time of apocentre typically has some intermediate value between 600 and 900 kpc (Fig. 3), we estimate that uncertainty in the behaviour of  $\mathbf{g}_{\text{ext}}$  causes an  $\approx 10 \text{ km s}^{-1}$  uncertainty in the PM of M31. This would not much alter the present MW–M31 orbital pole because  $v_{t,783} = 34.12 \text{ km s}^{-1}$  in our best-fitting POR model. Indeed, Fig. 6 shows that  $\hat{\mathbf{h}}$  does not change much due to the EFT already included in POR, so uncertainties in the EFT should have only a small effect. Interestingly,  $\hat{\mathbf{h}}$  also hardly precesses due to the flyby, which is indicated with a vertical solid grey line at pericentre (the later solid grey line represents apocentre). As a result, the final  $\hat{\mathbf{h}}$  is fairly similar to its initial orientation. The more rapid precession of  $\hat{\mathbf{h}}$  after the most recent apocentre is caused by the MW–M31 orbit becoming almost radial (notice the low value of  $h$  at that time in Fig. 3). It could also be related to tidal debris falling back on to the MW (Section 3.3.3).

Uncertainty in the EFT is mitigated by the geometric factors being well known: the MW–M31 line at apocentre must be quite close to its presently observed direction as they are on a nearly radial orbit (van der Marel et al. 2012, 2019; Salomon et al. 2021).  $\mathbf{g}_{\text{ext}}$  is also constrained by the observed motion of the LG with respect to the CMB/surface of last scattering (section 2.2 of BRZ18). Since the EF on the LG mostly arises from rather distant sources, the direction of  $\mathbf{g}_{\text{ext}}$  would not have changed much in the last 5 Gyr. As a result, the EFT at late times can only affect the PM of M31 along one particular direction. However, it is not too useful to speculate further about this because the PM of M31 still has an uncertainty  $\gg 10 \text{ km s}^{-1}$  (Fig. 4).



**Figure 6.** The MW–M31 orbital pole in our POR simulation, shown as Galactic latitude (blue) and Galactic longitude less  $180^\circ$  (red). The deviation from the initial orbital pole is shown using a solid black line. The vertical solid grey lines around 7 and 13 Gyr show the times of perigalacticon and the most recent apogalacticon, respectively, while the dotted grey line shows the present time of 13.82 Gyr. Notice the lack of precession around pericentre. Combined with the magnitude of the MW–M31 angular momentum hardly changing then (bottom panel of Fig. 3), it is clear that the MW–M31 interaction has little torquing effect on their orbit. Their orbital pole precesses more significantly around apocentre due to the EFT (Section 3.2.2). The present orbital pole direction ( $209.77^\circ, 3.29^\circ$ ) is very similar to that of the M31 SP.

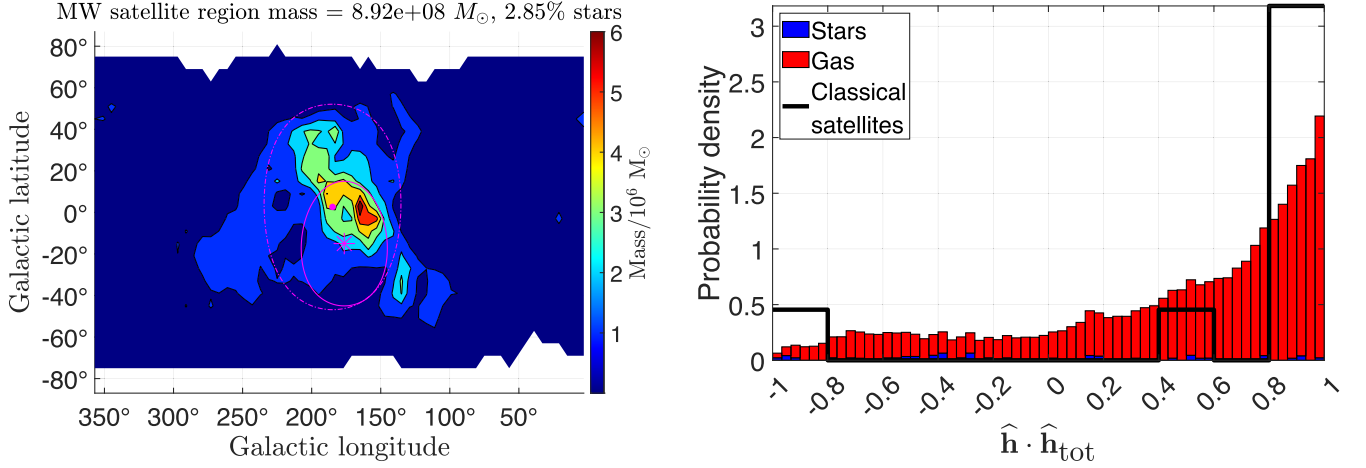
### 3.3 Tidal debris

We now turn to the distribution of tidal debris around the MW and M31 disc remnants. One of the main goals of our POR simulations is to check whether the tidal debris around each galaxy prefers a particular orbital pole, and if so, to compare this with the observed orbital pole of its SP. This requires us to define a ‘satellite region’ around each disc, which we take to be within 250 kpc of the barycentre found in Section 3.2 and  $z_{\text{max}} = 50$  kpc from the disc plane. For simplicity, we assume that the orientation of each disc matches observations – we show later that this is a fairly good assumption (Section 3.4) due to iterative adjustments to the initial orientation of each disc (Section 2.3.1). Even if the actual disc orientations differ somewhat, our choice of  $z_{\text{max}}$  should completely exclude the disc. However, only a small portion of the SPs would be lost as the satellite galaxies of the MW and M31 go out much further (see e.g. fig. 1 of Pawlowski 2018). We note that using a spherical excluded region for the disc does not work very well since e.g. excluding the central 30 kpc still leaves a considerable amount of material close to the disc plane at larger radii. This makes it very difficult to clearly disentangle the disc and SP, at least unless a much larger inner radius is considered – which then loses much of the satellite region. We therefore consider only a disc-shaped excluded region in what follows, or alternatively focus on precisely this region when analysing the disc remnants (Section 3.4). Appendix D shows the effect of varying  $z_{\text{max}}$ .

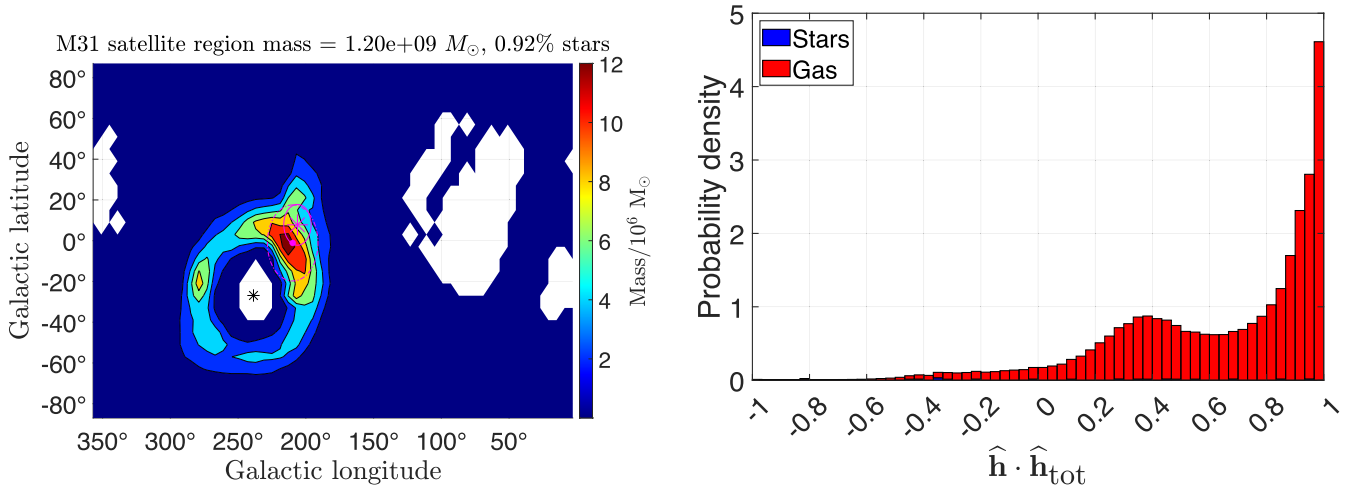
#### 3.3.1 Orbital pole distribution and mass

We find the angular momentum of all particles and gas cells in the satellite region of each galaxy relative to the galaxy’s barycentre. The important quantity for us is the direction of this angular momentum, which we use to build up a distribution in Galactic coordinates. The procedure is the same as that described in section 4.1 of BRZ18.





**Figure 7.** Left: The orbital pole distribution of material in the MW satellite region in the 8.2 Gyr snapshot of our POR simulation, shown in Galactic coordinates. The title shows the total mass in the satellite region and the fraction of this in stars. The colour of each pixel indicates the mass in a square of  $(l, b)$  with sides of length  $6^\circ$ . The pink star shows the observed orbital pole of the Galactic SP (Table 9). The solid pink curve shows a cone around this direction with opening angle of  $30^\circ$ . The pink dot shows the preferred orbital pole in the simulation (equation 19), with the dot-dashed pink curve showing the estimated orbital pole dispersion (equation 20). The lack of material at very low and high Galactic latitudes is caused by excluding material within  $z_{\max} = 50$  kpc of the disc plane. Results appear similar for a different  $z_{\max}$  (Appendix D). Right: The red bars show the cosine distribution of the angle between the orbital pole  $\hat{\mathbf{h}}$  and  $\hat{\mathbf{h}}_{\text{tot}}$ , the preferred orbital pole of the simulated SP (equation 19). The open black bars show the orbital pole distribution of classical MW satellites at Galactocentric distances of 50–250 kpc based on the positions and orbital poles listed in table 2 of Pawłowski & Kroupa (2020). The angles shown here are relative to the observationally preferred direction listed in Table 9.



**Figure 8.** Similar to Fig. 7, but showing results for M31. Its disc spin vector is shown as a black star. The lack of material with an orbital pole close to this direction or its opposite is caused by the exclusion of material close to the M31 disc plane. Notice the smaller simulated disc-SP misalignment compared to the MW, which is likely due to M31 having a higher mass and a shorter scale length for the outer disc (Table 2). The simulated and observed orbital pole dispersions are also much smaller than for the MW. Results appear similar for a different  $z_{\max}$  (Appendix D).

The resulting orbital pole distribution is shown in Fig. 7 for the MW and in Fig. 8 for M31. The observed orbital poles of the MW and M31 SPs (summarized in Table 9) are obtained from section 3 of Pawłowski & Kroupa (2013) and section 4 of Pawłowski et al. (2013), respectively. These are shown as pink stars on the left-hand panels, which our POR model matches fairly well. In Appendix D, we show that the appearance remains quite similar if we alter  $z_{\max}$  to 40 or 60 kpc.

An important aspect of our analysis is an estimate for the dispersion in orbital pole directions. On the observational side, we use  $30^\circ$  for the MW based on section 4 of Pawłowski & Kroupa (2013). For M31, we note that its SP has an aspect ratio  $3\times$  smaller than for the MW (table 3 of Pawłowski et al. 2013). We therefore adopt an orbital

**Table 9.** The observed orbital pole of the MW SP (section 3 of Pawłowski & Kroupa 2013) and of the M31 SP (section 4 of Pawłowski et al. 2013) in Galactic coordinates. We also show the angle between the spin vector of each SP and that of its parent disc, and the angle between the SP spin vectors. The MW–M31 orbital pole (shown in Fig. 6) aligns fairly closely with that of the simulated M31 SP (shown in Fig. 8).

Galaxy	MW	M31
SP spin vector	$(176.4^\circ, -15.0^\circ)$	$(206.2^\circ, 7.8^\circ)$
Disc-SP misalignment	$75^\circ$	$47^\circ$
Angle between SP spins	$37^\circ$	

pole dispersion of  $10^\circ$  for M31. These dispersions are illustrated by drawing a cone with an opening angle equal to the estimated dispersion and an axis aligned with the observed SP orbital pole direction listed in Table 9. These cones are shown on the left-hand panels of Figs 7 and 8 using solid pink curves.

The dashed pink curves on these figures show analogous results for the simulated tidal debris, whose preferred orbital pole  $\hat{\mathbf{h}}_{\text{tot}}$  is shown with a pink dot in each case. We estimate this using an iterative procedure where the initial guess is the centre of the pixel in  $(l, b)$  with the most mass. We then find

$$\hat{\mathbf{h}}_{\text{tot}} \propto \sum_i m_i \hat{\mathbf{h}}_i, \quad (19)$$

where each particle or gas cell in the satellite region has mass  $m_i$  and orbital pole direction  $\hat{\mathbf{h}}_i$  relative to its host galaxy. The sum is taken over only those particles whose  $\hat{\mathbf{h}}_i$  aligns with  $\hat{\mathbf{h}}_{\text{tot}}$  to better than  $30^\circ$  for M31 or  $90^\circ$  for the MW, i.e.  $3\times$  the observed orbital pole dispersion in both cases. This restriction causes  $\hat{\mathbf{h}}_{\text{tot}}$  to influence which particles and gas cells contribute to the sum, so we need to repeat the process a few times until convergence is reached. We find that only a handful of iterations are required to reach convergence in  $\hat{\mathbf{h}}_{\text{tot}}$  to within machine precision. We then calculate the orbital pole dispersion  $\theta_{\text{rms}}$  using

$$\theta_{\text{rms}} = \sqrt{\frac{\sum_i m_i \theta_i^2}{\sum_i m_i}}, \quad (20)$$

where  $\theta_i \equiv \cos^{-1}(\hat{\mathbf{h}}_i \cdot \hat{\mathbf{h}}_{\text{tot}})$  is the angle between  $\hat{\mathbf{h}}_{\text{tot}}$  and the orbital pole of particle  $i$ . The sum is again taken over only those particles or gas cells in the satellite region whose  $\hat{\mathbf{h}}_i$  is within the above-mentioned cone around  $\hat{\mathbf{h}}_{\text{tot}}$ . The so-obtained orbital pole dispersion is  $\theta_{\text{rms}} = 49^\circ$  for the MW and  $19^\circ$  for M31, so our model naturally yields a lower  $\theta_{\text{rms}}$  around M31. The simulated dispersions are larger than the observed ones, which we ascribe to the somewhat high 372 kK temperature floor of our POR simulations due to resolution limitations. It is also likely that individual TDGs would form only in the densest regions, leading to a narrower spread of orbital poles than for the tidal debris considered as a whole.

The left-hand panels of Figs 7 and 8 reveal the expected gap in the orbital pole distribution around the disc spin vector and the opposite direction arising from our definition of the satellite region. We clarify this in the M31 case by displaying its observed disc spin vector ( $238.65^\circ, -26.89^\circ$ ) as a black star in Fig. 8 (section 2.1 of Banik & Zhao 2018c). This is omitted for the MW because by definition its disc spin points towards the south Galactic pole, leading to a lack of material at very low and high Galactic latitudes in Fig. 7.

The right panels of these figures show the mass-weighted distribution of  $\cos \theta_i$ , which should be uniform for a completely isotropic distribution. In the MW case, we use open black bars to show the observed distribution for classical satellites at Galactocentric distances of 50–250 kpc (table 2 of Pawlowski & Kroupa 2020). The result is similar to what we obtain in our POR simulation: both the model and observations show a significant clustering of orbital poles, though observationally the clustering is somewhat tighter. This could be related to the resolution and temperature floor of our model. The orbital poles are much more clustered for the tidal debris around M31 than for the MW, which as explained earlier is in line with observations as the M31 SP is much thinner than that of the MW.

The titles of Figs 7 and 8 indicate that the satellite region of each galaxy has a mass of  $\approx 10^9 M_\odot$ . Since we expect the SP material to mostly have been quite far out initially (section 5.2.4 of BRZ18), our estimated SP masses are quite sensitive to the reliability of

extrapolating the assumed exponential disc law to large radii. The relatively small amount of material here means that altering the initial mass distribution at large radii would hardly affect the gravitational field, leaving the simulated SP orientations unchanged. Therefore, the SP masses are not a strong test of our model. Bearing this in mind, we note that an SP mass of  $10^9 M_\odot$  is reasonable for the MW because the LMC dominates the baryonic mass in the Galactic SP. The RC of the LMC has a flatline level of  $\approx 70 \text{ km s}^{-1}$  (Alves & Nelson 2000; van der Marel & Kallivayalil 2014; Vasiliev 2018), which in a MOND context implies a mass of  $1.5 \times 10^9 M_\odot$ . While this is in reasonable agreement with our model, another important aspect of the LMC is its rather high specific angular momentum  $h$  (Kallivayalil et al. 2013). This is quite difficult to reproduce in our model, with only a small fraction of the tidal debris having a higher  $h$  than that of the LMC (Section 3.3.4). It is difficult to solve this problem by postulating a much larger amount of tidal debris as then the Galactic disc would need to be damaged much more significantly. Instead, the problem may lie with resolution and the temperature floor: colder gas would need more support from rotational motion, which should lead to higher  $h$  initially. It is not presently clear whether this will solve the problem of the LMC, so for the time being its properties remain somewhat problematic for our flyby scenario.

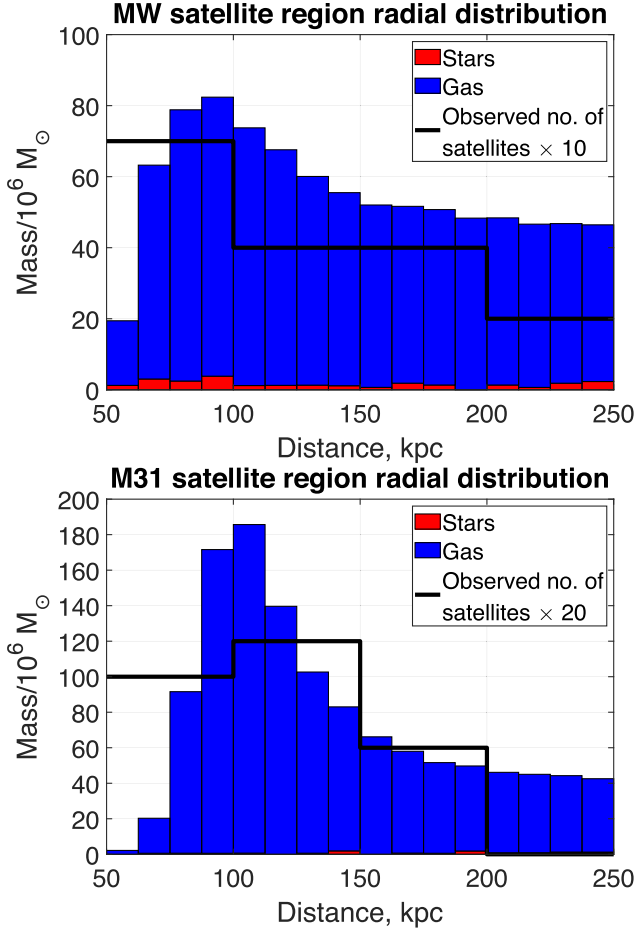
In the M31 case, the simulated SP mass suggests that M32 is likely part of this structure (see section 5.2.3 of BRZ18), and/or that all of the mass in its SP has not condensed into individual satellites. This is quite possible given that the orbital pole distribution is a non-uniform ring rather than a single point (Fig. 8). It could well be that only in the densest part of this ring was the density high enough for the gas to undergo Jeans collapse into TDGs. Moreover, gas accreted on to a newly formed TDG could be subsequently expelled by feedback, with perhaps only a small fraction ending up in stars.

To address these issues more thoroughly, a higher resolution simulation would be required in which bound satellite galaxies form out of the tidal debris, which is beyond the scope of this project. A previous high-resolution hydrodynamical MOND simulation of interacting disc galaxies suggests that TDGs should form out of the tidal debris (Renaud et al. 2016). Their work focused on the Antennae (Mirabel et al. 1992), but our work indicates that a similar process could well have played out in the much better observed LG. We hope that the initial conditions of our best-fitting model (Table 4) serve as a starting point for further work on the LG in MOND.

### 3.3.2 Radial distribution

We use the top and bottom panel of Fig. 9 to show the radial distribution of material in the satellite region of the simulated MW and M31, respectively. Each bar is divided into a red part indicating stars and a blue part indicating gas. It is clear that the satellite regions are completely dominated by gas. For the MW, this might be caused by its initial distribution of gas being more extended than that of its stars (Table 2). However, the stars and gas in M31 have the same initial surface density profile. The dominance of gas in the satellite region could indicate that this is more easily removed from the disc than stars due to ram pressure effects.

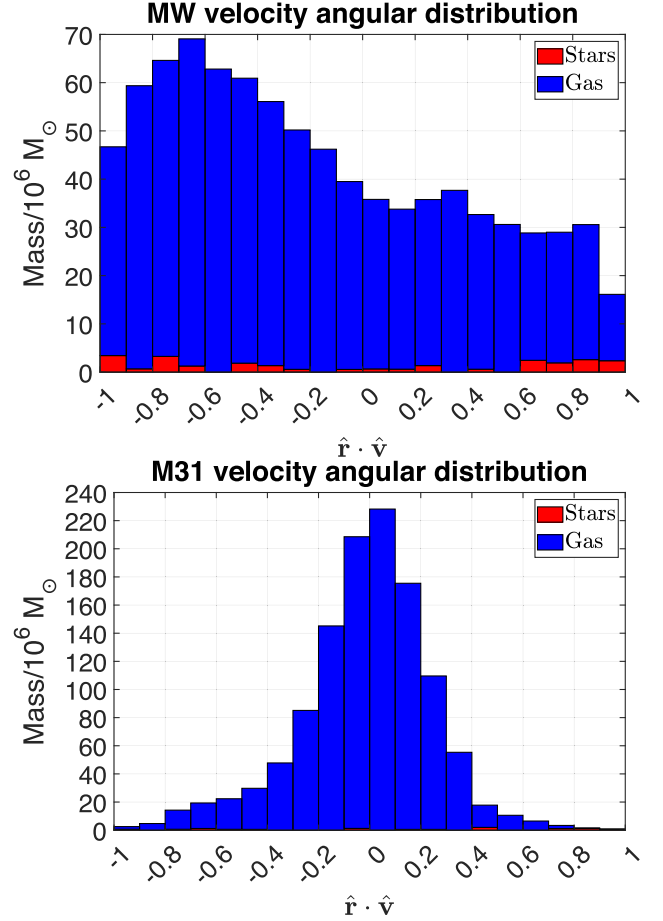
Our results can tentatively be compared with the observed distances of LG SP members from their host galaxy. We obtain these from the bold entries in table 4 of Pawlowski et al. (2013), restricting further to only those satellites within 50–250 kpc of their parent galaxy. These observational results are shown in Fig. 9 using a solid



**Figure 9.** The radial distribution of material in the MW (top) and M31 (bottom) satellite regions, showing contributions from stars (red) and gas (blue). The observed satellite distribution is shown using a stepped solid black line representing the number of satellites in four equally wide bins covering 50–250 kpc. These were obtained from the bold entries in table 4 of Pawlowski et al. (2013) and scaled for clarity as indicated in the legend. Our POR simulation suggests that the SPs would initially have been dominated by gas, though this need not be the case today as we do not simulate star formation and feedback.

black line. There is good overall agreement with our model, but observationally there is a lack of observed satellites 200–250 kpc from their host. This could be due to the tidal debris at such large distances being too sparse to form TDGs. Selection effects could also play a role, especially around the MW where more distant satellites would be fainter. While this effect is less significant around M31, the limited size of the Pan-Andromeda Archaeological Survey (Ibata et al. 2014a; McConnachie et al. 2018) means that observations of M31 do not extend to galactocentric distances much beyond 150 kpc.

Our simulations do not allow star formation, so the high gas fraction in the satellite regions merely indicates that the flyby created an initially gas-rich distribution of tidal debris. In a more advanced simulation, the gas may well condense into a small number of TDGs which then form stars (Renaud et al. 2016). While the formation of TDGs and their internal evolution is beyond the scope of this project, their overall distribution should be similar in a higher resolution simulation that allows star formation. In this respect, it is interesting that the radial distribution of simulated material in the satellite region of each galaxy broadly agrees with that of its actual SP members.



**Figure 10.** The cosine distribution of the angle between the position and velocity of each particle relative to its host galaxy, shown for the MW (top) and M31 (bottom) satellite regions. Contributions are shown from stars (red) and gas (blue). If the velocities were oriented randomly relative to the radial direction, then the distribution would be uniform. The asymmetry between  $\pm \hat{\mathbf{r}} \cdot \hat{\mathbf{v}}$  for the MW indicates that material is still falling towards it. Notice the strong preference for tangential motion in the M31 case.

### 3.3.3 Orbital eccentricity

An important aspect of the LG satellite planes is their small dispersion of orbital pole directions (Pawlowski 2021). Another aspect to consider is the motions of satellite galaxies within their preferred plane. To quantify this, we determine the position  $\mathbf{r}$  and velocity  $\mathbf{v}$  of each particle or gas cell relative to its parent galaxy (Section 3.2). We then find the unit vectors  $\hat{\mathbf{r}}$  and  $\hat{\mathbf{v}}$ , where  $\hat{\mathbf{n}} \equiv \mathbf{n}/n$  for any vector  $\mathbf{n}$ . The distribution of the angle between  $\hat{\mathbf{r}}$  and  $\hat{\mathbf{v}}$  provides a measure of whether the velocity dispersion tensor is isotropic or tangentially biased, and also whether there is any net radial motion that shows up as an asymmetry between  $\pm \hat{\mathbf{r}} \cdot \hat{\mathbf{v}}$ .

We use Fig. 10 to plot the distribution of  $\hat{\mathbf{r}} \cdot \hat{\mathbf{v}}$ . The top panel shows the result for the MW. If the  $\hat{\mathbf{v}}$  for each particle or gas cell were distributed randomly relative to the outwards radial direction  $\hat{\mathbf{r}}$ , then  $\hat{\mathbf{r}} \cdot \hat{\mathbf{v}}$  would be distributed uniformly over the range  $(-1, 1)$ . This is a reasonable description of the figure, so there is not much tendency for the orbits to be tangentially biased, somewhat at odds with observations (Cautun & Frenk 2017; Riley et al. 2019; Hammer et al. 2021). This is probably due to the high gas temperature floor of 372 kK, which is required given numerical limitations. In this sense, our result for the MW is similar to fig. 8 of BRZ18. One



difference is that unlike in their work, our results show an asymmetry between  $\pm \hat{\mathbf{r}} \cdot \hat{\mathbf{v}}$  such that radial inflow is preferred over radial outflow. This indicates that according to our simulation, some tidal debris should still be falling back towards the MW. This could explain the properties of the Fornax dwarf spheroidal satellite (Yang et al. 2022).

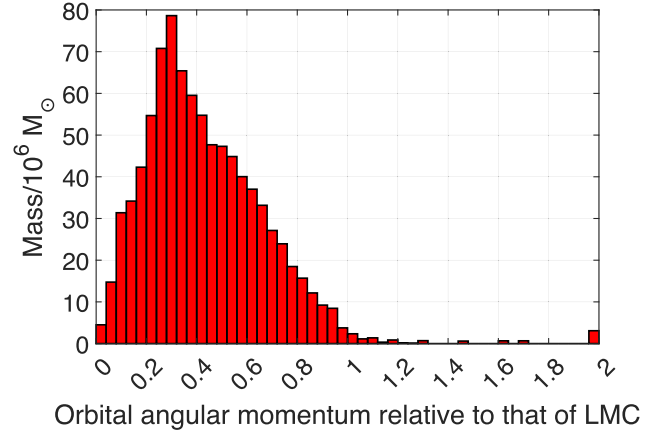
The bottom panel of Fig. 10 shows the  $\hat{\mathbf{r}} \cdot \hat{\mathbf{v}}$  distribution for M31. There is a sharp peak near 0, indicating a strong preference for tangential motion. This is quite different to the result obtained by BRZ18, whose fig. 8 shows only a mild preference for tangential motion for the tidal debris around M31. The hydrodynamical nature of our simulations is probably the main reason for this difference. Since a higher resolution simulation with a lower temperature floor should have even more dissipation, our POR simulation indicates that the M31 SP members should have rather low orbital eccentricities. This is in line with the observed PMs of the only two M31 SP members whose PMs are currently known (Sohn et al. 2020; Pawlowski & Tony Sohn 2021). There is not much tendency for radial inflow as opposed to outflow: at low values of  $|\hat{\mathbf{r}} \cdot \hat{\mathbf{v}}| \lesssim 0.5$ , there is more material at positive values, indicating a tendency for outflow. However, this trend is reversed at high values of  $|\hat{\mathbf{r}} \cdot \hat{\mathbf{v}}| \gtrsim 0.5$ . Thus, the M31 SP should be more nearly virialized than the MW SP. The strong tendency for tangential motion in the M31 SP also suggests that it should be dynamically colder than the MW SP, thereby having a lower aspect ratio and orbital pole dispersion. This is indeed the case in our model (equation 20) and observationally (e.g. Pawlowski et al. 2013). The full 6D phase space structure of the M31 SP is still unclear because this would require accurate PMs for its relatively faint member satellites, which have only recently become available in two cases (Sohn et al. 2020). While these are indicative of corotation, a more detailed analysis will need to await further data (Pawlowski & Tony Sohn 2021).

### 3.3.4 Analogues to the LMC

The phase space structure of the Galactic SP is much better known because of its proximity. This is clearly dominated by the LMC, so a realistic model should have material on an orbit similar to the LMC. To investigate this, we conduct a very similar analysis to section 5.4 of BRZ18. The LMC is fairly close to pericentre at the moment (Kallivayalil et al. 2013), which could well be something that even a quite realistic model does not reproduce simply because the considered snapshot time is not when the LMC analogue is at pericentre. Indeed, analogues to the Magellanic Clouds are not very common observationally, so it may be quite difficult to reproduce certain aspects of them even with the correct theory of nature (Robotham et al. 2012). Therefore, we focus on quantities which should be rather more robust around the whole orbit.

The high tangential velocity of the LMC for its position indicates a rather high  $h$  in the Galactocentric frame. To explain this, we should have material in the satellite region with a similar  $h$ . The distribution of  $h$  is shown in Fig. 11. The LMC value is based on table 5 of Kallivayalil et al. (2013), which implies a Galactocentric tangential velocity of  $314 \pm 24 \text{ km s}^{-1}$  at a distance of  $r_{\text{LMC}} = 49.39 \text{ kpc}$  (Pietrzyński et al. 2013). More recent observations indicate a similar LMC PM (Gaia Collaboration 2018). It is clear that the high  $h$  of the LMC is somewhat challenging for our model given its mass, though the simulated  $h$  distribution for material in the MW satellite region does extend beyond the LMC value.

An important aspect of the LMC is its high orbital eccentricity. This is not necessarily captured merely by looking at  $h$ , because  $h$



**Figure 11.** The specific angular momentum distribution of material in the MW satellite region, shown relative to that of the LMC. The rightmost bar includes all material with higher values.

would be rather high even for material on a circular orbit if it is sufficiently far out. We address this by following equation (51) of BRZ18 in defining an adjusted specific energy

$$\tilde{E} \equiv \Phi(r) + \frac{1}{2}v^2 - \frac{h^2}{2r_{\text{LMC}}^2}, \quad (21)$$

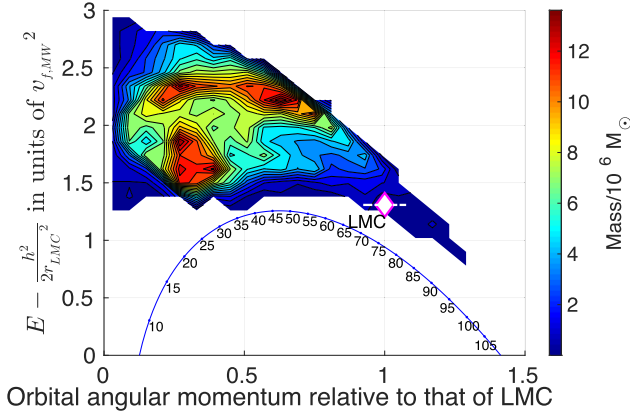
where the distance  $r$  and velocity  $v$  are Galactocentric.  $\tilde{E}$  is the specific radial kinetic energy of a particle at  $r = r_{\text{LMC}}$ , with  $\tilde{E} < 0$  indicating that the particle is not capable of reaching this distance. It would have little difficulty doing so if  $r \approx r_{\text{LMC}}$  in the simulation. But if  $r \gg r_{\text{LMC}}$ , then the particle would need to undergo substantial radial oscillations to ever resemble the LMC at any point along its orbit. If it does so, then ideally it should match the  $\tilde{E}$  value of the LMC, i.e. have an RV of  $64 \text{ km s}^{-1}$  at  $r = r_{\text{LMC}}$ . To calculate if this is the case, we need to assume some form for the Galactic potential  $\Phi(r)$ . We find this by treating the MW as an isolated point mass, neglecting M31 and the EFE. This is justified because the relevant distances are much larger than the extent of the MW disc (Bovy & Rix 2013), but much smaller than the distance to M31 (McConnachie 2012) or the distance beyond which the EFE dominates (Banik & Zhao 2018b). The form of  $\Phi$  follows from integrating the simple interpolating function (Famaey & Binney 2005) for a single point mass using a hyperbolic substitution.

$$\Phi(r) = \sqrt{GMa_0} \left[ \ln \left( 1 + \sqrt{1 + \tilde{r}^2} \right) - \frac{1}{\tilde{r}} - \sqrt{\frac{1}{\tilde{r}^2} + 1} \right],$$

$$\tilde{r} \equiv \frac{2r}{r_M}, \quad r_M \equiv \sqrt{\frac{GM}{a_0}}. \quad (22)$$

Here,  $M = 9.15 \times 10^{10} M_\odot$  is the initial mass of the MW, while  $r_M$  is its MOND radius, the distance beyond which MOND effects become significant. Since  $r_M \approx 10 \text{ kpc}$ , it is not much smaller than the LMC distance, so it is not accurate to assume that the problem is in the deep-MOND limit.

The joint distribution of  $\tilde{E}$  and  $h$  is shown in Fig. 12, which also has a blue curve showing values for particles on circular orbits at different  $r$  (indicated in kpc). Our results show that if considering both parameters jointly, the observed LMC value is near the edge of the simulated distribution. Its  $\tilde{E}$  is somewhat on the low side, while its  $h$  is on the high side. None the less, the observed combination of  $(\tilde{E}, h)$  for the LMC is still within the simulated distribution for



**Figure 12.** Distribution of the adjusted energy (equation 21) and specific angular momentum relative to that of the LMC, shown for material in the MW satellite region assuming a point mass Galactic potential (equation 22). The LMC value is shown using a white diamond, with the horizontal dashed line through it showing a  $24 \text{ km s}^{-1}$  uncertainty in its Galactocentric tangential velocity (see the text). We neglect uncertainty in its RV because uncertainty in its PM is much larger. The solid blue curve shows values for particles on circular orbits, with the orbital radius indicated in kpc as a text label next to the corresponding filled blue dot.

material in the MW satellite region. Therefore, the LMC does not obviously present an insurmountable difficulty for our model. Given also that material in the satellite region shows a sharp concentration of orbital poles at about the right direction (Fig. 7), our model provides a plausible explanation for how one could get a large satellite with a rather high angular momentum on a quite eccentric nearly polar orbit aligned with the SP.

Our model strongly suggests that the LMC formed as a TDG out of the MW–M31 flyby. Its high mass could indicate that it formed at the tip of the Galactic tidal tail, where a massive TDG is more likely to form due to the weaker tidal stress and EFE combined with the tendency for material to pile up at apocentre. In this scenario, the LMC must have been orbiting the MW for  $\approx 7$  Gyr. This is quite feasible in a MOND context as there is almost no dynamical friction 40 kpc from a purely baryonic MW. Moreover, the time dependence of the gravity from M31 could cause perigalacticon distances to vary somewhat between orbits (see e.g. fig. 20 of Banik & Zhao 2018d). Assuming that the SMC also formed as a TDG on a similar orbit to the LMC as part of the Galactic SP, one can imagine that the LMC and SMC subsequently underwent one or more interactions, helping to explain the observed properties of the Magellanic Stream (Hammer et al. 2015; Lucchini et al. 2020; Lucchini, D’Onghia & Fox 2021). Indeed, it seems difficult to altogether avoid close interactions between a large number of TDGs confined within a 2D structure by virtue of how they formed.

The top panel of Fig. 9 indicates that the Galactic SP would initially have been mostly gas, though a small contribution from stars is also expected. This might explain how the LMC came to have stars older than the flyby (Harris & Zaritsky 2009; Nidever et al. 2021). Our scenario would struggle to explain a significant fraction of stars older than the flyby, though this depends on the efficiency with which gas in the tidal arm condenses into stars in the LMC. A lower efficiency would increase the relative importance of stars formed before the flyby. The star formation histories of MW satellites are discussed further in Section 4.1.

A common origin during a past MW–M31 flyby provides a natural explanation for the alignment between the Galactic orbit of the

LMC, the LMC–SMC orbit, and the MW SP (fig. 1 of Pawlowski, McGaugh & Jerjen 2015). These planes also align well with that defined by the Magellanic Stream (fig. 1 of Pawlowski & Kroupa 2020). If instead the LMC and SMC were recently accreted by the MW due to dynamical friction with its CDM halo (Besla et al. 2007), then these geometric alignments would be fortuitous. This is because the orientation of the MW SP would a priori be unrelated to the LMC–SMC binary orbit, which could itself be oriented differently to their barycentre’s motion around the MW (section 5.3.1 of Kroupa 2015). One exception is if the dynamical influence of the LMC is responsible for most Galactic satellites having orbital poles in a very narrow range of directions, as proposed recently in the  $\Lambda$ CDM context where the LMC should be more massive relative to the MW than in MOND (Garavito-Camargo et al. 2021).<sup>9</sup> However, their idealized simulation only works because the test particles they consider have a very low specific angular momentum, so even a very small tidal torque can significantly change the orbital pole. This scenario is untenable for the classical MW satellites as they have quite high specific angular momenta, which is related to their strong preference to be moving tangentially rather than radially (Hammer et al. 2021). Consequently, including the influence of the LMC in a backwards integration (Correa Magnus & Vasiliev 2022) still leaves the MW satellites with a clustered orbital pole distribution prior to the LMC infall, implying that this scenario does not explain the clustered orbital pole distribution of MW satellites in the  $\Lambda$ CDM context (Pawlowski et al. 2021).

The LMC could also be related to the MW SP rather more directly if the LMC brought in its own retinue of satellites (Samuel et al. 2021). However, their claim to explain the MW SP in  $\Lambda$ CDM contradicts the fact that group infall is already included in cosmological  $\Lambda$ CDM simulations.<sup>10</sup> Based on these, it was previously argued that ‘having most satellites accreted as a single group or along a single filament is unlikely to explain the MW disc of satellites’ (Shao et al. 2018), an important reason being that dwarf galaxy groups are typically much larger than the thickness of the LG SPs (fig. 1 of Metz et al. 2009b). It is difficult for an LMC-mass galaxy to bring in too many satellites due to its lower mass than the MW, so the Galactocentric orbit of the LMC should be aligned with a pre-existing chance alignment of Galactic satellites into a planar structure. A recently fallen in group would also span only a narrow range of angles on the sky (Santos-Santos et al. 2021), whereas the MW SP members go around most of the sky (fig. 2 of Pawlowski & Kroupa 2020). Group infall is already included in cosmological  $\Lambda$ CDM simulations like Illustris, so postulating it as a reasonably likely explanation for the MW SP is not meaningful when the low likelihood of this structure in a  $\Lambda$ CDM universe has already been demonstrated in a cosmological context (Pawlowski & Kroupa 2020), especially since high resolution hydrodynamical simulations indicate that the impact of baryonic physics is small (Samuel et al. 2021). Moreover, the velocities of MW satellites are biased tangentially (Hammer et al. 2021), even though a radial bias is expected from infall (Angus, Diaferio & Kroupa 2011). Our model is quite promising in this respect because it yields a significant tangential bias to the velocities for the tidal debris around M31, though it does not achieve this for the MW (Fig. 10). However, a dissipative origin for the Galactic

<sup>9</sup>Conroy et al. (2021) recently obtained tentative evidence for the perturbation induced by a massive LMC on the Galactic stellar halo.

<sup>10</sup>An exception is if MW analogues very rarely have an LMC analogue around them, but these rare cases are also those with a satellite system similar to that of the MW. In this case,  $\Lambda$ CDM would not be consistent with the LMC.

**Table 10.** Spin vectors of the MW and M31 discs at the start of our POR simulation and in its 8.2 Gyr snapshot, comparison of which yields the angle by which each disc has precessed. We also show the observed spin vector of each disc, which for M31 follows from its observed inclination and kinematic position angle (section 2.1 of Banik & Zhao 2018c).

Galaxy	MW	M31
Initial	(55.11°, −84.89°)	(238.38°, −33.71°)
Final	(231.05°, −85.63°)	(237.72°, −26.15°)
Observed	≡ (0°, −90°)	(238.65°, −26.89°)
Precession	9.47°	7.59°
Mismatch with observations	4.37°	1.11°

SP remains a promising explanation for why its members have tangentially biased orbits, especially when considering the numerical limitations of our simulation.

A successful theory should explain not only the MW SP but also the SPs around M31 (Ibata et al. 2013, 2014b) and Cen A (Müller et al. 2018a, 2021) without postulating a highly unusual chance alignment in all cases. Rather, it should provide a mechanism for generating the significant observed anisotropy in the only three systems where the 3D distribution of satellites is well known. Ideally, this would be done with minimal assumptions, e.g. based on the known existence of dissipative baryons and the known formation of phase space-correlated TDGs (Mirabel et al. 1992; Bournaud et al. 2004). In keeping with Occam’s Razor, it is also preferable to use only one hypothetical galaxy interaction to explain both SPs in the LG.

### 3.4 The simulated MW and M31 discs today

Our results in Section 3.3 show that only a small fraction of the initial  $3.05 \times 10^{11} M_{\odot}$  in the MW and M31 discs ends up in their satellite regions. This indicates that the interaction is not very strong and should still leave recognizable disc remnants, as also suggested by Figs 1 and 2. The reason is that the perigalacticon distance of 81.47 kpc is much more than the disc scale length of either galaxy (Table 2), which also causes there to be only a small amount of dynamical friction (Fig. 3). This is encouraging for our scenario because a strong encounter that transforms the MW and M31 discs into ellipticals would not be realistic. In this section, we take a closer look at the MW and M31 disc remnants (Sections 3.4.1 and 3.4.2, respectively). In each case, our analyses focus on the region within 250 kpc of the galaxy’s barycentre and within 50 kpc of its disc plane. Combining our analyses of the disc and satellite regions therefore covers the entire 250 kpc sphere around each galaxy.

We begin by finding the total angular momentum of all material within the disc region of each galaxy relative to the galaxy’s barycentre. This gives the disc spin vector of each galaxy in the analysed POR simulation snapshot, which we list in Table 10 in Galactic coordinates. This also shows the initial and observed disc spin vector in each case, and the amount by which the disc has precessed during the simulation. As discussed in section 4.2.5 of Banik et al. (2020), this precession is partly due to the EFE, though we expect the flyby to also induce disc precession. The total precession is  $\approx 10^\circ$  in both cases, so differential precession of different parts of the MW/M31 disc should not warp it too much. Importantly, the simulated discs end up oriented similarly to observations thanks to our iterative adjustments (Section 2.3.1), especially in the M31 case where the disc-SP misalignment is smaller (Fig. 8). In what follows, we analyse each disc in a reference frame rotated so the  $z$ -axis corresponds to the observed disc orientation as listed in Table 10. We

have checked explicitly that the results of our analyses differ very little if we instead align the  $z$ -axis to the spin vector of each simulated disc as given in Table 10. We avoid doing this for simplicity as it would require iterative adjustments to the simulated spin vector of each disc.

#### 3.4.1 MW stellar disc remnant

The top left panel of Fig. 13 shows the face-on view of the stars in the MW 8.2 Gyr after the start of our POR simulation. This shows a regular appearance. We also show the MW in the cylindrical  $rz$  projection system described in section 3.2 of Banik et al. (2020). The so-obtained  $rz$  view of the stars in the MW is shown in the top right panel. This confirms that a disc remains despite the flyby and other simulated effects.

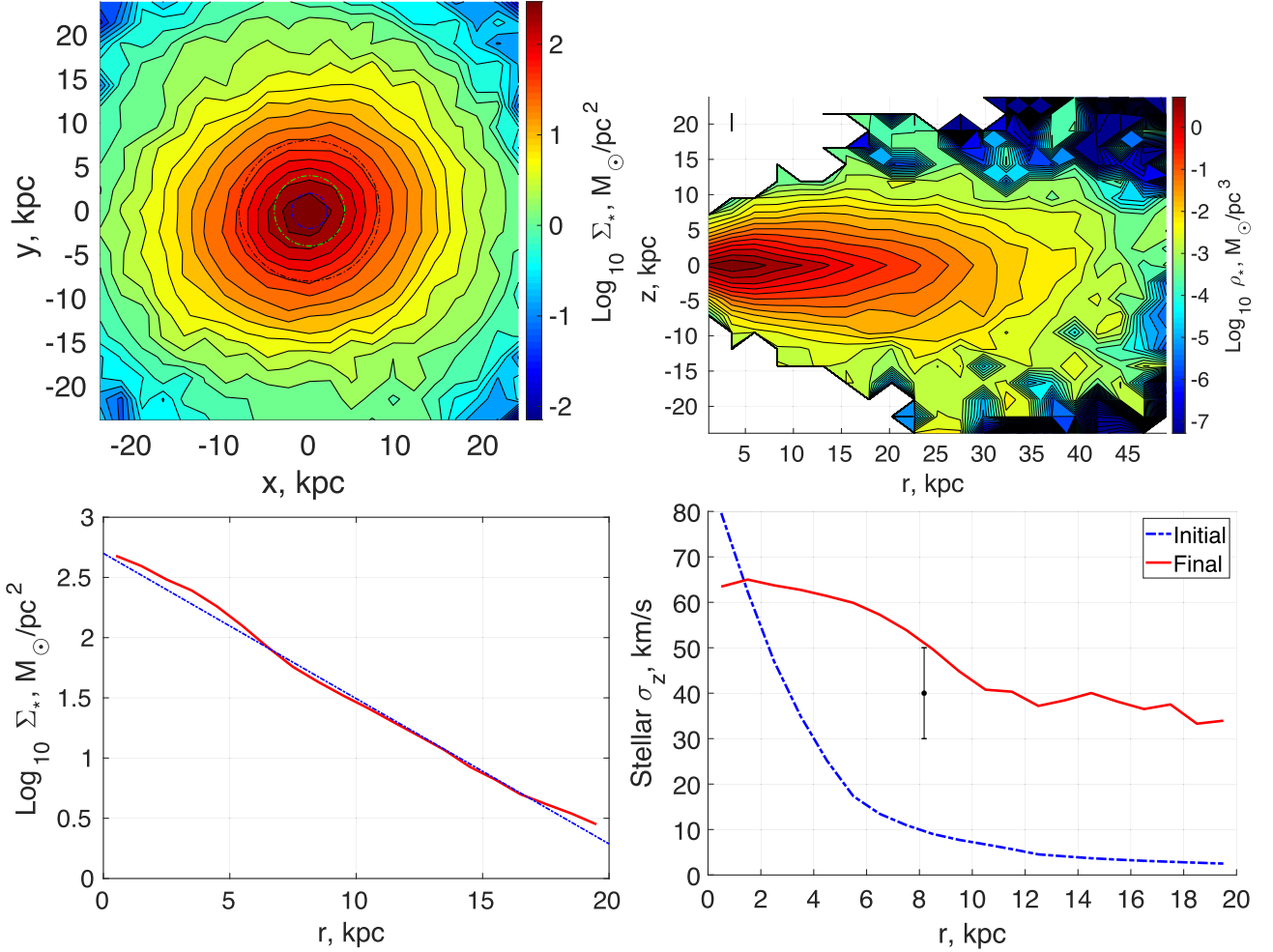
Our results allow us to obtain the surface density profile of the MW’s stellar component. Since our simulation does not allow star formation, this should correspond to something akin to the old Galactic thick disc (Gilmore & Reid 1983), whose stars pre-date the flyby. By binning the results in cylindrical polar radius, we obtain the stellar surface density profile  $\Sigma_{\star}(r)$  shown in the bottom left panel of Fig. 13. The red line shows the simulated result, while the blue line shows an exponential law with a scale length of 3.6 kpc (Jurić et al. 2008). This is consistent with Jayaraman et al. (2013), who found that the thick disc scale length is  $\approx 4$  kpc (see their section 4.4). It is also in line with Li & Zhao (2017) and Mateu & Vivas (2018).

Another test of our model is the vertical velocity dispersion  $\sigma_z$  at different  $r$ , especially at the Solar circle. We find this by binning the stars in  $r$  and finding their mass-weighted  $\sigma_z$  using an iterative outlier rejection procedure with threshold of  $3.29\sigma$ , corresponding to the 0.1 per cent tail of a Gaussian distribution (the procedure is described further in section 3.2.1 of Banik et al. 2020). Our result is shown in the bottom right panel of Fig. 13. At the Solar circle of 8.2 kpc (McMillan 2017; Gravity Collaboration 2019), the simulated  $\sigma_z$  rises from an initial value of  $\approx 10 \text{ km s}^{-1}$  to a present value of  $\approx 50 \text{ km s}^{-1}$ . Note that since our model does not allow star formation, this should be compared with only those stars which existed at the time of the MW–M31 flyby. In this regard, we mention that the old metal-poor sample at large height in e.g. fig. 4 of Yu & Liu (2018) has  $\sigma_z \approx 40 \text{ km s}^{-1}$ , though results differ by  $\approx 10 \text{ km s}^{-1}$  depending on the exact sample used. We therefore show a data point at  $40 \pm 10 \text{ km s}^{-1}$  at the Solar circle. In a higher resolution simulation with all other parameters fixed,  $\sigma_z \approx 40 \text{ km s}^{-1}$  at the Solar circle – we present this in Appendix E. While results at this level of detail may differ in a more advanced simulation, the broad agreement in the disc scale length and  $\sigma_z$  is encouraging.

#### 3.4.2 M31 stellar disc remnant

The face-on view of M31 is shown in the top left panel of Fig. 14, revealing a bar. Observationally, M31 is indeed a barred spiral galaxy (Beaton et al. 2007), with the bar having a length of  $\approx 4$  kpc and a ratio of corotation radius/bar length of  $\mathcal{R} = 1.6 \pm 0.2$  (Blaña Díaz et al. 2018). Detailed analysis of the bar goes beyond the scope of this contribution, though bars have been considered before in MOND (Tiret & Combes 2007, 2008; Combes 2014). We refer the reader to Banik et al. (2020) for a recent study of the M33 bar in MOND, and to Sellwood et al. (2019) for a similar study in  $\Lambda$ CDM. A general study of galactic bars in CDM and different theoretical frameworks is available in Roshan et al. (2021a), which focused on





**Figure 13.** Top left: The stellar surface density map of the MW as viewed from along its observed disc spin vector, which is very similar to that in our POR simulation (Table 10). The positions are relative to the MW barycentre, found as described in Section 3.2. The dot-dashed circles show radii of 2, 4, and 8 kpc. Top right: The MW in the cylindrical  $rz$  projection (Banik et al. 2020). Bottom left: The stellar surface density profile of the MW (red curve), found by binning the stellar particles in cylindrical polar radius. The blue line shows an exponential profile whose scale length is an observationally motivated 3.6 kpc (Jurić et al. 2008). Bottom right: The initial (dot-dashed blue) and final (solid red) stellar  $\sigma_z$  of the MW disc as a function of radius. The data point at the Solar circle radius of 8.178 kpc (Gravity Collaboration 2019) shows  $\sigma_z = 40 \pm 10 \text{ km s}^{-1}$  based on results for different stellar samples in Yu & Liu (2018). Appendix E shows a version of this panel for a higher resolution simulation.

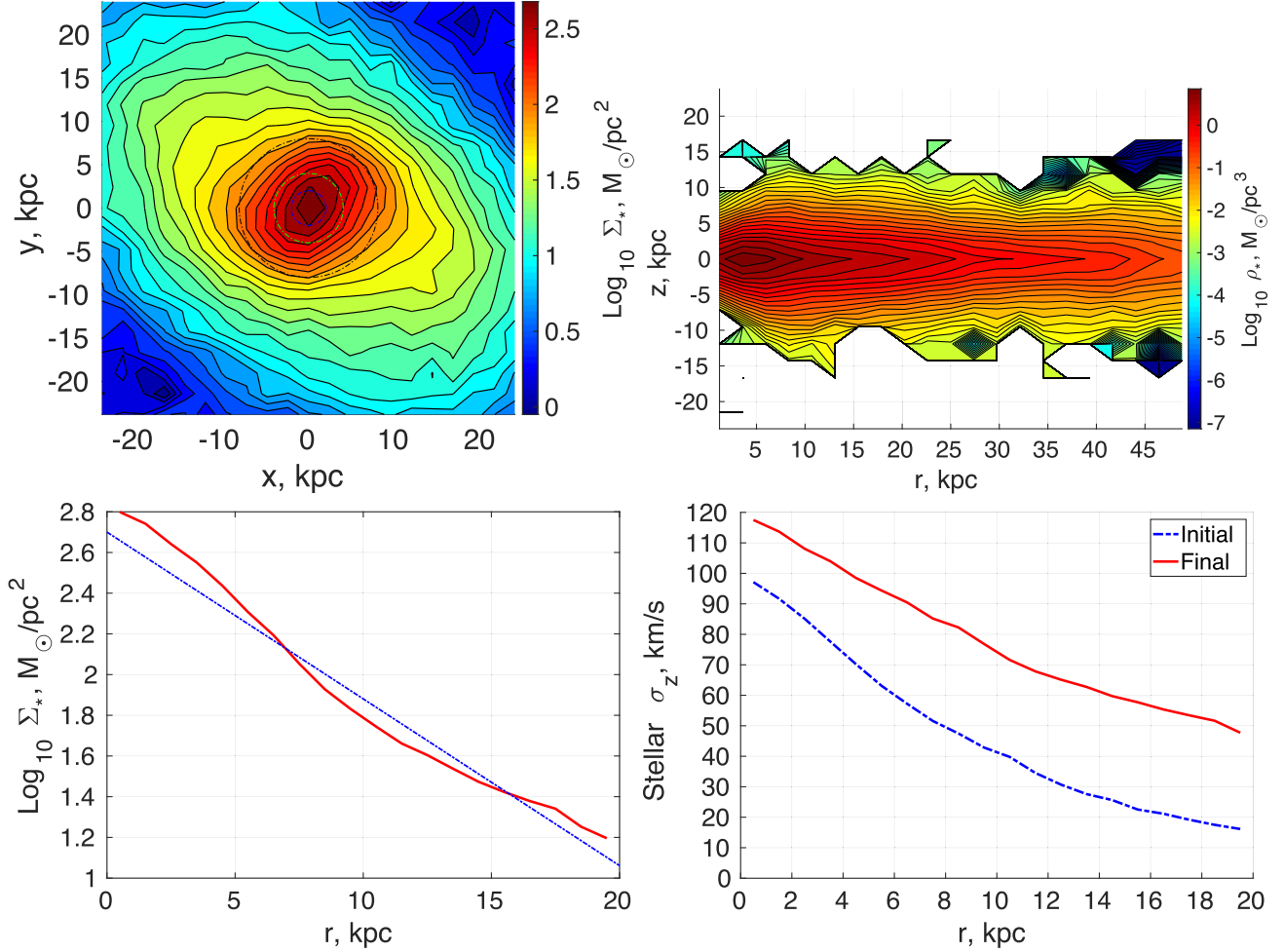
galaxies with a central surface density similar to the MW and M31.<sup>11</sup> Their work showed that in isolated MOND simulations, we would typically expect such a galaxy to have a bar with  $\mathcal{R} \approx 1\text{--}1.2$  (see their fig. 21), so the bar of M31 is somewhat slower but consistent within uncertainties.

The  $rz$  projection of M31 (top right panel of Fig. 14) indicates that our simulation retains a rather thin M31 disc. Its surface density profile is shown in the bottom left panel, which also shows a blue line representing an exponential law with an observationally motivated scale length of 5.3 kpc (Courteau et al. 2011). This provides a fairly good description of our simulated M31 disc remnant, especially when bearing in mind that their estimated scale length has an uncertainty of 0.5 kpc.

<sup>11</sup>In an acceleration-dependent theory like MOND, there is a critical surface density. Thus, galaxies with a similar central surface density should behave similarly, albeit with dimensional quantities like lengths appropriately scaled (see section 5.2 of Roshan et al. 2021a).

Finally, the most interesting result about the M31 remnant is shown in the bottom right panel of Fig. 14, which plots the initial and present simulated  $\sigma_z$  of the M31 stellar disc at different radii. It is dynamically much hotter than the MW disc (Fig. 13), which is consistent with the results of Collins et al. (2011). This tendency is weakened somewhat in a higher resolution model (Appendix E), but even then, the M31 disc is still hotter than the MW disc at all radii. Therefore, the M31 disc in our simulation is broadly similar to the real M31 disc.

While the higher  $\sigma_z$  in M31 is partly due to the initial conditions, the evolution of  $\sigma_z$  differs between the MW and M31. In the M31 case,  $\sigma_z$  increases by  $\approx 20 \text{ km s}^{-1}$  towards the centre and  $\approx 30 \text{ km s}^{-1}$  at 20 kpc, but in general, the whole  $\sigma_z(r)$  curve moves upwards by a similar amount at all radii. Even in the higher resolution simulation (Appendix E), the MW disc does not heat up very much at small radii ( $\sigma_z$  rises by  $\lesssim 10 \text{ km s}^{-1}$ ). The increase in  $\sigma_z$  is  $\approx 30 \text{ km s}^{-1}$  at large radii, which is similar to M31. Thus, the rise in  $\sigma_z$  at any fixed radius is larger for M31 over the entire disc. However, an increase of  $30 \text{ km s}^{-1}$  is a proportionately much larger increase for the outer



**Figure 14.** Similar to Fig. 13, but for M31. Top left: Face-on view of M31, showing a bar (unlike the simulated MW). Top right: Cylindrical  $rz$  projection of M31. Bottom left: The stellar surface density profile of M31 (red curve). The blue line shows an exponential law with an observationally motivated scale length of 5.3 kpc (Courteau et al. 2011). Bottom right: The initial (dot-dashed blue) and final (solid red) stellar  $\sigma_z$  of M31 as a function of radius, showing that it is expected to be dynamically hotter than the MW (c.f. Appendix E).

regions of the MW. These differences may be caused by the greater extent to which especially the outer MW disc is affected by the flyby, as evident in the larger disc-SP misalignment compared to M31 (Table 6).

### 3.4.3 Gas disc remnant surface density profiles

Our POR simulations do not include star formation, making it somewhat difficult to compare the gas disc remnants with observations. None the less, we use Fig. 15 to show the surface density profiles of the MW and M31 gas disc remnants, which we find using similar procedures to those used for the stellar disc remnants.

The outer gas disc of the MW is reasonably well fit with an exponential profile of scale length 7 kpc, which works well observationally (see table 1 of McMillan 2017). This is shown as a dot-dashed blue line in the top panel of Fig. 15 with an arbitrary normalization. It provides a reasonably good match to the simulated distribution, which is shown as a solid red line. Thus, our simulation seems to give a realistic MW gas disc surface density distribution at late times.

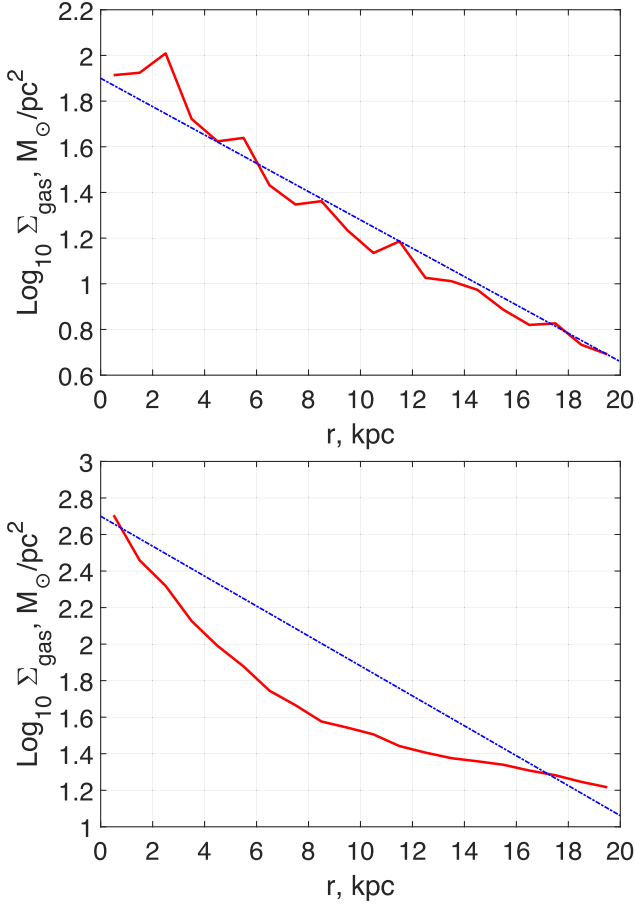
The results for M31 are shown in the bottom panel of Fig. 15, where the dot-dashed blue line shows a 5.3 kpc scale length exponential

profile with an arbitrary normalization. Our results show that this is not a good fit to the simulated gas disc, which cannot be well fit by a single exponential profile. However, we should bear in mind that a 5.3 kpc scale length is shown because it is a good fit to the observed M31 stellar disc (Courteau et al. 2011). The M31 gas disc has a rather irregular surface density profile (see fig. 16 of Chemin et al. 2009). Moreover, it is inevitable that the M31 gas disc is subject to significant additional processes in the MOND scenario since in addition to star formation, M31 seems to have experienced a more active interaction history (e.g. Fardal et al. 2013; D’Souza & Bell 2018).

## 4 DISCUSSION

### 4.1 The star formation histories of LG satellites

Our POR simulation indicates that the satellite regions of the MW and M31 should have been dominated by gas shortly after their flyby. Since SAM puts the flyby 7.2 Gyr ago (Fig. 3), we might expect members of the MW and M31 SPs to contain very few stars formed at earlier times. This can be altered somewhat, e.g. a higher Cen A mass could push the flyby earlier by  $\approx 1$  Gyr. None the less, it is difficult to



**Figure 15.** Top: Surface density profile of the simulated MW gas disc remnant 8.2 Gyr into our POR simulation (solid red curve). The dot-dashed blue line shows an exponential profile with 7 kpc scale length, which roughly describes the observed outer Galactic gas disc (see table 1 of McMillan 2017). Bottom: Similar to the top panel, but for M31. The dot-dashed blue line shows an exponential profile with a 5.3 kpc scale length. Observationally, the M31 gas disc has a rather irregular surface density profile (see fig. 16 of Chemin, Carignan & Foster 2009).

push the flyby much further back because this would require a much greater asymmetry between the MW–M31 orbit before and after the flyby (section 5.1.2 of BRZ18).<sup>12</sup> As a result, our flyby scenario can definitely not easily explain why a significant fraction of the stellar mass in the LMC seems to be older than 10 Gyr (Harris & Zaritsky 2009; Nidever et al. 2021). This is also apparent from detailed studies of Sculptor (de Boer et al. 2012), and of Galactic satellites more generally (Weisz et al. 2014).

The tidal tails expelled from the MW and M31 during the flyby would contain both stars and gas, so some stars from before the flyby are certainly to be expected in the TDGs. A larger proportion of stars from before the flyby could arise if gas accreted on to a newly formed TDG gets expelled by feedback. This would require more mass overall in the satellite region as the observed stellar masses of the LG satellites are known fairly well. Indeed, it was argued in section 5.2.4 of BRZ18 that even a small change to the initial

Galactic surface density profile could put much more mass in the satellite region with little effect on the overall potential. In this case, a lower star formation efficiency would be needed to explain the present mass in the Galactic SP members.

It is also possible that there are systematic uncertainties in the stellar ages of MW satellites. The age of the ancient globular cluster NGC 5904 has an uncertainty of 1 Gyr even though it is only 7.4 kpc away (Gontcharov, Mosenkov & Khovritchev 2019). More massive systems like the LMC have a wider range of stars that allow for more constraints on their age, including from RR Lyrae. However, these could be younger than generally thought if the gas cloud from which they formed contained a higher helium abundance than the primordial value (Savino et al. 2020). This seems likely to some extent as the flyby scenario involves TDGs forming out of gas already processed in the MW/M31, which would have been enriched in metals and helium prior to the flyby. In addition, considering binary stars generally leads to lower age estimates (Stanway & Eldridge 2018).

Thus, some combination of a slightly earlier flyby and reduced age estimates for the stellar populations in SP members could reduce the proportion of stars older than the flyby. The proportion might then become plausibly consistent with expectations, since after all there should be stars in a TDG that pre-date its formation from tidal debris. None the less, it is clear that the very ancient nature of the Galactic satellites is by far the biggest challenge to the particular flyby scenario studied here.

## 4.2 Evidence for the flyby beyond the SPs

In this section, we consider what the LG was like  $\approx 9$  Gyr ago and how this might relate to an MW–M31 encounter around that time.

### 4.2.1 Near the MW

A past encounter with M31 could have triggered the rapid formation of the Galactic bulge (Ballero et al. 2007) by means of tidal torques driving gas into the central regions of the MW. This may be related to the formation and subsequent buckling of the Galactic bar (Grady, Belokurov & Evans 2020), which could also have formed the similarly old thick disc (Kilic et al. 2017). Its pattern of enhanced  $\alpha$ -element abundances (Mashonkina et al. 2019) could indicate a very early flyby or a starburst triggered by the flyby (c.f. Renaud et al. 2016).

Looking to the satellite region, the young halo globular clusters (YH GCs) in the Galactic halo likely formed out of the inner part of the tidal debris cloud around the MW because they are spatially distributed similarly to the Galactic SP (Pawlowski et al. 2012).<sup>13</sup> Interestingly, the YH GCs have a bimodal age distribution (Mackey & Gilmore 2004). In addition to the ancient peak  $\approx 12$  Gyr ago, their fig. 9 shows a clear secondary peak  $\approx 9$  Gyr ago. One advantage of their study is that the ages are relative to that of M92, which should cancel some systematic uncertainties. To obtain absolute ages, those authors assumed that M92 is 12.55 Gyr old based on averaging two earlier estimates (Salaris & Weiss 2002). The more recent peak in the age distribution could be due to the flyby triggering the formation of stars and globular clusters (c.f. Renaud et al. 2016).

<sup>12</sup>The asymmetry could be caused by a higher past  $a_0$ , but section 5.2.3 of Haslbauer et al. (2020) argued against this using in particular the CMB, which would enter the MOND regime if  $a_0$  then exceeded its present value by  $\gtrsim 20\times$ .

<sup>13</sup>The old halo globular clusters have a smaller radial extent and are more nearly isotropically distributed than the YH GCs.

#### 4.2.2 The NGC 3109 association

Beyond the MW, another important line of evidence for a past encounter with M31 is provided by the kinematics of the NGC 3109 association  $\approx 1.5$  Mpc away (Pawlowski & McGaugh 2014). Those authors discussed how the high RV of NGC 3109 implies that it should have been close to the MW  $\approx 9$  Gyr ago based on looking at the problem backwards in time. However, their simplified dynamical analysis of the LG prevents one from drawing strong conclusions as they did not even consider gravity. Using a detailed 3D timing argument analysis of the LG and the major galaxies and galaxy groups outside it up to 8 Mpc away, Banik & Zhao (2017) showed that the RV of NGC 3109 is  $110 \text{ km s}^{-1}$  higher than in the best-fitting  $\Lambda$ CDM model.<sup>14</sup> This conclusion was confirmed in Banik & Zhao (2018c) using a much more thorough search for trajectories that match observational constraints (see their section 4.1).

One deficiency of these few-body timing argument analyses is that they do not consider the possibility of a galaxy passing close to the MW while it was undergoing an interaction, perhaps gaining energy from the interaction in a three-body process. Such backplash galaxies certainly exist in  $\Lambda$ CDM, but very rarely do they have properties resembling NGC 3109 (Banik et al. 2021). The problem is worsened considerably when bearing in mind the filamentary nature of the NGC 3109 association, which suggests that it was a gravitationally bound galaxy group in the past with a total mass of  $\approx 3.2 \times 10^{11} M_{\odot}$  (Bellazzini et al. 2013). A close approach to a major LG galaxy would have created significant dynamical friction that precludes escape to a large distance.

The backplash process would be much more efficient in MOND because the past high-velocity MW–M31 flyby it requires could have gravitationally slingshot galaxies out to quite large distances (Banik & Zhao 2018c). Those authors studied the backplash process in MOND in much more detail, with their fig. 6 showing that it is not necessary to very closely approach either the MW or M31. As a result, dynamical friction between the purely baryonic galaxies would be negligible and the disc of NGC 3109 might well have been preserved. However, it is likely that the NGC 3109 association as a whole would have become unbound, consistent with its unbound nature today (Kourkchi & Tully 2017). Therefore, the LG SPs are not the only anomalies faced by  $\Lambda$ CDM in the LG for which MOND has an explanation by means of a past MW–M31 flyby.

#### 4.3 Additional complications in the real LG

So far, we have focused on the MW and M31. Apart from their interaction 7.2 Gyr ago and the EFE from large-scale structure, they evolve more or less in isolation. This is of course not necessarily correct in the actual LG, where we expect additional satellite and non-satellite galaxies to play some role. Although many of the LG satellites should have formed out of tidal debris from the MW and M31 in a past encounter scenario, our POR simulation does not explicitly form individual TDGs, leaving instead a flattened gas distribution in each galaxy’s halo with a preferred orbital pole aligned similarly to the observed SP (Figs 7 and 8). If individual TDGs had formed, then these could perturb the disc of the parent galaxy in a way that is not captured by our simulations. Moreover, the MW and M31 have been subjected to various other perturbations that should be considered when studying the present LG.

In the M31 case, one important source of disturbance is M32, which is only 23 kpc from M31 (table 4 of Weisz et al. 2014). It is unclear whether M32 formed in the flyby, with BRZ18 arguing for a separate origin in their section 5.2.3. An origin unrelated to the MW–M31 flyby is very likely for the giant M31 southern stream as it formed in a much more recent galactic interaction (Fardal et al. 2013). It is also possible that M31 interacted with M33 at some point (Tepper-García, Bland-Hawthorn & Li 2020). M33 likely did not form as a TDG born out of the MW–M31 flyby because M33 lies outside the M31 SP (Ibata et al. 2013). We have implicitly assumed that M33 has been orbiting M31 since at least the start of our POR simulation and that it did not interfere with the MW–M31 flyby.

Turning to the MW, we know that the Galactic disc has been perturbed by the Sagittarius dwarf spheroidal satellite galaxy (Ibata, Gilmore & Irwin 1994), leaving behind the most prominent stream in the Galactic halo (see Thomas et al. 2017 in the MOND context). Since for every action there is an equal and opposite reaction, the Galactic disc has also been affected by Sagittarius (e.g. Laporte et al. 2018). Its orbital motion may have imprinted oscillations on the MW’s star formation history (Ruiz-Lara et al. 2020).

The most massive MW satellite is the LMC, which in principle should have formed out of tidal debris expelled from the MW during the interaction with M31 (Section 3.3.4). As our simulation does not form individual TDGs, it is possible that the formation of the LMC leads to additional subtle effects not captured in our model. The EFE from the LMC on the Galactic disc might be responsible for its warp (Brada & Milgrom 2000). In general, a Milgromian disc is expected to warp oppositely to the applied external field (see section 4.5.1 of Banik et al. 2020). The warping of the MW disc was also explored by Bílek et al. (2018) in  $N$ -body models of the MW–M31 flyby in MOND, showing that it could in principle explain the observed warp, though the orientation of the warp is likely to have changed substantially since the flyby (Poggio et al. 2020). There might also be other causes of disc warping.

The MW seems to have undergone a minor merger with a satellite known as Gaia-Sausage-Enceladus (GSE;  $M_{\star} \approx 5 \times 10^8 M_{\odot}$ ) around the time when the thick disc formed (Kilic et al. 2017). Stars and globular clusters from GSE were deposited in the Galactic halo with high radial anisotropy (Belokurov et al. 2018; Deason et al. 2018; Haywood et al. 2018; Helmi et al. 2018; Myeong et al. 2018). The timing of the merger is quite consistent with our estimate for the MW–M31 flyby, suggesting a connection. GSE might originally have been a satellite of M31. If instead it was a non-satellite dwarf in the LG, then the gravitational focusing effect of the combined MW+M31 gravity might have caused it to closely encounter the MW around the time of the flyby (Banik & Zhao 2018c). This could also be the case for other tentative halo structures potentially associated to galaxies accreted at roughly the same time, e.g. Sequoia (Myeong et al. 2019) and Thamnos (Koppelman et al. 2019).

In any case, it is clear that the history of the real LG has been more complex than in the simulations presented here. Those nevertheless display some generic features which are encouraging, especially with regards to the LG SPs but also the disc remnants and the overall MW–M31 orbit.

## 5 CONCLUSIONS

The existence of a thin plane of satellite galaxies around the MW is highly unexpected in  $\Lambda$ CDM cosmology (Kroupa et al. 2005; Pawlowski & Kroupa 2020), with the problem compounded by a similar SP around M31 (Ibata et al. 2013, 2014b; Sohn et al. 2020; Pawlowski & Tony Sohn 2021) that aligns fairly well with the MW SP

<sup>14</sup>Similar results were obtained by Peebles (2017) using a similar algorithm.



(Table 9; see also Pawlowski et al. 2013). In this contribution, we consider whether their SPs could have formed as TDGs born in a past flyby encounter between the MW and M31  $\approx 7$  Gyr ago. Such a flyby is required in MOND (Zhao et al. 2013) due to the rather strong MW–M31 mutual gravity acting on their nearly radial orbit (van der Marel et al. 2012, 2019; Salomon et al. 2021). For the first time, we simulate this flyby with hydrodynamic simulations by building on the earlier restricted  $N$ -body simulations of BRZ18 (gravity sourced by two point masses), where it was shown that the MW–M31 trajectory in MOND can plausibly be made consistent with the cosmological initial condition of little peculiar velocity at high redshift (the timing argument; Kahn & Woltjer 1959).

In this contribution, we conducted our hydrodynamical simulations of the flyby using POR, extending the  $N$ -body flyby models of Bílek et al. (2018). The MW and M31 disc templates were initialized similarly to Banik et al. (2020), which presented a fairly realistic MOND model of M33 that avoided some of the problems that arise when simulating it in Newtonian gravity with a live CDM halo (Sellwood et al. 2019). The barycentric position and velocity of the MW and M31 were initialized similarly to BRZ18. Although our POR simulation does not allow star formation and adopts a rather high gas temperature of 465 kK, it should be enough to determine the orbital pole distribution of the tidal debris around each galaxy. Our main goal was to check if they prefer a particular orbital pole, and if so, whether this aligns with the actually observed SP of the relevant galaxy. More detailed simulations are necessary to follow the formation of individual TDGs out of the tidal debris.

We advanced the MW and M31 for 8.2 Gyr, with  $\approx 1$  Gyr of this before the flyby. The timing argument mass of the LG in this model is  $3.457 \times 10^{11} M_{\odot}$  as estimated by SAM, which is slightly on the high side but still reasonable (Section 3.1). The galaxies separate to a large distance and retain rather thin discs, with very little material ending up outside both discs (Figs 1 and 2). This is because the relatively distant encounter compared to the disc scale lengths (Fig. 3) limits dynamical friction and tidal disruption during the flyby. The MW–M31 orbital geometry in our best-fitting model is consistent with the observed M31 PM (Section 3.2). This is a non-trivial success because it was not considered a priori when exploring the parameter space and because the latest constraints (Salomon et al. 2021) are tight enough that not all orbital geometries are allowed for fixed  $h$  (Fig. 5).

The tidal debris around the MW prefer a particular orbital pole, with the preferred direction aligning fairly well with the observed orbital pole of its SP (Fig. 7). The same is true for M31 (Fig. 8). Our model naturally yields a more concentrated distribution of orbital poles around M31, in line with observations (Pawlowski et al. 2013). These results are not much affected by the precise definition of the satellite region (Appendix D).

The encounter does not destroy the discs of the MW and M31, both of which retain a flattened disc by the end of the simulation (top right panels of Figs 13 and 14, respectively). The spin vectors of their discs are quite similar to observations (Table 10). In the flyby scenario, we expect that the Galactic thick disc formed by dynamical heating of a pre-existing stellar disc (c.f. the thickening of both discs in the  $N$ -body simulations of Bílek et al. 2018). Bearing in mind that the simulated Galactic stellar disc today corresponds to its observed thick disc due to the lack of star formation in our models, the simulated  $\Sigma_*$  profiles of the disc remnants are broadly consistent with an exponential law with a reasonable scale length (Section 3.4). This also applies to the Galactic gas disc remnant. Most importantly, our model indicates that the M31 disc should be dynamically hotter than the MW disc, as observed (e.g. Collins et al. 2011).

While our model is promising in many respects, a higher resolution simulation would be required to form individual TDGs and compare their properties with observations of LG satellites. For instance, it would be interesting to determine the mass–metallicity relation, which is expected to be rather similar to primordial satellites because the MW–M31 flyby was so long ago (Collins et al. 2015; Recchi et al. 2015) and because the tidal debris would mostly come from the outskirts of the MW and M31 discs (BRZ18). Any TDGs formed out of the MW–M31 flyby could have undergone significant enrichment at later times, especially given their strong self-gravity in MOND. This is related to the star formation histories of the TDGs, which are potentially problematic for our model (Weisz et al. 2014) because we would in general expect only a small fraction of their stars to be substantially older than the flyby. A significant fraction of the stellar mass in the LMC seems to be older than plausible estimates for when the flyby occurred (Harris & Zaritsky 2009; Nidever et al. 2021). The solution might lie in a slightly earlier flyby combined with a low star formation efficiency in the tidal tails, increasing the relative importance of stars drawn from the parent galaxy (Section 4.1).

Although our study focused on the LG SPs, a similar scenario could be responsible for the Cen A SP (Müller et al. 2018a, 2021). This might have formed due to a past M83 flyby, or it could be related to the major merger that Cen A likely experienced  $\approx 2$  Gyr ago (Wang et al. 2020). In the latter case, the SP members should be more metal-rich (Duc et al. 2014), though this need not be the case in the LG due to the much more ancient interaction (Recchi et al. 2015). Forming an SP out of a merger in principle allows a much wider range of initial orbital configurations because the merger geometry is not constrained independently of the SP. There is also only one SP to match rather than the two in the LG. As a result, similar modelling attempts around Cen A are likely underdetermined unless additional observational constraints become available. Alternatively, the Cen A SP might be related to a past interaction with M83, whose satellite system also appears to be flattened (Müller, Rejkuba & Jerjen 2018b). One difficulty is that distance uncertainties make it challenging to identify a distant plane of satellites viewed close to face-on. Further observations are needed to clarify the situation for M83 and to better characterize the SP around Cen A, whose properties are less well known than for the SPs in the LG.

The LG SPs (reviewed in Pawlowski 2018, 2021) pose a severe challenge to  $\Lambda$ CDM due to the combination of their extreme flatness and coherent rotation (suggestive of dissipation), their mutual alignment within  $37^\circ$  (suggestive of a common origin), and the high  $\sigma_{\text{int}}$  of their member satellites (ruling out that they are TDGs obeying Newtonian gravity). The issue is not much affected by baryonic physics (Pawlowski & Kroupa 2020; Samuel et al. 2021), suggesting a fundamental problem. We therefore considered a MOND model of the LG (for a discussion of larger scale issues, see e.g. section 3.1 of Haslbauer et al. 2020). We showed here that the past close MW–M31 flyby required by this framework (Zhao et al. 2013) leads to the formation of tidal debris around each galaxy, with a preferred orbital pole in each case. This direction aligns with that of the observed SP for both the MW and M31. The two galaxies reach a realistic post-flyby separation with a reasonable relative PM in accordance with the latest constraints. The MW and M31 also retain thin discs with realistic properties by the end of the simulation. Therefore, it might be possible to explain their SPs in a Milgromian framework while remaining consistent with other constraints, the only truly significant challenge for the particular flyby scenario studied here being the ages of the stellar populations within LG satellite galaxies.

## ACKNOWLEDGEMENTS

IB is supported by Science and Technology Facilities Council grant ST/V000861/1. He acknowledges support from a ‘Pathways to Research’ fellowship from the University of Bonn in 2021 after an Alexander von Humboldt Foundation postdoctoral research fellowship (2018–2020). IT acknowledges support through the Stellar Populations and Dynamics research group at the University of Bonn. GC acknowledges support from Chile’s National Fund for Scientific and Technological Development (FONDECYT) Regular No. 1181708. BF and RI acknowledge funding from the Agence Nationale de la Recherche (ANR project ANR-18-CE31-0006 and ANR-19-CE31-0017) and from the European Research Council (ERC) under the European Union’s Horizon 2020 Framework Programme (grant agreement number 834148). MSP was supported by the Leibniz-Junior Research Group grant J94/2020 via the Leibniz Competition, and a Klaus Tschira Boost Fund provided by the Klaus Tschira Stiftung and the German Scholars Organization. IB is grateful to Hongsheng Zhao for supervising a masters project in which RT adapted DICE to prepare a MOND disc template algorithm that forms the basis for this and other work. The authors are grateful to the anonymous referee for comments which helped to improve this publication.

## DATA AVAILABILITY

The data underlying this article are available in the article. The algorithms used to prepare and run POR simulations of disc galaxies and to extract their results into human-readable form are discussed further in Banik et al. (2020) and are publicly available.<sup>15</sup> A user manual describing the operation of these algorithms is also available (Nagesh et al. 2021).

## REFERENCES

- Ahmed S. H., Brooks A. M., Christensen C. R., 2017, *MNRAS*, 466, 3119
- Aiola S. et al., 2020, *J. Cosmol. Astropart. Phys.*, 2020, 047
- Algorry D. G. et al., 2017, *MNRAS*, 469, 1054
- Alves D. R., Nelson C. A., 2000, *ApJ*, 542, 789
- Angus G. W., 2009, *MNRAS*, 394, 527
- Angus G. W., 2010, *J. Cosmol. Astropart. Phys.*, 2010, 026
- Angus G. W., Famaey B., Diaferio A., 2010, *MNRAS*, 402, 395
- Angus G. W., Diaferio A., Kroupa P., 2011, *MNRAS*, 416, 1401
- Angus G. W., Diaferio A., Famaey B., van der Heyden K. J., 2013, *MNRAS*, 436, 202
- Asencio E., Banik I., Kroupa P., 2021, *MNRAS*, 500, 5249
- Athanassoula E., 2002, *ApJ*, 569, L83
- Ballero S. K., Matteucci F., Origlia L., Rich R. M., 2007, *A&A*, 467, 123
- Banik I., Zhao H., 2016, *MNRAS*, 459, 2237
- Banik I., Zhao H., 2017, *MNRAS*, 467, 2180
- Banik I., Zhao H., 2018a, *SciFed J. Astrophys.*, 1, 1000008
- Banik I., Zhao H., 2018b, *MNRAS*, 473, 419
- Banik I., Zhao H., 2018c, *MNRAS*, 473, 4033
- Banik I., Zhao H., 2018d, *MNRAS*, 480, 2660
- Banik I., Zhao H., 2022, preprint ([arXiv:2110.06936](https://arxiv.org/abs/2110.06936))
- Banik I., O’Ryan D., Zhao H., 2018, *MNRAS*, 477, 4768 (BRZ18)
- Banik I., Thies I., Famaey B., Candlish G., Kroupa P., Ibata R., 2020, *ApJ*, 905, 135
- Banik I., Haslbauer M., Pawlowski M. S., Famaey B., Kroupa P., 2021, *MNRAS*, 503, 6170
- Barnes J. E., Hernquist L., 1992, *Nature*, 360, 715
- Beaton R. L. et al., 2007, *ApJ*, 658, L91
- Begeman K. G., Broeils A. H., Sanders R. H., 1991, *MNRAS*, 249, 523
- Bekenstein J., Milgrom M., 1984, *ApJ*, 286, 7
- Bellazzini M., Oosterloo T., Fraternali F., Beccari G., 2013, *A&A*, 559, L11
- Belokurov V., Erkal D., Evans N. W., Koposov S. E., Deason A. J., 2018, *MNRAS*, 478, 611
- Besla G., Kallivayalil N., Hernquist L., Robertson B., Cox T. J., van der Marel R. P., Alcock C., 2007, *ApJ*, 668, 949
- Bilek M., Thies I., Kroupa P., Famaey B., 2018, *A&A*, 614, A59
- Bilek M., Müller O., Famaey B., 2019, *A&A*, 627, L1
- Bilek M., Thies I., Kroupa P., Famaey B., 2021, *Galaxies*, 9, 100
- Blaña Díaz M. et al., 2018, *MNRAS*, 481, 3210
- Bournaud F., Duc P. A., Amram P., Combes F., Gach J. L., 2004, *A&A*, 425, 813
- Bovy J., Rix H.-W., 2013, *ApJ*, 779, 115
- Brada R., Milgrom M., 2000, *ApJ*, 531, L21
- Brimouille F., Seitz S., Lerchster M., Bender R., Snigula J., 2013, *MNRAS*, 432, 1046
- Brouwer M. M. et al., 2017, *MNRAS*, 466, 2547
- Brouwer M. M. et al., 2021, *A&A*, 650, A113
- Buch J., Leung J. S. C., Fan J., 2019, *J. Cosmol. Astropart. Phys.*, 2019, 026
- Candlish G. N., 2016, *MNRAS*, 460, 2571
- Casas R. A., Arias V., Peña Ramírez K., Kroupa P., 2012, *MNRAS*, 424, 1941
- Cautun M., Frenk C. S., 2017, *MNRAS*, 468, L41
- Chae K.-H., Bernardi M., Domínguez Sánchez H., Sheth R. K., 2020a, *ApJ*, 903, L31
- Chae K.-H., Lelli F., Desmond H., McGaugh S. S., Li P., Schombert J. M., 2020b, *ApJ*, 904, 51
- Chae K.-H., Desmond H., Lelli F., McGaugh S. S., Schombert J. M., 2021, *ApJ*, 921, 104
- Chemin L., Carignan C., Foster T., 2009, *ApJ*, 705, 1395
- Collins M. L. M. et al., 2011, *MNRAS*, 413, 1548
- Collins M. L. M. et al., 2015, *ApJ*, 799, L13
- Combes F., 2014, *A&A*, 571, A82
- Conn A. R. et al., 2013, *ApJ*, 766, 120
- Conroy C., Naidu R. P., Garavito-Camargo N., Besla G., Zaritsky D., Bonaca A., Johnson B. D., 2021, *Nature*, 592, 534
- Corbelli E., Lorenzoni S., Walterbos R., Braun R., Thilker D., 2010, *A&A*, 511, A89
- Correa Magnus L., Vasiliev E., 2022, *MNRAS*, 511, 2610
- Courteau S., Widrow L. M., McDonald M., Guhathakurta P., Gilbert K. M., Zhu Y., Beaton R. L., Majewski S. R., 2011, *ApJ*, 739, 20
- D’Souza R., Bell E. F., 2018, *Nat. Astron.*, 2, 737
- Deason A. J., Belokurov V., Koposov S. E., Lancaster L., 2018, *ApJ*, 862, L1
- Debatista V. P., Sellwood J. A., 2000, *ApJ*, 543, 704
- de Boer T. J. L. et al., 2012, *A&A*, 539, A103
- Di Valentino E., 2021, *MNRAS*, 502, 2065
- Duc P.-A., Paudel S., McDermid R. M., Cuillandre J.-C., Serra P., Bournaud F., Cappellari M., Emsellem E., 2014, *MNRAS*, 440, 1458
- Efstathiou G., Sutherland W. J., Maddox S. J., 1990, *Nature*, 348, 705
- Famaey B., Binney J., 2005, *MNRAS*, 363, 603
- Famaey B., McGaugh S. S., 2012, *Living Rev. Relativ.*, 15, 10
- Famaey B., McGaugh S., Milgrom M., 2018, *MNRAS*, 480, 473
- Fardal M. A. et al., 2013, *MNRAS*, 434, 2779
- Fattahi A., Navarro J. F., Frenk C. S., Oman K. A., Sawala T., Schaller M., 2018, *MNRAS*, 476, 3816
- Fletcher R., Powell M. J. D., 1963, *Comput. J.*, 6, 163
- Fragkoudi F., Grand R. J. J., Pakmor R., Springel V., White S. D. M., Marinacci F., Gomez F. A., Navarro J. F., 2021, *A&A*, 650, L16
- Francis C., Anderson E., 2014, *Celest. Mech. Dyn. Astron.*, 118, 399
- Freeman K. C., 1970, *ApJ*, 160, 811
- Freundlich J., Famaey B., Oria P.-A., Bilek M., Müller O., Ibata R., 2022, *A&A*, 658, A26
- Gaia Collaboration, 2018, *A&A*, 616, A12
- Garavito-Camargo N., Patel E., Besla G., Price-Whelan A. M., Gómez F. A., Laporte C. F. P., Johnston K. V., 2021, *ApJ*, 923, 140
- Gentile G., Famaey B., de Blok W. J. G., 2011, *A&A*, 527, A76
- Genzel R. et al., 2017, *Nature*, 543, 397

<sup>15</sup><https://bitbucket.org/SrikanthTN/bonnpor/src/master/>

This includes a version of RDRAMSES rated to work in parallel.

- Genzel R. et al., 2020, *ApJ*, 902, 98
- Gilmore G., Reid N., 1983, *MNRAS*, 202, 1025
- Gontcharov G. A., Mosenkov A. V., Khovritchev M. Y., 2019, *MNRAS*, 483, 4949
- Grady J., Belokurov V., Evans N. W., 2020, *MNRAS*, 492, 3128
- Gravity Collaboration, 2019, *A&A*, 625, L10
- Haghi H., Bazkiaei A. E., Zonoozi A. H., Kroupa P., 2016, *MNRAS*, 458, 4172
- Haghi H. et al., 2019a, *MNRAS*, 487, 2441
- Haghi H., Amiri V., Hasani Zonoozi A., Banik I., Kroupa P., Haslbauer M., 2019b, *ApJ*, 884, L25
- 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
- Hammer F., Yang Y. B., Flores H., Puech M., Fouquet S., 2015, *ApJ*, 813, 110
- Hammer F., Wang J., Pawlowski M. S., Yang Y., Bonifacio P., Li H., Babusiaux C., Arenou F., 2021, *ApJ*, 922, 93
- Harris J., Zaritsky D., 2009, *AJ*, 138, 1243
- Harrison C. M. et al., 2017, *MNRAS*, 467, 1965
- Haslbauer M., Dabringhausen J., Kroupa P., Javanmardi B., Banik I., 2019, *A&A*, 626, A47
- Haslbauer M., Banik I., Kroupa P., 2020, *MNRAS*, 499, 2845
- Haslbauer M., Banik I., Kroupa P., Wittenburg N., Javanmardi B., 2022, *ApJ*, 925, 183
- Haywood M., Di Matteo P., Lehnert M. D., Snaith O., Khoperskov S., Gómez A., 2018, *ApJ*, 863, 113
- Hees A., Folkner W. M., Jacobson R. A., Park R. S., 2014, *Phys. Rev. D*, 89, 102002
- Hees A., Famaey B., Angus G. W., Gentile G., 2016, *MNRAS*, 455, 449
- Helmi A., Babusiaux C., Koppelman H. H., Massari D., Veljanoski J., Brown A. G. A., 2018, *Nature*, 563, 85
- Ibata R. A., Gilmore G., Irwin M. J., 1994, *Nature*, 370, 194
- Ibata R. A. et al., 2013, *Nature*, 493, 62
- Ibata R. A. et al., 2014a, *ApJ*, 780, 128
- Ibata R. A., Ibata N. G., Lewis G. F., Martin N. F., Conn A., Elahi P., Arias V., Fernando N., 2014b, *ApJ*, 784, L6
- Jayaraman A., Gilmore G., Wyse R. F. G., Norris J. E., Belokurov V., 2013, *MNRAS*, 431, 930
- Jurić M. et al., 2008, *ApJ*, 673, 864
- Kahn F. D., Woltjer L., 1959, *ApJ*, 130, 705
- Kallivayalil N., van der Marel R. P., Besla G., Anderson J., Alcock C., 2013, *ApJ*, 764, 161
- Katz H., McGaugh S., Teuben P., Angus G. W., 2013, *ApJ*, 772, 10
- Keenan R. C., Barger A. J., Cowie L. L., 2013, *ApJ*, 775, 62
- Kilic M., Munn J. A., Harris H. C., von Hippel T., Liebert J. W., Williams K. A., Jeffery E., DeGennaro S., 2017, *ApJ*, 837, 162
- Klessen R. S., Kroupa P., 1998, *ApJ*, 498, 143
- Kogut A. et al., 1993, *ApJ*, 419, 1
- Koppelman H. H., Helmi A., Massari D., Price-Whelan A. M., Starkenburg T. K., 2019, *A&A*, 631, L9
- Kourkchi E., Tully R. B., 2017, *ApJ*, 843, 16
- Kroupa P., 1997, *New Astron.*, 2, 139
- Kroupa P., 2015, *Can. J. Phys.*, 93, 169
- Kroupa P., Theis C., Boily C. M., 2005, *A&A*, 431, 517
- Kroupa P. et al., 2018a, *Nat. Astron.*, 2, 925
- Kroupa P. et al., 2018b, *Nature*, 561, E4
- Kunkel W. E., Demers S., 1976, in *The Galaxy and the Local Group*. Royal Greenwich Observatory Bulletins, London, UK, p. 241
- Laporte C. F. P., Johnston K. V., Gómez F. A., Garavito-Camargo N., Besla G., 2018, *MNRAS*, 481, 286
- Lelli F., McGaugh S. S., Schombert J. M., Pawlowski M. S., 2017, *ApJ*, 836, 152
- Lelli F., de Breuck C., Falkendal T., Fraternali F., Man A. W. S., Nesvadba N. P. H., Lehnert M. D., 2018, *MNRAS*, 479, 5440
- Lelli F., McGaugh S. S., Schombert J. M., Desmond H., Katz H., 2019, *MNRAS*, 484, 3267
- Lelli F., Di Teodoro E. M., Fraternali F., Man A. W. S., Zhang Z.-Y., de Breuck C., Davis T. A., Maiolino R., 2021, *Science*, 371, 713
- Li C., Zhao G., 2017, *ApJ*, 850, 25
- Li P., Lelli F., McGaugh S., Schombert J., 2018, *A&A*, 615, A3
- Li H., Hammer F., Babusiaux C., Pawlowski M. S., Yang Y., Arenou F., Du C., Wang J., 2021, *ApJ*, 916, 8
- Lucchini S., D'Onghia E., Fox A. J., Bustard C., Bland-Hawthorn J., Zweibel E., 2020, *Nature*, 585, 203
- Lucchini S., D'Onghia E., Fox A. J., 2021, *ApJ*, 921, L36
- Lüghausen F., Famaey B., Kroupa P., 2015, *Can. J. Phys.*, 93, 232
- Lynden-Bell D., 1976, *MNRAS*, 174, 695
- Lynden-Bell D., 1982, *The Observatory*, 102, 202
- Ma C. et al., 1998, *AJ*, 116, 516
- Mackey A. D., Gilmore G. F., 2004, *MNRAS*, 355, 504
- McConnachie A. W., 2012, *AJ*, 144, 4
- McConnachie A. W. et al., 2018, *ApJ*, 868, 55
- McGaugh S. S., 2012, *AJ*, 143, 40
- McGaugh S. S., 2021, *Stud. Hist. Phil. Sci.*, 88, 220
- McGaugh S., Milgrom M., 2013a, *ApJ*, 766, 22
- McGaugh S., Milgrom M., 2013b, *ApJ*, 775, 139
- McGaugh S. S., Wolf J., 2010, *ApJ*, 722, 248
- McGaugh S. S., Lelli F., Schombert J. M., 2016, *Phys. Rev. Lett.*, 117, 201101
- McGaugh S. S., Lelli F., Schombert J. M., Li P., Visgaitis T., Parker K. S., Pawlowski M. S., 2021, *AJ*, 162, 202
- McMillan P. J., 2011, *MNRAS*, 414, 2446
- McMillan P. J., 2017, *MNRAS*, 465, 76
- Mashonkina L. I., Neretina M. D., Sitnova T. M., Pakhomov Y. V., 2019, *Astron. Rep.*, 63, 726
- Mateu C., Vivas A. K., 2018, *MNRAS*, 479, 211
- Metz M., Kroupa P., Jerjen H., 2007, *MNRAS*, 374, 1125
- Metz M., Kroupa P., Libeskind N. I., 2008, *ApJ*, 680, 287
- Metz M., Kroupa P., Jerjen H., 2009a, *MNRAS*, 394, 2223
- Metz M., Kroupa P., Theis C., Hensler G., Jerjen H., 2009b, *ApJ*, 697, 269
- Milgrom M., 1983, *ApJ*, 270, 365
- Milgrom M., 1986, *ApJ*, 302, 617
- Milgrom M., 2010, *MNRAS*, 403, 886
- Milgrom M., 2012, *Phys. Rev. Lett.*, 109, 131101
- Milgrom M., 2013, *Phys. Rev. Lett.*, 111, 041105
- Milgrom M., 2015, *Phys. Rev. D*, 91, 044009
- Milgrom M., 2017, preprint ([arXiv:1703.06110](https://arxiv.org/abs/1703.06110))
- Mirabel I. F., Dottori H., Lutz D., 1992, *A&A*, 256, L19
- Müller O., Jerjen H., Pawlowski M. S., Binggeli B., 2016, *A&A*, 595, A119
- Müller O., Pawlowski M. S., Jerjen H., Lelli F., 2018a, *Science*, 359, 534
- Müller O., Rejkuba M., Jerjen H., 2018b, *A&A*, 615, A96
- Müller O. et al., 2021, *A&A*, 645, L5
- Myeong G. C., Evans N. W., Belokurov V., Sanders J. L., Koposov S. E., 2018, *ApJ*, 863, L28
- Myeong G. C., Vasiliev E., Iorio G., Evans N. W., Belokurov V., 2019, *MNRAS*, 488, 1235
- Nagesh S. T., Banik I., Thies I., Kroupa P., Famaey B., Wittenburg N., Parziale R., Haslbauer M., 2021, *Can. J. Phys.*, 99, 607
- Neeleman M., Prochaska J. X., Kanekar N., Rafelski M., 2020, *Nature*, 581, 269
- Nidever D. L. et al., 2021, *AJ*, 161, 74
- Oria P. A. et al., 2021, *ApJ*, 923, 68
- Ostriker J. P., Steinhardt P. J., 1995, *Nature*, 377, 600
- Pawlowski M. S., 2018, *Mod. Phys. Lett. A*, 33, 1830004
- Pawlowski M. S., 2021, *Galaxies*, 9, 66
- Pawlowski M. S., Kroupa P., 2013, *MNRAS*, 435, 2116
- Pawlowski M. S., Kroupa P., 2020, *MNRAS*, 491, 3042
- Pawlowski M. S., McGaugh S. S., 2014, *MNRAS*, 440, 908
- Pawlowski M. S., Tony Sohn S., 2021, *ApJ*, 923, 42
- Pawlowski M. S., Kroupa P., de Boer K. S., 2011, *A&A*, 532, A118
- Pawlowski M. S., Pflamm-Altenburg J., Kroupa P., 2012, *MNRAS*, 423, 1109
- Pawlowski M. S., Kroupa P., Jerjen H., 2013, *MNRAS*, 435, 1928
- Pawlowski M. S. et al., 2014, *MNRAS*, 442, 2362
- Pawlowski M. S., McGaugh S. S., Jerjen H., 2015, *MNRAS*, 453, 1047



- Pawlowski M. S. et al., 2017, *Astron. Nachr.*, 338, 854
- Pawlowski M. S., Oria P.-A., Taibi S., Famaey B., Ibata R., 2021, preprint (arXiv:2111.05358)
- Peebles P. J. E., 2017, preprint (arXiv:1705.10683)
- Peebles P. J. E., 2020, *MNRAS*, 498, 4386
- Peebles P. J. E., Nusser A., 2010, *Nature*, 465, 565
- Perret V., Renaud F., Epinat B., Amram P., Bournaud F., Contini T., Teyssier R., Lambert J.-C., 2014, *A&A*, 562, A1
- Peschken N., Łokas E. L., 2019, *MNRAS*, 483, 2721
- Pietrzyński G. et al., 2013, *Nature*, 495, 76
- Planck Collaboration VI, 2020, *A&A*, 641, A6
- Planck Collaboration XVI, 2014, *A&A*, 571, A16
- Poggio E., Drimmel R., Andrae R., Bailer-Jones C. A. L., Fouesneau M., Lattanzi M. G., Smart R. L., Spagna A., 2020, *Nat. Astron.*, 4, 590
- Recchi S., Kroupa P., Ploekinger S., 2015, *MNRAS*, 450, 2367
- Renaud F., Famaey B., Kroupa P., 2016, *MNRAS*, 463, 3637
- Riess A. G., 2020, *Nat. Rev. Phys.*, 2, 10
- Riess A. G., Casertano S., Yuan W., Bowers J. B., Macri L., Zinn J. C., Scolnic D., 2021a, *ApJ*, 908, L6
- Riess A. G. et al., 2021b, preprint (arXiv:2112.04510)
- Riley A. H. et al., 2019, *MNRAS*, 486, 2679
- Rizzo F., Vegetti S., Powell D., Fraternali F., McKean J. P., Stacey H. R., White S. D. M., 2020, *Nature*, 584, 201
- Robotham A. S. G. et al., 2012, *MNRAS*, 424, 1448
- Roshan M., Banik I., Ghafourian N., Thies I., Famaey B., Asencio E., Kroupa P., 2021a, *MNRAS*, 503, 2833
- Roshan M., Ghafourian N., Kashfi T., Banik I., Haslbauer M., Cuomo V., Famaey B., Kroupa P., 2021b, *MNRAS*, 508, 926
- Ruiz-Lara T., Gallart C., Bernard E. J., Cassisi S., 2020, *Nat. Astron.*, 4, 965
- Salaris M., Weiss A., 2002, *A&A*, 388, 492
- Salomon J. B., Ibata R., Reylé C., Famaey B., Libeskind N. I., McConnachie A. W., Hoffman Y., 2021, *MNRAS*, 507, 2592
- Samuel J., Wetzel A., Chapman S., Tollerud E., Hopkins P. F., Boylan-Kolchin M., Bailin J., Faucher-Giguère C.-A., 2021, *MNRAS*, 504, 1379
- Sanders R. H., 1998, *MNRAS*, 296, 1009
- Sanders R. H., 2019, *MNRAS*, 485, 513
- Santos-Santos I. M., Domínguez-Tenreiro R., Pawlowski M. S., 2020, *MNRAS*, 499, 3755
- Santos-Santos I. M. E., Fattahi A., Sales L. V., Navarro J. F., 2021, *MNRAS*, 504, 4551
- Savino A., Koch A., Prudil Z., Kunder A., Smolec R., 2020, *A&A*, 641, A96
- Schrabback T. et al., 2021, *A&A*, 646, A73
- Sellwood J. A., Shen J., Li Z., 2019, *MNRAS*, 486, 4710
- Shao S., Cautun M., Frenk C. S., Grand R. J. J., Gómez F. A., Marinacci F., Simpson C. M., 2018, *MNRAS*, 476, 1796
- Sharma G., Salucci P., Harrison C. M., van de Ven G., Lapi A., 2021, *MNRAS*, 503, 1753
- Shelest A., Lelli F., 2020, *A&A*, 641, A31
- Skordis C., Złotnik T., 2019, *Phys. Rev. D*, 100, 104013
- Skordis C., Złotnik T., 2021, *Phys. Rev. Lett.*, 127, 161302
- Snaith O., Haywood M., Di Matteo P., Lehnert M. D., Combes F., Katz D., Gómez A., 2015, *A&A*, 578, A87
- Sofue Y., 2015, *PASJ*, 67, 75
- Sohn S. T., Patel E., Fardal M. A., Besla G., van der Marel R. P., Geha M., Guhathakurta P., 2020, *ApJ*, 901, 43
- Stanway E. R., Eldridge J. J., 2018, *MNRAS*, 479, 75
- Stott J. P. et al., 2016, *MNRAS*, 457, 1888
- Tepper-García T., Bland-Hawthorn J., Li D., 2020, *MNRAS*, 493, 5636
- Teyssier R., 2002, *A&A*, 385, 337
- Teyssier R., Chapon D., Bournaud F., 2010, *ApJ*, 720, L149
- Thomas G. F., Famaey B., Ibata R., Lüghausen F., Kroupa P., 2017, *A&A*, 603, A65
- Tiret O., Combes F., 2007, *A&A*, 464, 517
- Tiret O., Combes F., 2008, *A&A*, 483, 719
- Toomre A., 1964, *ApJ*, 139, 1217
- Tully R. B., Libeskind N. I., Karachentsev I. D., Karachentseva V. E., Rizzi L., Shaya E. J., 2015, *ApJ*, 802, L25
- van der Marel R. P., Kallivayalil N., 2014, *ApJ*, 781, 121
- van der Marel R. P., Fardal M., Besla G., Beaton R. L., Sohn S. T., Anderson J., Brown T., Guhathakurta P., 2012, *ApJ*, 753, 8
- van der Marel R. P., Fardal M. A., Sohn S. T., Patel E., Besla G., del Pino A., Sahlmann J., Watkins L. L., 2019, *ApJ*, 872, 24
- Vasiliev E., 2018, *MNRAS*, 481, L100
- Wang J., Hammer F., Rejkuba M., Crnojević D., Yang Y., 2020, *MNRAS*, 498, 2766
- Weisz D. R., Dolphin A. E., Skillman E. D., Holtzman J., Gilbert K. M., Dalcanton J. J., Williams B. F., 2014, *ApJ*, 789, 147
- Wetzstein M., Naab T., Burkert A., 2007, *MNRAS*, 375, 805
- Wittenburg N., Kroupa P., Famaey B., 2020, *ApJ*, 890, 173
- Yang Y., Hammer F., Jiao Y., Pawlowski M., 2022, *MNRAS*, 512, 4171
- Yu J., Liu C., 2018, *MNRAS*, 475, 1093
- Zhao H., Li B., Bienaymé O., 2010, *Phys. Rev. D*, 82, 103001
- Zhao H., Famaey B., Lüghausen F., Kroupa P., 2013, *A&A*, 557, L3

## APPENDIX A: SETTING UP THE PARTICLE MASSES IN THE SIMULATION

An important new aspect of the present POR simulations is that we need to consider the outer parts of the simulated discs in much greater detail, since we expect these regions to be the original source material for the SPs (BRZ18). Material at an initial galactocentric distance of  $\approx 50$  kpc would be very poorly resolved with a feasible number of equal mass particles, so we need to vary the particle masses in our disc templates. This differs from previous POR simulations (e.g. Banik et al. 2020). We therefore adjust the DICE algorithm as described below.

To maintain an exponential surface density profile for each component, we require that the particle mass  $m$  at some location is

$$m(x) \propto \frac{x \exp(-x)}{P(x)}, \quad (\text{A1})$$

where  $P dx$  is the probability of a randomly chosen particle having some particular  $x \pm dx/2$ , with  $x$  being the galactocentric radius in units of the exponential scale length for the considered component. For different regions to be resolved similarly well despite the exponential profile,  $P$  should be roughly constant.  $P$  can theoretically have any non-zero value at any  $x$  in the range we wish to model, because this can always be compensated by adjusting  $m$  so as to satisfy equation (A1) and maintain the surface density profile. We found that setting  $P$  to a constant everywhere leads to disastrous numerical effects at low  $x$  because here we would need  $m \rightarrow 0$ . Since the central regions of a disc are most unstable to self-gravity, it is essential to minimize artificial relaxation effects here. This can be done if all particles have a similar mass in the central few scale lengths. Our compromise solution is

$$P \propto \tanh(\alpha x). \quad (\text{A2})$$

Setting the parameter  $\alpha = 3$  keeps  $m$  nearly constant over as large a range of  $x$  as possible.

We assign each particle a scaled radius  $x$  using a Monte Carlo technique to realize the distribution

$$P(x) = \frac{\alpha \tanh(\alpha x)}{\ln \cosh(\alpha x_{\max})}, \quad (\text{A3})$$

where  $x_{\max} \approx 25$  is the maximum normalized extent of the considered component. The azimuthal angle is drawn randomly from the

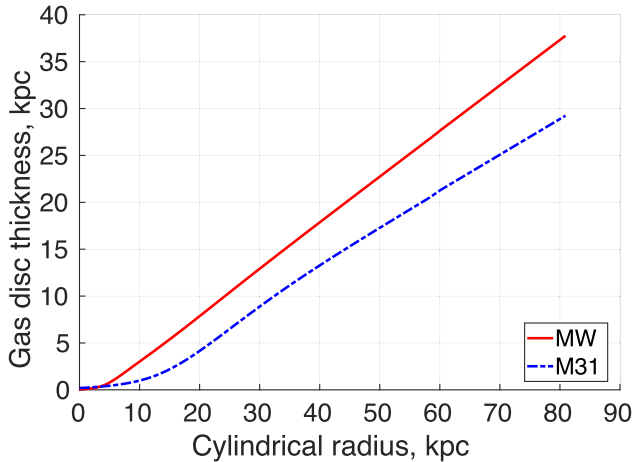


range  $(0-2\pi)$ , while vertically we use a  $\text{sech}^2$  profile. Both the radial and the vertical probability distributions can be integrated analytically, simplifying the code. Particle masses are then assigned using equation (A1). To ensure the correct normalization, we multiply the masses of all particles in the considered component by a common factor. These adjustments to our modified DICE are enabled by setting the flag *Equal\_mass\_particles* to 0, while using the default value of 1 causes the particles to have an equal mass within each component (as done in Banik et al. 2020).

In this way, we are able to adequately resolve the outer parts of the discs while minimizing numerical relaxation effects in their central regions. Relaxation is much less important in the outskirts, where the very weak self-gravity means that any particle is essentially just a test particle orbiting the central region. This is probably why radial variations in  $m$  in the outskirts of each disc did not cause numerical difficulties.

## APPENDIX B: INITIAL GAS DISC THICKNESS

At each radius within the MW or M31, the vertical profile of the gas disc follows a  $\text{sech}^2$  law with characteristic thickness as shown in Fig. B1. The thickness profile is similar to that used by Banik et al. (2020) in their 100 kK model of M33 (see their fig. 3).



**Figure B1.** Initial  $\text{sech}^2$  thickness of the MW (red) and M31 (blue) gas discs as a function of cylindrical polar radius, as found by our adapted DICE algorithm (Section 2.3).

## APPENDIX C: EXTRACTING THE MW–M31 TRAJECTORY FROM THE SIMULATION

We avoid the need to save simulation snapshots except that of interest by writing each galaxy’s centre of mass position and velocity to a text file at every coarse time-step of the POR simulation. For concreteness, we focus on how this is calculated for the MW – the same procedure is used for M31. While only a limited analysis is possible ‘on the fly’, this is sufficient to gain a reasonably good understanding of the MW–M31 trajectory as a basis for a more detailed analysis.

We exploit the fact that each particle has a unique identifier, allowing us to find the barycentre of all particles that were initially part of the MW. We then find their root mean square (rms) distance from the MW barycentre in both position and velocity, which we denote  $r_{\text{rms}}$  and  $v_{\text{rms}}$ , respectively. The ratio between the two is used to define a characteristic time-scale

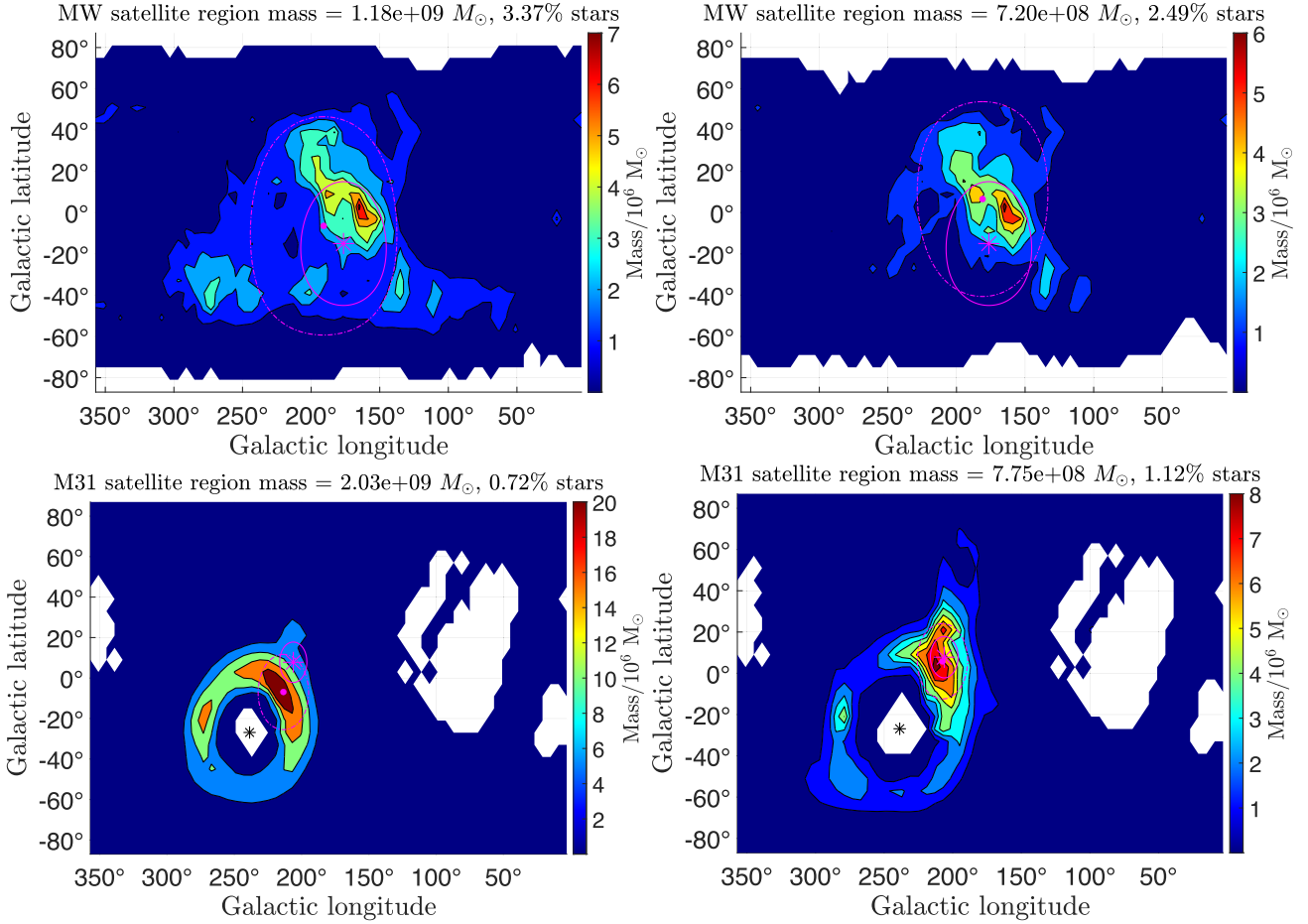
$$t_{\text{rms}} \equiv \frac{r_{\text{rms}}}{v_{\text{rms}}}. \quad (\text{C1})$$

We then iteratively repeat the calculation of all these quantities, each time considering only particles whose 6D position is within some threshold distance  $r_{\text{max}}$  from the barycentre identified at the previous step. Differences in velocity are converted to equivalent distances using  $t_{\text{rms}}$ .

After the first iteration, we consider all particles when finding the MW barycentre, since this could in principle be affected by material that has been accreted from M31. We set  $r_{\text{max}}$  to 14.1 per cent of the MW–M31 separation as calculated on the first iteration, i.e. by considering only particles that were initially part of each galaxy. For safety, we also restrict  $r_{\text{max}}$  to the range 10–70 kpc. The convergence condition is that the 6D position of the MW barycentre should move by  $< 0.01 r_{\text{rms}}$  between successive iterations, with the above-mentioned conversion between velocities and distances. We do not consider the gas as this would be extremely difficult to analyse while the simulation is running, but none the less obtain very accurate results using particles alone. The gas is considered when analysing the desired simulation snapshot in more detail.

## APPENDIX D: TIDAL DEBRIS ORBITAL POLE DISTRIBUTIONS

In Section 3.3.1, we presented the orbital pole distributions of material in the MW and M31 satellite regions. This required us to exclude the disc remnant, which we defined as all material within 250 kpc of the respective galaxy’s barycentre and up to  $z_{\text{max}} = 50$  kpc from its disc plane. To assess the impact of varying the adopted  $z_{\text{max}}$ , we use Fig. D1 to show the MW satellite region orbital pole

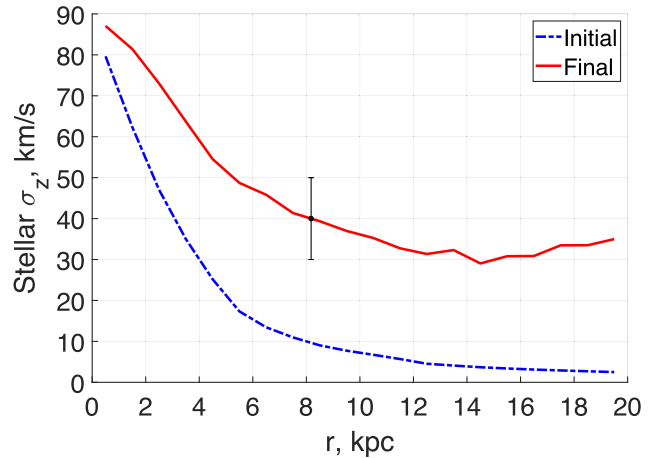


**Figure D1.** The orbital pole distribution of material in the satellite region of the MW (top) and M31 (bottom), shown similarly to Figs 7 and 8 but with an excluded region up to  $z_{\max} = 40$  kpc (left) or  $z_{\max} = 60$  kpc (right) to remove the disc remnant. The exact choice of  $z_{\max}$  has little effect on our results.

distribution if instead  $z_{\max} = 40$  kpc (top left panel) or 60 kpc (top right panel). The bottom panels of Fig. D1 show the corresponding results for M31. It is clear that for both galaxies, the preferred SP orbital pole direction and the orbital pole dispersion are little affected by the choice of  $z_{\max}$ .

## APPENDIX E: A HIGHER RESOLUTION SIMULATION

To check the numerical consistency of our results, we rerun the POR simulation with the maximum number of refinement levels raised from 12 to 13 (Section 2.6.2). This reduces the size of the smallest gas cell from 1.5 to 0.75 kpc. The results remain very similar, except for a small difference to the  $\sigma_z$  profile of the MW. We therefore show this in Fig. E1. The main difference is in the central region, presumably due to the improved resolution.  $\sigma_z$  is now higher than its initial value at all radii. It also has a much steeper decline away from the disc centre. This makes the shape of the  $\sigma_z$  profile more similar to that of M31 (Fig. 14), which changes little due to the higher resolution (not shown). However,  $\sigma_z$  is still higher for M31 than for the MW at low radii. This is also true further out, where the improved resolution slightly reduces the Galactic  $\sigma_z$ , leading to better agreement with observations (Yu & Liu 2018). A dynamically hotter stellar disc for M31 compared to the MW therefore seems to be a robust feature of our model, which also accounts for the opposite situation with their SPs.



**Figure E1.** Evolution of the Galactic  $\sigma_z$  profile, shown similarly to the bottom right panel of Fig. 13 but for a higher resolution simulation in which we allow up to 13 levels of refinement rather than 12. The result now matches the data point at  $40 \pm 10 \text{ km s}^{-1}$  for the Solar circle (Yu & Liu 2018).