The collision between the Milky Way and Andromeda

T. J. Cox* and Abraham Loeb*
Harvard–Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA

Accepted 2008 January 29. Received 2007 December 9; in original form 2007 May 8

ABSTRACT
We use an N-body/hydrodynamic simulation to forecast the future encounter between the Milky Way and the Andromeda galaxies, given present observational constraints on their relative distance, relative velocity, and masses. Allowing for a comparable amount of diffuse mass to fill the volume of the Local Group, we find that the two galaxies are likely to collide in a few billion years – within the Sun’s lifetime. During the interaction, there is a chance that the Sun will be pulled away from its present orbital radius and reside in an extended tidal tail. The likelihood for this outcome increases as the merger progresses, and there is a remote possibility that our Sun will be more tightly bound to Andromeda than to the Milky Way before the final merger. Eventually, after the merger has completed, the Sun is most likely to be scattered to the outer halo and reside at much larger radii (>30 kpc). The density profiles of the stars, gas and dark matter in the merger product resemble those of elliptical galaxies. Our Local Group model therefore provides a prototype progenitor of late-forming elliptical galaxies.

Key words: methods: N-body simulations – Galaxy: evolution – galaxies: evolution – galaxies: formation – galaxies: interactions – Local Group.

1 INTRODUCTION
It is well known that the Milky Way (MW) and Andromeda (M31) are the two largest members of the Local Group of galaxies. Together with their ~40 smaller companions, the MW and Andromeda comprise our Galactic neighbourhood and, as such, represent the nearest laboratory, and therefore the most powerful tool, to study the formation and evolution of Galactic structure.

Like most extragalactic groups, the Local Group is very likely to be decoupled from the cosmological expansion and is now a gravitationally bound collection of galaxies. This notion is supported by the observed relative motion between its two largest galaxies; namely, the MW and Andromeda are moving towards each other at ∼120 km s⁻¹ (Binney & Tremaine 1987). Unfortunately, this motion alone does not indicate whether the Local Group is bound or not. The unknown magnitude of Andromeda’s transverse velocity adds uncertainty into the present-day orbital parameters and therefore the past and future evolution of the Local Group.

Barring the uncertain transverse velocity of Andromeda, a considerable amount of information can be inferred about the Local Group, provided a plausible set of assumptions. Nearly 50 yr ago, Kahn & Woltjer (1959) pioneered the ‘timing argument’, in which the MW and Andromeda are assumed to form within close proximity to each other, during the dense early stages of the Universe, before they were pulled apart by the general cosmological expansion. They have subsequently reversed their path and are approaching one another owing to their mutual gravitational attraction. According to the timing argument, the MW and Andromeda have now traced out nearly a full period of their orbital motion which is governed by Kepler’s laws. By assuming that the system has no angular momentum and, given the present separation, velocity of approach, and the age of the Universe, the timing argument yields estimates for the mass of the Local Group (>3 × 10¹² M⊙), the semimajor axis of the orbit (<580 kpc), and the time of the next close passage (>4 Gyr) (see section 10.2 of Binney & Tremaine 1987).

While the seminal results of Kahn & Woltjer (1959) were an early indication of the large mass-to-light ratio in the Local Group and therefore the presence of dark matter, they also began a nearly five decade long quest to understand the past, present, and future of our Local Group. In particular, a number of studies have extended the original timing argument by allowing for various angular momenta, by including more realistic or time-dependent mass distributions, by adding the effects of mass at scales beyond that of the Local Group, or testing its validity using numerical simulations (see e.g. Peebles et al. 1989; Fich & Tremaine 1991; Valtonen et al. 1993; Peebles 1994; Peebles et al. 2001; Loeb et al. 2005; Sawa & Fujimoto 2005; Li & White 2008; van der Marel & Guhathakurta 2008).

One of the most-intriguing developments stemming from the various studies of the Local Group is an estimate of the transverse velocity of Andromeda. By employing the action principle to the motions of galaxies within and near (<20 Mpc) the Local Group, Peebles et al. (2001) concluded that the transverse velocity of Andromeda is less than 200 km s⁻¹. Using the well-measured transverse velocity of M33 (Brunthaler et al. 2005) and numerical simulations that tracked the potential tidal disruption during M33’s past encounters...
with Andromeda, Loeb et al. (2005) found an even smaller estimate, \( \sim 100 \text{ km s}^{-1} \), for the transverse velocity. While future astrometric observations using SIM\(^1\) and GAIA\(^2\) will be able to accurately measure the proper motion of Andromeda, the low values favoured by these papers suggest that the Local Group is indeed a gravitationally bound system.

Provided that the Local Group is gravitationally bound, and that the MW and Andromeda are heading towards each other, one must admit the possibility that they will eventually interact and merge. This outcome appears inevitable, given the massive haloes of dark matter that likely surround the MW and Andromeda. Numerical experiments have robustly concluded that dark matter haloes can exert significant dynamical friction, and are sponges that soak up energy and angular momentum leading to a rapid merger (Barnes 1988).

Even though the eventual merger between the MW and Andromeda is common lore in Astronomy, the merger process has not been addressed by a comprehensive numerical study. The one exception is a paper by Dubinski, Mihos & Hernquist (1996) that presented a viable model for the Local Group and numerically simulated the eventual merger between the MW and Andromeda. However, Dubinski et al. (1996) utilized this Local Group model and its numerical evolution to study the production of tidal tails during such an encounter and the possibility to use the structure of this tidal material to probe the dark matter potential. While the study by Dubinski et al. (1996) provided the first enticing picture of the future encounter between the MW and Andromeda (for a more recent and higher resolution version of this simulation, see Dubinski 2006), it was designed neither to detail the merger dynamics including intergalactic material, nor to outline the possible outcomes for the dynamics of our Sun, nor to quantify properties of the merger remnant. In addition, the last decade has produced a number of improved models for the structure of the MW and Andromeda as well as the properties of the intragroup medium.

In this paper, we quantitatively predict when the interaction and merger of the MW and Andromeda will likely occur and forecast the probable dynamics of the Sun during this event. We achieve this goal by constructing a model for the Local Group in Section 2 that satisfies all observational constraints. We then evolve this model using a self-consistent N-body/hydrodynamic simulation, as described in Section 3. The generic properties of the merger, including the merger time-scale, the possible evolution of our Solar system, and properties of the merger remnant, are outlined in Section 4. Finally, we conclude in Section 5.

### 2 A MODEL OF THE LOCAL GROUP

The distribution of mass within our Local Group of galaxies has been a long-standing question in astrophysics. It is clear that much of the matter is associated with the two largest galaxies in the Local Group: the MW and Andromeda. Moreover, these two spiral galaxies are likely to be embedded in an ambient medium of dark matter and gas.

#### 2.1 The MW and Andromeda

There are a number of different models for both the MW and Andromeda galaxies (see e.g. Klypin, Zhao & Somerville 2002; Widrow & Dubinski 2005; Seigar, Barth & Bullock 2006, and reference therein). These studies generally enlist a myriad of observational data to infer the distribution of baryons, while the dark matter, which dominates the gravitational potential, is set to match distributions extracted from cosmological N-body simulations (e.g. Navarro, Frenk & White 1996). Together, these models specify the total mass distribution out to the virial radius (\( \sim 200–300 \text{kpc} \)).

In our model of the Local Group we start by adopting the models for the MW and Andromeda favoured by Klypin et al. (2002). Within these models, the baryons are contained entirely within the rotationally supported exponential disc and central bulge. These components are then surrounded by a massive dark matter halo, which has nearly 20 times the mass as the baryons, as specified by the mass fractions, \( m_b \) and \( m_d \), defined as the bulge and disc mass, respectively, divided by the total mass. The exponential disc, of radial scale radius \( R_d \), also contains a set fraction \( f \) of its mass in collisional gas that can cool and form stars. Both the bulge and dark halo components are assumed to follow the Hernquist (1990) profile. The bulge scale radius \( a \) is fixed to be 20 per cent of the radial disc scale radius \( R_d \).

#### Table 1. Properties of the MW and Andromeda (M31) models used in this work (see the text for definitions).

<table>
<thead>
<tr>
<th>Property</th>
<th>MW</th>
<th>Andromeda</th>
</tr>
</thead>
<tbody>
<tr>
<td>( V_{200} ) (\text{km s}^{-1})</td>
<td>145</td>
<td>170</td>
</tr>
<tr>
<td>( M_{200} ) (10^{12} \text{M}_\odot)</td>
<td>1.0</td>
<td>1.6</td>
</tr>
<tr>
<td>( c )</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>0.031</td>
<td>0.036</td>
</tr>
<tr>
<td>( m_d )</td>
<td>0.041</td>
<td>0.044</td>
</tr>
<tr>
<td>( R_d )</td>
<td>2.2</td>
<td>3.6</td>
</tr>
<tr>
<td>( f )</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>( m_b )</td>
<td>0.008</td>
<td>0.012</td>
</tr>
<tr>
<td>( a )</td>
<td>0.4</td>
<td>0.7</td>
</tr>
<tr>
<td>( N_{dm} )</td>
<td>475 500</td>
<td>755 200</td>
</tr>
<tr>
<td>( N_{disc} )</td>
<td>14 350</td>
<td>24 640</td>
</tr>
<tr>
<td>( N_{gas} )</td>
<td>6100</td>
<td>10 560</td>
</tr>
<tr>
<td>( N_{halo} )</td>
<td>4000</td>
<td>9600</td>
</tr>
</tbody>
</table>

2 The orbit

Given the adopted parameters of the two largest galaxies in the Local Group, we must now define their orbital parameters and any ambient medium in which the system will be embedded. There are a few empirical constraints that must be considered. First, at the present epoch, the separation between the MW and Andromeda is 780 kpc (McConnachie et al. 2005; Ribas et al. 2005). Secondly, the MW and Andromeda are approaching each other at a radial speed of 120 km s\(^{-1}\), assuming a local circular speed of 220 km s\(^{-1}\) (see section 10.2 of Binney & Tremaine 1987). These observational facts tightly constrain any dynamical model of the Local Group since their fractional error bars are estimated to be less than \( \sim 5 \text{ per cent} \).
Less well constrained is the present estimate for the transverse velocity of Andromeda. As mentioned in Section 1, speeds of <200 km s\(^{-1}\) are favoured by recent models (Peebles et al. 2001; Loeb et al. 2005; van der Marel & Guhathakurta 2008), but depend on assumptions regarding the distribution of mass within the Local Group and its initial state. We will therefore gauge the success of our Local Group model by its ability to reproduce the above three observations, but it should be kept in mind that the first two have significantly less leeway than the third.

Constraints on the spin orientation of both the MW and Andromeda with respect to the orbital plane of the merger originate from the present position and orientation of Andromeda in the night sky. While these details are necessary for a complete model of the Local Group, they do not significantly influence the timing of the merger between the MW and Andromeda. The spin orientation will, however, affect the disc morphology during the merger, a detail that will be addressed when we attempt to track the possible dynamical fate of our Sun in Section 4.3.2.

### 2.3 The local intragroup medium

One plausible starting model for the Local Group is to follow the logic originally employed by the timing argument (Kahn & Woltjer 1959), that is, that the mass of the Local Group is entirely contained within the MW and Andromeda and their motion is a simple two body problem governed by Kepler’s equations. In practice, however, most of the recent implementations of the timing argument (see e.g. Binney & Tremaine 1987; Fich & Tremaine 1991; Li & White 2008) generally yield masses for the Local Group (>3 × 10\(^{12}\) M\(_\odot\)) that exceed the total masses in our MW and Andromeda models (2.6 × 10\(^{12}\) M\(_\odot\)). This discrepancy suggests that the MW and Andromeda do not contain the entire quantity of mass in the Local Group and are instead the most massive concentrations of mass within a larger all-encompassing medium. This point of view is very natural in a cosmological context where galaxies are not isolated islands, but rather mountain peaks within a vast continent of land.

For these reasons, our Local Group model supplements the MW and Andromeda galaxy models with a diffuse and extended intragroup medium, as schematically depicted in Fig. 1. Put within a cosmological framework, this initial configuration may represent a point in time when the entire Local Group has decoupled from the general Hubble flow and is in the process of collapsing to become a virialized structure.

For simplicity, we assume that the Local Group medium is initially a constant density cube of 1.5 Mpc on a side composed of both dark matter and gas. The total diffuse mass within this cube is set equal to the total mass of the two galaxy haloes today, 2.6 × 10\(^{12}\) M\(_\odot\), yielding a net mass interior to the MW/M31 orbit that is consistent with the timing argument. Our choice is motivated by other cosmological simulations (Gao et al. 2004) which indicate that a substantial fraction of the dark matter is likely to be diffuse and resides in between virialized haloes (within the unvirialized Local Group) at the initial time of our simulation ∼5 Gyr ago.\(^3\)

We also postulate that the Local Group volume contains close to the cosmic mean value of baryons, namely 16 per cent (Spiegel et al. 2003). Since the MW and Andromeda galaxy models only include baryons in the Galactic disc and bulge components, they are far short of this value. We therefore set the Local Group medium to 20 per cent primordial gas, by mass, so that the entire region approaches the cosmic mean value. The gaseous component, like the dark matter, is initialized to have a constant density, and its temperature is fixed to be 3 × 10\(^{4}\) K, consistent with estimates for the warm–hot intergalactic medium (WHIM) at comparable overdensities from cosmological simulations (see e.g. fig. 6 in Davé et al. 2001) and with observations of apparent RAM pressure stripping in the Local Group dwarf galaxy Pegasus (McConnachie et al. 2007).

We note that the initial gas temperature is far below the virialized temperature of the Local Group, however this assumption is consistent with the post-turnaround, and pre-virialized initial conditions that we adopt. The constant density is also not in line with the assumed growth of the initial stellar discs from a virialized gaseous halo (see e.g. Mo, Mao & White 1998, and references therein). Within a halo dynamical time, some of the initial intragroup gas is expected to be accreted by each galaxy, and most of the remaining intragroup gas will end up being shock-heated during the collapse and relaxation of the Local Group. The dynamics of the merger between the MW and Andromeda is dominated by the distribution of dark matter and stellar discs, and so the basic results of our simulation, involving, for example, the timing of the merger, are not directly affected by the initial gas temperature and density distribution.

Since the total mass of the diffuse medium is equivalent to the two galaxies, the two components are represented with an equivalent number of particles: 1.3 million (1.04 million dark matter, and 260 000 gas). With these assumptions, the overdensity of our initial Local Group model is δ\(_{LG}\) = \(ρ_{LG}/ρ_{crit}\) ≈ 10.

### 2.4 Other considerations

Unfortunately, once the Local Group contains this extended mass distribution the relative motion of the MW and Andromeda no longer becomes a trivial application of Kepler’s laws. The diffuse mass will

---

\(^3\) The initial scale of the Local Group in our simulation is an order of magnitude larger than the virial radius \(r_{200}\) of the MW or Andromeda galaxies. If one were to extend the Navarro, Frenk & White (1997) density profile of the envelope around each galaxy (with an asymptotic radial dependence of \(r^{-6}\)) out to our initial Local Group scale, then one would roughly double the mass found within the virial radius of each halo. The added mass in that case would be similar to the amount we indeed assume for the intragroup medium.
steadily extract orbital energy and angular momentum owing to dynamical friction, and the deep potential wells of the galaxies will slowly accrete diffuse matter. Both of these effects act to gradually extract energy from the binary system, hardening its orbit, and accelerating the merging process.

Owing to the complicated nature of the orbit, and its deviation from simple two-body motion, we are left with a fair degree of ambiguity regarding the initial state of our Local Group model. Because the configuration envisaged by Fig. 1 resembles the cosmological collapse of an overdense region of the Universe, a natural starting point is when this fluctuation decoupled from the general cosmological expansion, turned around, and began to collapse under its own self gravity. Like the original timing argument, we then presume that this scenario also occurred for the MW and Andromeda as well, that is, we initialize their radial velocity to be zero as if they have just turned around and are now about to begin their gravitational collapse.

While some insight about the value of the turn around radius can be gained from an application of the timing argument, in practice we find that the position and velocity of the MW and Andromeda quickly deviated from simple two-body motion owing to the diffuse intragroup medium. In general, the velocities quickly grew larger than the original orbit specified and the direction became highly radial. We therefore adopt a trial-and-error approach where we build an initial model, run it until the MW and Andromeda are separated by 780 kpc (as this is their present separation), and inspect the relative velocity between the two galaxies to assess the validity of the model.

Employing this procedure results in a number of models that are all honest representations of the Local Group at the present time. In practice, we find a trade-off between the initial separation and the initial orbital energy such that models which start with a larger separation require a smaller initial eccentricity. In all cases, the apparent orbit of the two galaxies becomes (or was to begin with) nearly parabolic at the present time, consistent with recent estimates of the MW–Andromeda orbit (see e.g. van der Marel & Guhathakurta 2008). While the rest of this paper will primarily present the results of one particular model, we will also show that all models yield similar estimates for the eventual merger between the MW and Andromeda. We will argue in Section 4.3.1 that this convergence results naturally from our assumed intragroup medium.

The model we choose to focus upon begins with an initial separation of 1.3 Mpc, and initializes the MW and Andromeda on an eccentric orbit $\epsilon = 0.494$, with a distance at perigalacticon of 450 kpc. With this orbit the initial angular velocity is 65 km s$^{-1}$, which could likely originate from tidal torques (Gott & Thuan 1978; Raychaudhury & Lynden-Bell 1989). This particular model begins with the largest separation of all our models and therefore is likely to be the most insensitive to its initial configuration.

3 NUMERICAL METHODS

To simulate the evolution of our Local Group and in particular the interaction between the MW and Andromeda use the publicly available N-body/hydrodynamic code GADGET2 (Springel 2005). This version of the code employs the ‘conservative-entropy’ formulation of smoothed particle hydrodynamics (SPH, Springel & Hernquist 2002) that conserves both energy and entropy (unlike earlier versions of SPH; see e.g. Hernquist 1993b), while improving shock-capturing. We assume that the gas is of primordial composition, and include the effects of radiative cooling.

Star formation and its associated feedback are both included in a manner very similar to that described in Cox et al. (2006b). As is commonly assumed, stars are stochastically formed at a rate determined by the SPH gas density (see e.g. Katz 1992; Springel & Hernquist 2003; Springel et al. 2005; Cox et al. 2006b) with an efficiency set to match the observed correlation between star formation and gas density (Kennicutt 1998).

Feedback from stellar winds and supernovae is treated in a very simplistic manner, namely the SPH particles that have sufficient density to form stars are fixed to have an effective temperature of $10^5$ K. This methodology is similar in principle to most of the presently favoured models for feedback (see e.g. Springel 2000; Springel & Hernquist 2003; Cox et al. 2006b; Stinson et al. 2006), and is easy to implement. Since the focus of this work is the large-scale evolution of the Local Group and the generic dynamics of the collision between the MW and Andromeda, the detailed treatment of the interstellar medium does not influence our primary conclusions.

Numerical resolution is a significant consideration for any computational problem. For our purposes here, we require sufficient resolution to reliably follow the interaction and merger of the MW and Andromeda, while maintaining the ability to perform a number of simulations with the available computational resources. These considerations motivated the particle number choices outlined in Section 2. Given the large number of components in these simulations (stellar discs, dark haloes, and intragroup medium) and the desire to reduce two-body effects, we required all particles to have an identical mass of $2 \times 10^7 M_\odot$ and we employed a universal gravitational softening length of 150 pc. To test the sensitivity of the results to these parameter choices, we also ran a higher resolution version of one model with 30 times the baryonic disc mass resolution (and therefore number of particles) and two times the resolution of the dark matter and intragroup medium. In this case, we also decreased the gravitational softening length of the baryons by a factor of 3, and increased that of the dark matter by a factor of 3. While this test yielded much better resolution of the stellar discs and in particular the tidal material, the general merger dynamics were identical to the low-resolution version.

4 THE COLLISION BETWEEN THE MW AND ANDROMEDA

4.1 Generic features of the merger

In Figs 2–6, we present the basic properties of the dynamical evolution of our Local Group, from 5 Gyr in the past and until 10 Gyr into the future, beyond the merger time between the MW and Andromeda. Most of the features present in these figures are generic to binary galaxy interactions, and have been described in detail by prior studies (see e.g. Toomre & Toomre 1972; Barnes & Hernquist 1991, 1992; Mihos & Hernquist 1996; Cox et al. 2006b). However, we will review some of the details that are particularly relevant to the Local Group, and subsequently highlight the unique status of our own Sun which will be a participant in this galaxy interaction. The future evolution of structures beyond the Local Group was simulated elsewhere (Busha et al. 2003, 2005; Nagamine & Loeb 2003, 2004).

To begin, Figs 2 and 3 present the entire evolution of the Local Group from the point of view of a distant observer. These images
Collision between the MW and Andromeda

Figure 2. Time-sequence of the projected stellar density, shown in red-scale, for the merger of the MW and Andromeda (M31). Andromeda is the larger of the two galaxies and begins the simulation in the upper right-hand panel. The MW begins on the edge of the image in the lower left-hand panel. Each panel is 1.5 Mpc², and the simulation time, in Gyr relative to today, appears on the top left-hand label of each panel. The trajectories of the MW and Andromeda are depicted by the blue solid lines.

begin at the start of our simulation, when the MW and Andromeda are separated by 1.3 Mpc, and include the present state of the Local Group (labelled ‘Today’) and the eventual merger of the MW and Andromeda. As a guide to the eye, each panel includes the trajectory of both the MW and Andromeda.

Shown in Fig. 2 is the evolution of the stellar component, which in our simulation only has contributions from the MW and Andromeda as we ignore any structure smaller than the two largest galaxies in the Local Group.

Fig. 3 presents the projected gas distribution during the interaction, with panels shown at the same times as in Fig. 2. Here the colour-scale has been stretched to emphasize the abundant quantity of low-density gas that is spread throughout the Local Group. The initial condition of our Local Group model assumes a uniform distribution of warm gas, however the gas quickly responds to the non-uniform potential. In particular gas is accreted and shocked to form a hydrostatic halo of warm gas around the MW and Andromeda galaxies. The gas distribution is also clearly affected by the interaction itself, as shocks develop once the galaxy haloes begin to interpenetrate at close separation.

While a detailed analysis of the diffuse gaseous intragroup medium is beyond the primary focus of our work, the basic properties of this component appear to be consistent with observations. In particular, our model predicts that, at the present state, the intragroup medium is predominantly warm $10^5$–$10^6$ K, has a fairly low density, $10^{-4}$–$10^{-6}$ cm$^{-3}$, and fills the entire volume that we simulate. We note that this medium is also expected to extend to larger scales into what has been termed the ‘WHIM’ (Hellsten, Gnedin
Figure 3. The same as Fig. 2 but only the projected gas density is shown.

While Fig. 2 presented the dynamical evolution of the stellar mass on large scales, this vantage point makes it difficult to distinguish the tell-tale signs of a galaxy merger. We therefore zoom into the central regions of the Local Group and specifically show the merger between the MW and Andromeda in Fig. 4. From this viewpoint, the classic signatures of a galaxy interaction, such as tidal tails, plumes, and shells are clearly evident.

The physical separation between the MW and Andromeda is presented in Fig. 5. The present state of the Local Group occurs $\sim 5$ Gyr after the start of our simulation. For the standard set of cosmological parameters (Tegmark 2006), this implies that we initiate the simulation at a redshift of $z \approx 0.5$, around the time when the Sun was born in the MW disc. Fig. 5 also clearly shows the prediction that our models make for the future collision between the MW and Andromeda. Their first close passage will occur less than 2 Gyr from the present, and the centres are fully coalesced in less than

4 Note that extending the simulation to significantly earlier times is not adequate since stellar ages imply that the two Galactic discs (and presumably their haloes) have not been fully assembled at $z \geq 2$ (Wyse 2007).
5 Gyr. These time-scales are comparable to the lifetime of our Sun (Sackmann, Boothroyd & Kraemer 1993) and admit the possibility that an observer in the Solar system will witness some (or all) of the galaxy collision. We will return to this possibility in Section 4.3.2.

Finally, we present the relative velocity between the MW and Andromeda in Fig. 6. As in Fig. 5 we clearly delineate the present state of the Local Group with a hatched region whose size corresponds to the errors estimated with the velocity measurements. We note that the velocities presented are relative to each galaxy’s centre of mass, and do not correct for any motion relative to that.

Now that the general features of our Local Group model have been outlined, we next explore the validity of our model, the merger dynamics, the star formation during the interaction, and the properties of the merger remnant.

4.2 The present state of the Local Group

Although our model for the Local Group presents its past, present, and future evolution, the only way to test its validity is by comparing it to the empirical data on its present-day state. As we focus primarily on the evolution of the two largest galaxies in the Local Group, the MW and Andromeda, it is only the relative separation and motion of these two galaxies that can be compared to data. As mentioned in both Sections 1 and 4, the separation and line-of-sight velocity of the MW and Andromeda are presently measured to be 780 kpc (McConnachie et al. 2005; Ribas et al. 2005, and references therein) and $-120 \text{ km s}^{-1}$ (Binney & Tremaine 1987), respectively. The proper motion of Andromeda perpendicular to our line of sight is less well constrained, but present estimates suggest that it is $<200 \text{ km s}^{-1}$ (Peebles et al. 2001; Loeb et al. 2005).

Figure 4. Time-sequence of the projected surface mass density of stars, during the final merger between Andromeda and the MW. Panels have varying scales as specified by the label, in kpc, on the lower right-hand side of each panel. The top left-hand panel is identical to the ‗Today‘ panel in Fig. 2. The simulation time, with respect to today, is shown in the top left-hand side of each panel.
Figure 5. The separation between the centres of Andromeda and the MW during the course of their merger. The present separation of $\sim 780$ kpc is shown with a horizontal dashed line and occurs at $T \approx 4.7$ Gyr.

Figure 6. The relative velocity between the centres of Andromeda and the MW galaxies during the course of their merger. The present state of the Local Group occurs at $T \approx 4.7$ Gyr.

Figs 5 and 6 present the direct comparisons between the observational constraints and our model, and clearly demonstrate that it is viable. In particular, at a time of $4.7$ Gyr after the start of the simulation (close to the present cosmic time) the separation between the MW and Andromeda is $780$ kpc, while the relative radial velocity is $135$ km s$^{-1}$ and the tangential velocity is $132$ km s$^{-1}$. While the tangential velocity is well within the limits presently favoured, the line-of-sight velocity is slightly larger than (but within $2\sigma$ of) the observations. Even though a number of our model assumptions can be manipulated to reduce these values, for example, the initial separation and eccentricity of the fiducial orbit, we note that our model is a good fit to the velocity when the separation between the MW and Andromeda is larger than $780$ kpc. Given the observational uncertainties in these values, we feel that there are likely to be a large number of models that can simultaneously fit all the data within its $2\sigma$ error bars.

4.2.1 An ensemble of models

As mentioned in Section 2.4, the model described up to this point is one of 20 Local Group models that we have simulated. Since the primary conclusion of this paper, involving the time-scale for the eventual merger between the MW and Andromeda, may be influenced by any number of our model assumptions, we explicitly show the separation between the MW and Andromeda and therefore the time of the merger, for all of our models in Fig. 7.

Fig. 7 demonstrates several interesting features. Nearly all of the models provide a similar outcome for time when the two galaxies make their first passage (ensemble average and standard deviation are $T = 2.8 \pm 0.5$ Gyr). The same holds true for the final merger time ($T = 5.4 \pm 0.4$ Gyr). We also note that these average values are slightly larger than one provided by the model described up to this point. This highlights a general trend, namely the mergers that start with a larger separation and have to traverse a longer path through the intragroup medium, usually (but not always) have a quicker merger dynamics. Regardless of the relatively small differences between merger times in these models, they are all completely coalesced by $6.2$ Gyr from today. In the following section, we will argue that this result is a direct byproduct of our inclusion of an intragroup medium.

4.3 Merger dynamics

4.3.1 Time-scale

One of the most-intriguing characteristics of our Local Group model is the relatively quick time-scale for the interaction and merger between the MW and Andromeda. As stated in the previous section, their first close passage will occur in less than $2$ Gyr, and the final coalescence will occur in less than $5$ Gyr. While the following section will specifically address the possible dynamics of the Sun
the merger time-scales extracted from binary merger simulations. Significant overdensities prior to their interaction event. In this sense, the background density, it may still change the distribution of orbits in action estimates once the haloes overlap and therefore dominate the dynamical friction on the diffuse medium. While this omission does not significantly alter the dynamical friction, it may still change the distribution of orbits in high-density environments where the galaxies traverse through significant overdensities prior to their interaction event. In this sense, the merger time-scales extracted from binary merger simulations during the merger, we show here that the cause for the short merger time-scale is the intragroup medium.

Our model for the Local Group assumes that the MW and Andromeda are embedded in a diffuse medium of dark matter and gas. The medium itself (minus the evolved galaxies) is unvirialized and of low overdensity ($\delta \sim 5$) corresponding to a region of the Universe that has decoupled from the Hubble flow and started its evolution towards a collapsed virialized state. This medium exerts dynamical friction on the MW and Andromeda and speeds up the merger dynamics by soaking orbital energy and angular momentum. This is shown explicitly in Fig. 8, which compares the MW--Andromeda separation in our fiducial model to an identical model without a diffuse medium. Without the intragroup medium, the merger time-scale is nearly three times longer than with it.

The effect of the intragroup medium should scale similarly to the standard (Chandrasekhar) formula for dynamical friction (Binney & Tremaine 1987, Equation 7-18), in which the deceleration of a massive object is proportional to the background matter density, $d\nu/dt \propto \rho$. Therefore, the rate at which angular momentum is extracted from the orbit depends on the assumed intragroup medium density, a quantity which is poorly constrained observationally. Once the dark matter haloes begin to interpenetrate, that is, when the MW--Andromeda separation is $\sim 100$ kpc, the merger completes relatively quickly because the dark matter haloes dominate over the background density.

This last point is particularly relevant to simulations of binary galaxy mergers, which typically omit any background overdensity. While this omission does not significantly alter the dynamical friction estimates once the haloes overlap and therefore dominate the background density, it may still change the distribution of orbits in high-density environments where the galaxies traverse through significant overdensities prior to their interaction event. In this sense, the merger time-scales extracted from binary merger simulations may be considered an upper limit that is most applicable to galaxies in the field.

4.3.2 The fate of our Solar System

An interesting consequence of the short time-scale for the merger between the MW and Andromeda is the possibility that a human (or decedent thereof) observer will witness the interaction and merger. At the heart of this issue is the comparison between the lifetime of the Sun and the interaction time-scale.

Present evolutionary models (see e.g. Sackmann et al. 1993) predict that the Sun will steadily increase its size and luminosity for the next 7 Gyr as it slowly consumes all available hydrogen and evolves towards a red giant phase. While this places a strong upper limit to the extent of life on Earth, it is likely that much smaller changes (<50 per cent) in the Sun’s luminosity will significantly alter the Earth’s atmosphere and thus its habitability within the next 1.1–3.5 Gyr (Kasting 1988). Korycansky, Laughlin & Adams (2001) suggested that the onset of these effects could be delayed by increasing the orbital radius of the Earth through a sequence of interactions with bodies in the outer Solar system, and we cannot rule out the possible colonization of habitable planets in nearby stars, especially long-lived M-dwarfs (Udry et al. 2007) whose lifetime may exceed $10^{12}$ yr (Adams, Bodenheimer & Laughlin 2005). In short, it is conceivable that life may exist for as little as 1.1 Gyr into the future or, if interstellar travel is possible, much longer.

Regardless of the prospects for life in the future, we can attempt to predict what any potential observer at the solar Galactic circle might see by tracking candidate Suns in our simulation. In particular, we flag all stellar particles with a galactocentric orbital radius of 8 – 0.5 kpc, corresponding to the observed value for the Sun (Eisenhauer et al. 2003), and subsequently follow the location of these particles forward in time. This procedure typically yields ~700 particles as ‘candidate Suns’ (although it should be kept in mind that owing to our limited resolution, each of the simulated particles is many orders of magnitude more massive than the Sun). The results are presented in Fig. 9 and demonstrate some of the possible outcomes for the future location of our Sun. Given the uncertainties in our model parameters and the fact that the orbital period of the Sun around the MW is much shorter than the merger time-scale, it is not possible to forecast reliably the actual phase of the Galactic orbit of the Sun at the time of closest approach to Andromeda. Therefore, we regard all the stellar particles at the galactocentric radius of the Sun as equally probable of representing the Sun. We will first outline some of the general features of our fiducial model, before showing similar distributions from a subset of models to assess the reliability of these results.

Fig. 9 outlines the wide variety of potential locations for our candidate Suns during the future evolution of the Local Group. For example, the top right-hand panel in Fig. 9 demonstrates the location of the candidate Suns after the first passage of Andromeda. At this point, the observer will most likely still be in the (now disturbed) disc of the MW, but there exists a 12 per cent chance that the Sun will be tidally ejected and take part in the tidal tail material that is $> 20$ kpc away from the MW centre.

The probability that the candidate Sun will be located farther than 20 kpc from the MW centre steadily increases as the interaction progresses. At the second passage, the percentage of Suns that are farther than 20 kpc is 30 per cent. This number increases to 48 per cent after the second passage, and is 68 per cent in the merger remnant. In fact, there is a 54 per cent chance that the Sun will be at

Figure 9. The possible location of the Sun during various stages of the merger between the MW and Andromeda. The top panel shows all stellar particles in our simulation that have a present-day galactocentric radius of $8 \pm 0.1$ kpc with a red cross and tracks their position into the future. The bottom panel presents a histogram of the radial distance from the centre of the MW.

radii larger than 30 kpc in the merger remnant. However, we caution that this probability is derived at one point in time. Individual stars will generally spend much of their orbital time at large radii, even if their orbit is eccentric and they come much closer to the Galactic Centre at other times.

One unexpected possibility for the location of the future Sun is demonstrated in the lower-middle panel of Fig. 9, namely the Sun may actually become bound to Andromeda instead of the MW before the two galaxies coalesce. Such a situation occurs when material becomes loosely bound after the first passage and is later captured by the gravitational potential of Andromeda during its second passage. While this outcome is unexpected and certainly exciting, only 2.7 per cent of the candidate Suns became bound to Andromeda and so this outcome is relatively unlikely.

In Fig. 10, we attempt to assess whether the percentages just quoted for our fiducial model are characteristic of all our models and are therefore robust. Each panel in Fig. 10 shows the distribution of galactocentric radii for candidate Suns during three periods of the interaction in several representative Local Group models, including the fiducial one. The similarity between the models highlights the generic features of the fiducial model, although there is a large spread in the percentages quoted in the previous paragraphs. The common distributions might have been expected, given the similar orbital evolution demonstrated in Fig. 7 and the tendency for mergers to preserve the hierarchy of the initial binding energies of the collisionless particles in the merger remnant.

We note that several of the interactions generate unique tidal features which result in subtle differences in the distributions shown in Fig. 10. These differences are an outcome of different disc-spin orientations that were examined in the different models. One model in particular yields a remnant where some of the candidate Suns reside in a shell-like structure at $\sim 50–60$ kpc (see the rightmost panel in Fig. 10). Only half of the models include the possibility for the Sun to be ‘stolen’ by Andromeda; however, one model yielded a probability that was nearly three times larger than in the fiducial case.

4.4 Star formation

There is mounting evidence that galaxy interactions are the predominant mechanism for producing large bursts of star formation, such as in the ultraluminous infrared galaxies (ULIRGs, Sanders &
Collision between the MW and Andromeda

Figure 10. The possible location of the Sun during the first passage (left-hand panel), second passage (middle panel), and final merger (right-hand panel) for several different models of the merger between the MW and Andromeda. The histograms are difficult to distinguish, which highlights the generic probability for the Sun to reside at various galactocentric radii. The distribution of radii from the fiducial model, which is presented in Fig. 9, is denoted by a hatched histogram.

Figure 11. The cumulative star formation rate during the merger of the MW and Andromeda compared to the star formation for models of the MW and Andromeda evolved in isolation.

While we have not explicitly tracked the black holes at the centre of the MW and Andromeda, it is interesting to speculate whether the merger will produce a luminous quasar, which many models argue is intricately linked to galaxy mergers and starbursts (Hopkins et al. 2006). Even though Fig. 11 demonstrates that there is not enough gas to fuel a powerful starburst, this gas content is clearly sufficient to ignite a luminous quasar if ~1 per cent of it is accreted by the black hole. While this work cannot address this possibility in detail, our model provides a framework to study the formation of quasars in the future.

4.5 The merger remnant

Galaxy mergers have become an area of intense study owing to their proposed role in shaping galaxy morphology. In particular, the ‘merger hypothesis’ (Toomre 1977) posits that the interaction and merger of two spiral galaxies leaves behind a remnant that is morphologically and kinematically similar to an elliptical galaxy. Since the future evolution of the Local Group contains such an event, it is natural to examine whether the Local Group will eventually consist of a single elliptical galaxy, and if so, how do the properties of this galaxy differ from present-day ellipticals which were formed at earlier epochs.

Note that the morphology of the MW and Andromeda merger remnant, which we abbreviate hereafter as Milkomeda, was presented in Fig. 4. This figure confirms the notion that galaxy mergers leave behind remnants that are spheroidal in shape, contain stars with a wide range of orbits and a large velocity dispersion, and possess very modest net rotation.

In Fig. 12, we present a series of plots that further support the assertion that Milkomeda resembles an elliptical galaxy and also serve to better quantify its properties. In panel (a), we present the spherically averaged mass profile, decomposed by the three components: dark matter, stellar and gaseous mass. At radii greater than ~2–3 kpc, the dark matter dominates the mass density, and at radii greater than ~20 kpc, the profile is well fitted by the NFW (Navarro et al. 1996) or Hernquist (Hernquist 1990) profile.

The projected mass distribution shown in panel (b) of Fig. 12 presents the first direct evidence that Milkomeda resembles an elliptical galaxy. Specifically, the stellar surface density is close to a pure $R^1/4$ distribution, with the exception of the inner ~400 pc (~2.5 times the gravitational softening length), where the surface density flattens to a nearly constant density core.
Figure 12. Properties of the MW and Andromeda merger remnant, which we abbreviate as Milkomeda. Panel (a) shows the spherically averaged distribution of dark matter, stars, and gas. The mass density is expressed in units of the present-day critical density, $\rho_{\text{crit}} = 10^{-29}$ g cm$^{-3}$. Panel (b) shows the projected distribution of stellar mass, including the contribution from stars formed during the course of the simulation which are termed ‘New Stars’, plotted against $R^{1/4}$, where $R$ is the projected radius. Panel (c) presents the kinematics of Milkomeda. For 195 projections, the ellipticity of the half-mass isophote $\epsilon$, the rotation velocity ($V$) along the major-axis $V_{\text{maj}}$ and the central velocity dispersion $\sigma$, are plotted as a coloured histogram, with the darker shades representing a higher concentration of projected properties. The solid line represents what is expected from an oblate isotropic rotation and overplotted are data extracted from the literature. Panel (d) shows the velocity anisotropy, defined as the ratio $V/\sigma$ normalized by the value expected for an oblate isotropic rotator, versus the isophotal shape parameter $a_4/a$, where positive values are discy isophotes and negative values are boxy. Overplotted in black are present data (Ralf Bender, private communication). Panel (e) shows the X-ray surface brightness profile, and panel (f) shows Milkomeda on the near-IR Fundamental Plane along with the relation observed by Pahre, de Carvalho & Djorgovski (1998).

We have also quantified the kinematics and isophotal shape of the Milkomeda galaxy following the methods of Cox et al. (2006a). Panel (c) in Fig. 12 presents the anisotropy diagram, a measure of the half-mass isophote ellipticity versus the maximum rotation along its major-axis divided by the central velocity dispersion. The shaded region represents the distribution of values for Milkomeda if viewed from 195 directions that uniformly sample the unit sphere (with angles selected using HEALPIX, Górski et al. 2005). Also plotted in this figure is the relation expected from an oblate isotropic rotator as a solid line. Owing to the high concentration of projections that closely track the solid line, this analysis demonstrates that Milkomeda is nearly an oblate isotropic rotator. In addition, the deviations from a perfect ellipse are also quantified and are presented in panel (d) in Fig. 12. Depending on the viewing direction, Milkomeda may appear to be either discy or boxy, with a slightly larger fraction of views producing discy isophotes.
Collision between the MW and Andromeda

We have also analyzed the properties of the hot gas in and around Milkomeda. The evolution of this component was presented in Fig. 3 and clearly demonstrates the formation of an extended gaseous halo, primarily accreted from the large reservoir of the intragroup medium. Although the gas temperature was originally $3 \times 10^4$ K, it has been shock-heated to the virial temperature of $\sim 3 \times 10^5$ K in Milkomeda and has a gradient to cooler temperatures at large radii. This hot gas leads to an X-ray surface brightness profile shown in panel (e) of Fig. 12 and a total X-ray luminosity of $\sim 10^{41}$ erg s$^{-1}$, which is consistent with present-day elliptical galaxies of equivalent $B$-band luminosity ($\sim 3 \times 10^{10}$ $L_\odot$).

In general, Milkomeda resembles the remnants of gas-rich major mergers, which in turn resemble the general population of low- and moderate-luminosity elliptical galaxies (Cox et al. 2006a; Naab, Jesseit & Burkert 2006). However, there are some systematic differences that likely arise because of the much smaller gas content of the MW and Andromeda when they merge. In particular, the inner regions of Milkomeda have a much lower stellar density than present-day ellipticals, which are often observed to have steep spikes of newly formed stars (Rothberg & Joseph 2004; Kormendy et al. 2007). These central excesses arise from high gas concentrations that fuel nuclear starbursts during the merger event (see e.g. Mihos & Hernquist 1994; Springel 2000; Cox et al. 2006b). This process does not occur during the formation of Milkomeda (see Section 4.4), for which the stars formed during the merger simulation have an identical profile to the entire stellar population (see panel b in Fig. 12).

The diffuse nature of Milkomeda is also evident if it is projected on to the near-IR Fundamental Plane as shown in panel (f) of Fig. 12, where it lies above the observed relation. The half-mass radius is 4.9 kpc, which is larger than the mean relation found in the Sloan Digital Sky Survey (Shen et al. 2003; Desroches et al. 2007) for galaxies of equivalent stellar mass ($1.3 \times 10^{12}$ $M_\odot$) or $r$-band absolute magnitude ($\sim 21.2$). These comparisons give credence to the claim that present-day ellipticals cannot have formed from the merger of present-day spirals (Naab & Ostriker 2007).

### 5 CONCLUSIONS

In this paper, we have used an N-body/hydrodynamic simulation to track the evolution of the Local Group, focusing primarily on the two most-massive galaxies: the MW and Andromeda. In contrast to most prior work, which typically employed models for the Local Group to infer its total mass or the proper motion of its constituents, we simulated the large-scale dynamics of all matter in the Local Group including dynamical friction on the intragroup medium, leading to the eventual merger of the MW and Andromeda.

Owing to the diffuse intragroup medium that was assumed to pervade the Local Group with a total mass comparable to that the galaxies, we have found that the interaction and merger between the MW and Andromeda will occur in less than 5 Gyr, a timescale comparable to the lifetime of the Sun (Sackmann et al. 1993). This Local Group model therefore admits the possibility that future astronomers in the Solar system will witness parts of, or the entire interaction and merger of the MW and Andromeda.

With this in mind we have calculated the probable location of our Solar system during specific points of the future interaction and find several interesting outcomes. First, there is a chance that the Sun will be ejected along with other tidal material into a long tidal tail following the first passage of Andromeda. Secondly, as a result of the disruptive effects of each close tidal passage, there is an increasing chance that the Sun will inhabit extended tidal features as the interaction proceeds. Moreover, there is a small chance that the Sun will be more tightly bound to Andromeda at some point during the merger. In such a case, Andromeda will capture the Sun and future astronomers in the Solar system might see the MW as an external galaxy in the night sky.\(^5\)

While this paper highlights the possible outcomes of the future interaction between the MW and Andromeda, we emphasize that our model is likely to be one of many plausible models within an ensemble of possibilities that span the uncertain value of the transverse velocity of Andromeda and the density (and other properties) of the intragroup medium. We have performed 20 additional runs in order to test the sensitivity of our results to various assumptions of our model – mainly involving the initial orbit of the MW and Andromeda. These runs yield similar estimates for the merger timescale as well as for the possible locations of the Sun in the future, provided that the intragroup medium is indeed similar to our fiducial case. While this gives us some confidence that our results are robust, an even larger suite of models, that spans a much wider set of model assumptions, will provide better statistics on these results.

In addition, employing higher resolution simulations with increased complexity may shed light on the nature of the intragroup medium, the soft X-ray background (Ossone et al. 2002), Galactic substructure (Willman et al. 2005), the origin of the Magellanic Clouds and Stream (Besla et al. 2007), and the future evolution of globular clusters (Forbes et al. 2000).

Finally, we note that the simulated views from the distribution of locations for the candidate Suns in the merger remnant (see Fig. 9), which we have termed Milkomeda, represent the only views available for a future local astronomer. Extragalactic astronomy will come to an end within 100 billion years if the cosmological constant will not evolve with time. Owing to the accelerated expansion caused by a steady cosmological constant, all galaxies not bound to the Local Group will eventually recede away from the Local Group and exit our event horizon (Loeb 2002). At that point, the merger product of the MW and Andromeda (with its bound satellites) will constitute the entire visible Universe (Nagamine & Loeb 2003).

### ACKNOWLEDGMENTS

The simulations were performed at the Centre for Parallel Astrophysical Computing at the Institute for Theory and Computation at the Harvard–Smithsonian Center for Astrophysics. This research was supported in part by a grant from The Foundational Questions Institute. We acknowledge helpful discussions with G. Besla, S. Dutta, L. Hernquist, P. Hopkins and B. Robertson. We thank Ralf Bender for kindly providing data used in Fig. 12.

### REFERENCES


\(^5\) Aside from the changed night sky during the interaction between the MW and Andromeda, future observers might witness enhanced comet showers due to the increased flux of stars passing by the Solar system and perturbing the Oort cloud. Additional observables would include a slightly enhanced star formation in the two galaxies and the production of hot gas by the shocks surrounding the interaction region.


Downloaded from https://academic.oup.com/mnras/article/386/1/461/978865 by guest on 21 July 2022