The Masses of the Milky Way and Andromeda Galaxies

Laura L. Watkins[1], N. Wyn Evans[1], Jin H. An[2,3,4]

[1]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge, CB3 0HA, UK*
[2] *National Astronomical Observatories, Chinese Academy of Sciences, A20 Datun Road, Chaoyang District, Beijing 100012, PR China,*
[3] *Dark Cosmology Centre, Niels Bohr Institutet, Københavns Universitet, Juliane Maries Vej 30, DK-2100 Copenhagen Ø, Denmark,*
[4] *Niels Bohr International Academy, Niels Bohr Institutet, Københavns Universitet, Blegdamsvej 17, DK-2100 Copenhagen Ø, Denmark.*

24 February 2010

ABSTRACT
We present a family of robust tracer mass estimators to compute the enclosed mass of galaxy haloes from samples of discrete positional and kinematical data of tracers, such as halo stars, globular clusters and dwarf satellites. The data may be projected positions, distances, line of sight velocities or proper motions. The estimators all assume that the tracer population has a scale-free density and moves in a scale-free potential in the region of interest. The circumstances under which the boundary terms can be discarded and the estimator converges are derived. Forms of the estimator tailored for the Milky Way galaxy and for M31 are given. Monte Carlo simulations are used to quantify the uncertainty as a function of sample size.

For the Milky Way galaxy, the satellite sample consists of 26 galaxies with line-of-sight velocities. We find that the mass of the Milky Way within 300 kpc is $M_{300} = 0.9\pm0.3\times10^{12}M_\odot$ assuming velocity isotropy. However, the mass estimate is sensitive to the assumed anisotropy and could plausibly lie between 0.7 - 3.4 $\times10^{12}M_\odot$, if anisotropies implied by simulations or by the observations are used. Incorporating the proper motions of 6 Milky Way satellites into the dataset, we find $M_{300} = 1.4 \pm 0.3 \times 10^{12}M_\odot$. The range here if plausible anisotropies are used is still broader, from 1.2 - 2.7 $\times 10^{12}M_\odot$. Note that our error bars only incorporate the statistical uncertainty. There are much greater uncertainties induced by velocity anisotropy and by selection of satellite members.

For M31, there are 23 satellite galaxies with measured line-of-sight velocities, but only M33 and IC 10 have proper motions. We use the line of sight velocities and distances of the satellite galaxies to estimate the mass of M31 within 300 kpc as $M_{300} = 1.4 \pm 0.4 \times 10^{12}M_\odot$ assuming isotropy. There is only a modest dependence on anisotropy, with the mass varying between 1.3 - 1.6 $\times 10^{12}M_\odot$. Incorporating the proper motion dataset does not change the results significantly. Given the uncertainties, we conclude that the satellite data by themselves yield no reliable insights into which of the two galaxies is actually the more massive.

Leo I has long been known to dominate mass estimates for the Milky Way due to its substantial distance and line-of-sight velocity. We find that And XII and And XIV similarly dominate the estimated mass of M31. As such, we repeat the calculations without these galaxies, in case they are not bound – although on the balance of the evidence, we favour their inclusion in mass calculations.

Key words: galaxies: general – galaxies: haloes – galaxies: kinematics and dynamics – galaxies: individual: M31 – dark matter

1 INTRODUCTION

The structure and extent of dark matter haloes have important implications for modern astrophysics, yet the determination of such properties is a difficult task and the results are often conflicting. A neat illustration is provided by the usage of Sagittarius Stream data to constrain the shape of the Milky Way dark halo. This has told us that the halo is nearly spherical (Fellhauer et al. 2006), prolate (Helmi 2004), oblate (Johnston et al. 2005) or triaxial (Law et al. 2009) in nature! The Milky Way is the closest halo available for

our study, the availability of data has improved substantially in recent years, and yet we are not able to determine its shape reliably.

Similarly, we are unable to measure the masses of the Milky Way, or its neighbour, the Andromeda Galaxy (M31) with any precision. Despite their proximity to us, their masses remain sketchily determined and there is some controversy as to which halo is more massive. Judged by criteria such as the surface brightness of the stellar halo or the numbers of globular clusters or the amplitude of the gas rotation curve, M31 is seemingly the more massive. Judged by criteria such as the velocities of the satellite galaxies and distant

globulars or tidal radii of the dwarf spheroidals, then the Milky Way is seemingly the more massive. For example, Evans et al. (2000) argued that the M31 halo is roughly as massive as that of the Milky Way, with the Milky Way marginally being the more massive of the two, while recent studies have found evidence favouring both the Milky Way (e.g. Evans & Wilkinson 2000; Gottesman et al. 2002) and M31 (e.g. Klypin et al. 2002; Karachentsev et al. 2009) as the more massive galaxy.

The masses of both haloes within a few tens of kiloparsecs are reasonably well constrained by gas rotation curve data (e.g. Rohlfs & Kreitschmann 1988; Braun 1991). However, these data only sample the inner parts of the haloes. In order to probe further out, we must turn to the kinematics of the satellite populations. Such tracers are a valuable tool for studying the dark matter haloes as their orbits contain important information about their host potential. Distance, radial velocity and proper motion data can be used to constrain halo extent, mass and velocity anisotropy (see e.g. Little & Tremaine 1987; Kochanek 1996; Wilkinson & Evans 1999).

The uncertainties in the mass estimates for the Milky Way and M31 are largely due to the fact that there is seldom proper motion data available to complement distance and radial velocity information. With only one velocity component to work with, the eccentricities of the orbits are poorly constrained. Statistical methods must be applied to determine masses and these methods suffer greatly from the small sample sizes available, even with the recent burst of satellite discoveries associated with both galaxies.

The projected mass estimator was introduced by Bahcall & Tremaine (1981). They assumed that only projected distance and line-of-sight velocity information was available. The estimator is also contained in the study of White (1981) on scale-free ensembles of binary galaxies. The analysis was extended by Heisler et al. (1985) and further modified by Evans et al. (2003) to consider the case of tracer populations. These previous studies successfully used the mass estimator to weigh M31. However, in its present form, the mass estimator is ill-suited for application to the Milky Way and such a study has not yet been attempted.

Here, we develop alternative forms of the estimator, and analyse the conditions under which they are valid. In addition, the census of satellites around M31 has increased significantly (Zucker et al. 2004, 2007; Martin et al. 2006; Majewski et al. 2007; Ibata et al. 2007; Irwin et al. 2008; McConnachie et al. 2008) since the last studies of this type were attempted and so we have more data at our disposal. Hence, we apply our estimator to M31 with these new data.

2   MASS ESTIMATORS

The projected mass estimator (Bahcall & Tremaine 1981) takes the form

$$M = \frac{C}{G} \left\langle v_{\text{los}}^2 R \right\rangle = \frac{C}{G} \frac{1}{N} \sum_{i=1}^{N} v_{\text{los},i}^2 R_i \tag{1}$$

for a set of N tracers objects (e.g planetary nebulae, stars, globular clusters, dwarf spheroidal galaxy satellites) with line-of-sight velocities v_{los} and projected distances R. Here, G is the gravitational constant and C is a constant determined by the host potential and the eccentricity of the orbits. They found that $C = 16/\pi$ for test particles with an isotropic velocity distribution orbiting a point mass and $C = 32/\pi$ for test particles moving on radial orbits.

This analysis was extended by Heisler et al. (1985) to consider

the case in which tracers may track the total mass (e.g. in galaxy groups). They found that $C = 32/\pi$ for particles with an isotropic velocity distribution and $C = 64/\pi$ for particles on radial orbits. A key assumption in this work is that the members/tracers track the mass of the group/host. This is not true for all tracer populations, particularly for those tracers which are commonly used to estimate the masses of ellipticals or the haloes of spiral galaxies.

2.1   Tracer Mass Estimator

Here, we give a formal derivation of our tracer estimators, so as to clarify the conditions under which they converge to the enclosed mass. Readers primarily interested in applications, and willing to take convergence on trust, should skip straight to the estimators themselves, namely eqns (16), (23), (24) and (26). We give formulae for the various cases in which true distances or projected distances, and line-of-sight velocities, or radial velocities or proper motions, are known for the tracers. The estimators are both simple and flexible.

Let us begin by supposing that the observations are discrete positions r and radial velocities v_r of N members of a tracer population. Here, r is measured from the centre of the host galaxy, whilst $v_r = \dot{r}$ is the radial velocity. We propose to combine the positional and kinematic data to give the enclosed mass M in the form

$$M = \frac{C}{G} \left\langle v_r^2 r^\lambda \right\rangle = \frac{C}{G} \frac{1}{N} \sum_{i=1}^{N} v_{r,i}^2 r_i^\lambda. \tag{2}$$

Here, unlike equation (1), the constant C is not necessarily dimensionless. Notice that *a priori* we do not know the best choice for λ. This will emerge from our analysis.

If f is the phase space distribution function of the tracers and σ_{r} the radial velocity dispersion, we see that under the assumption of spherical symmetry:

$$\langle v_r^2 r^\lambda \rangle = \frac{1}{M_{\text{t}}} \int d^3\boldsymbol{r} \, d^3\boldsymbol{v} \, f v_r^2 r^\lambda = \frac{4\pi}{M_{\text{t}}} \int \rho \sigma_{\text{r}}^2 r^{\lambda+2} \, dr \tag{3}$$

where M_{t} is the mass in the tracers

$$M_{\text{t}} = 4\pi \int r^2 \rho \, dr. \tag{4}$$

Now, let us assume that the tracer population is spherically symmetric and has a number density which falls off like a power-law

$$\rho(r) \propto r^{-\gamma}; \qquad \frac{d \log \rho}{d \log r} = -\gamma \tag{5}$$

at least within the radius interval $[r_{\text{in}}, r_{\text{out}}]$ where the data lie. Then, the estimator reduces to

$$\langle v_r^2 r^\lambda \rangle = \frac{1}{\mathcal{M}} \int_{r_{\text{in}}}^{r_{\text{out}}} r^{\lambda-\gamma+2} \sigma_{\text{r}}^2 \, dr \, ; \qquad \mathcal{M} = \begin{cases} \dfrac{r_{\text{out}}^{3-\gamma} - r_{\text{in}}^{3-\gamma}}{3-\gamma} & (\gamma \neq 3) \\[2ex] \log\left(\dfrac{r_{\text{out}}}{r_{\text{in}}}\right) & (\gamma = 3) \end{cases}, \tag{6}$$

where $\log x$ is the natural logarithm. Once the behaviour of σ_{r}^2 is found, we may relate this estimator to the dynamical halo mass $M(r)$. This can be achieved through solving the Jeans equation, which reads:

$$\frac{1}{\rho} \frac{d(\rho \sigma_{\text{r}}^2)}{dr} + \frac{2\beta \sigma_{\text{r}}^2}{r} = -\frac{GM(r)}{r^2}. \tag{7}$$

Here, we have introduced $\beta = 1 - \sigma_{\text{t}}^2/\sigma_{\text{r}}^2$, the Binney anisotropy

parameter, in which σ_t is the tangential velocity dispersion. Now, $\beta \to \infty$ corresponds to a circular orbit model, $\beta = 1$ corresponds to purely radial orbits and $\beta = 0$ is the isotropic case. We note that the Jeans equation (7) in a spherical system can be put into the form

$$Q\rho\sigma_r^2 = -\int Q\rho \, \frac{GM(r)}{r^2} \, dr \, ; \qquad \log Q = \int \frac{2\beta}{r} \, dr. \tag{8}$$

If β is independent of r, this simplifies to be $Q = r^{2\beta}$.

To proceed further, the underlying gravity field is assumed to be scale-free at least in the interval $[r_{\rm in}, r_{\rm out}]$, that is, the relative potential up to a constant is given by

$$\psi(r) = \begin{cases} \dfrac{v_0^2}{\alpha}\left(\dfrac{a}{r}\right)^\alpha & (\alpha \neq 0) \\[2ex] v_0^2 \log\left(\dfrac{a}{r}\right) & (\alpha = 0) \end{cases} \tag{9}$$

with $-1 \leq \alpha \leq 1$.[1] Here, a is a fiducial radius, which should lie in the region for which the power-law approximation for the relative potential is valid (i.e., $r_{\rm in} \leq a \leq r_{\rm out}$) and v_0 is the circular speed at that radius a. When $\alpha = 1$, this corresponds to the case in which the test particles are orbiting a point-mass; when $\alpha = 0$, the satellites are moving in a large-scale mass distribution with a flat rotation curve; when $\alpha = \gamma - 2$, the satellites track the total gravitating mass. We remark that our model of a scale-free tracer population of satellites in a scale-free potential has previously been used to study the mass of the Milky Way by Kulessa & Lynden-Bell (1992), although using the standard technique of maximum likelihood for parameter estimation.

The scale-free assumption is also equivalent to proposing the halo mass profile to be

$$\frac{M(r)}{M(a)} = \left(\frac{r}{a}\right)^{1-\alpha}, \tag{10}$$

and the local mass density $\propto r^{-(\alpha+2)}$. Consequently, if the power-law behaviour were allowed to be extended to infinity, the total mass of the dark halo would necessarily be infinite unless $\alpha = 1$. (However, if the halo density were to fall off faster than r^{-3} and so the total gravitating mass is finite, the leading term for the potential would be Keplerian. That is to say, for the case of a finite total mass halo, the gravity field experienced by the tracers may be approximated to be that of a point mass, given that $r_{\rm in}$ is chosen to be sufficiently large so that the gravitating mass inside the sphere of $r_{\rm in}$ dominates the mass within the shell region populated by the tracers.)

Combining this with the constant-anisotropy assumption, the Jeans equation integrated between r and $r_{\rm out}$ then reduces to

$$r^{2\beta-\gamma}\sigma_r^2(r) - r_{\rm out}^{2\beta-\gamma}\sigma_r^2(r_{\rm out}) = \frac{GM(a)}{a^{1-\alpha}}\int_r^{r_{\rm out}} \tilde{r}^{2\beta-\gamma-\alpha-1}d\tilde{r}. \tag{11}$$

provided that all our assumptions remain valid in the radius interval $[r_{\rm in}, r_{\rm out}]$ and $r, a \in [r_{\rm in}, r_{\rm out}]$.

Now, our goal is to find the total halo mass. In reality, the observed tracers are only populated up to a finite outer radius, and so, any mass distribution outside of that radius does not affect our observations in a strictly spherical system (Newton's theorem). We therefore extend the power-law potential assumption only up to the

finite outer radius (here $r_{\rm out}$), and set $a = r_{\rm out}$. In other words, the halo mass that we are interested in is that contained within the outer radius, $M = M(r_{\rm out})$. With $a = r_{\rm out}$, solving equation (11) for $\sigma_r^2(r)$ results in (here $s \equiv r/r_{\rm out}$)

$$\sigma_r^2 = \begin{cases} \dfrac{\sigma_r^2(r_{\rm out}) - \hat{v}_0^2}{s^{2\beta-\gamma}} + \dfrac{\hat{v}_0^2}{s^\alpha} & (\alpha + \gamma - 2\beta \neq 0) \\[3ex] \dfrac{\sigma_r^2(r_{\rm out}) - v_0^2 \log s}{s^\alpha} & (\alpha = 2\beta - \gamma) \end{cases} \tag{12}$$

where $v_0^2 = GM/r_{\rm out}$ is the circular speed at $r_{\rm out}$ whilst $\hat{v}_0^2 \equiv v_0^2/(\alpha + \gamma - 2\beta)$.

Then, substituting the result of equation (12) into equation (6) and explicitly performing the integration yields (ignoring particular parameter combinations that involve the logarithm)

$$\frac{\langle v_r^2 r^\lambda \rangle}{(3-\gamma)r_{\rm out}^\lambda} = \frac{v_0^2}{(\lambda-\alpha+3-\gamma)(\alpha+\gamma-2\beta)}\frac{1-u^{\lambda-\alpha+3-\gamma}}{1-u^{3-\gamma}}$$
$$+ \frac{1}{\lambda-2\beta+3}\left[\sigma_r^2(r_{\rm out}) - \frac{v_0^2}{\alpha+\gamma-2\beta}\right]\frac{1-u^{\lambda-2\beta+3}}{1-u^{3-\gamma}} \tag{13}$$

where $u \equiv r_{\rm in}/r_{\rm out}$. Notice now that the choice of $\lambda = \alpha$ makes the u-dependence of the first term in the right-hand side drop out. In fact, this could also have been deduced on dimensional grounds by requiring that our estimator is not dominated by datapoints at small radii or large radii.

The last terms in equation (13) basically constitute the surface 'pressure' support terms in the Jeans equation, which we wish to minimize as $u \to 0$. Here, we limit ourselves to the case that $\lambda = \alpha$, when the corresponding leading term is

$$\frac{1-u^{\alpha-2\beta+3}}{1-u^{3-\gamma}} \sim \begin{cases} 1 & 2\beta-\alpha, \gamma < 3 \\ -u^{-(2\beta-\alpha-3)} & \gamma < 3 < 2\beta-\alpha \\ -u^{\gamma-3} & 2\beta-\alpha < 3 < \gamma \\ u^{\alpha+\gamma-2\beta} & 3 < 2\beta-\alpha, \gamma \end{cases}. \tag{14}$$

In other words, provided that $\gamma > 3$ and $\gamma > 2\beta - \alpha$, the pressure term vanishes as $u \to 0$, and we obtain the scale-free Jeans solutions of Evans et al. (1997). In fact, since $\beta \leq 1$ and $-1 \leq \alpha \leq 1$, we find that $2\beta - \alpha \leq 3$ and thus the second condition here is essentially redundant. Consequently, provided that $\gamma > 3$, that is the tracer density falls off more quickly than r^{-3}, we find the estimator to be

$$\langle v_r^2 r^\alpha \rangle \simeq \frac{r_{\rm out}^\alpha}{\alpha+\gamma-2\beta}\frac{GM}{r_{\rm out}} + \mathcal{R} \tag{15}$$

where the remainder $\mathcal{R} \to 0$ vanishes as $r_{\rm in}/r_{\rm out} \to 0$ (here, $r_{\rm in}$ and $r_{\rm out}$ are the inner and outer radius of the tracer population).

Alternatively, if $\gamma < 3$ and $2\beta - \alpha < 3$, the remainder term tends to a constant as $u \to 0$. In a perfectly scale-free halo traced by again strictly scale-free populations, this constant must be zero. This is because, for such a system, σ_r^2 should also be scale-free. Yet equation (12) implies that this is possible only if $\sigma_r^2(r_{\rm out}) = \hat{v}_0^2$. Subsequently this also indicates that the coefficient for the remainder in equation (13) vanishes too. Even after relaxing the everywhere strict power-law behaviour, we would expect that $\sigma_r^2 \sim \hat{v}_0^2$ and consequently that $|\sigma_r^2 - \hat{v}_0^2| \ll \hat{v}_0^2$, provided that $2\beta - \alpha < \gamma$, which is required to ensure $\hat{v}_0^2 > 0$. That is to say, we expect that $\hat{v}_0^2 r_{\rm out}^\alpha \gg \mathcal{R}$ as $u \to 0$ in equation (15) for $2\beta - \alpha < \gamma < 3$, which is sufficient for justifying the applicability of our mass estimator.

In other words, we have obtained a very simple result

$$M = \frac{C}{G}\left\langle v_r^2 r^\alpha \right\rangle, \qquad C = (\alpha+\gamma-2\beta)\, r_{\rm out}^{1-\alpha}, \tag{16}$$

---

[1] $\alpha = -1$ corresponds to the gravitational field that pulls with an equal magnitude force regardless of radius, which is formally generated by a halo density falling off as r^{-1}. Provided we regard the scale-free potential as an approximation valid over a limited range and not extending to spatial infinity, we can permit $\alpha \geq -2$, since $\alpha = -2$ corresponds to the harmonic potential generated by a homogeneous sphere.

provided that $C > 0$ (the simple interpolative argument indicates that this is still valid for $\gamma = 3$). This corresponds to the case in which the tracers have known radial velocity components v_r resolved with respect to the centre of the galaxy, as well as actual distances r. For satellites of the Milky Way, the line of sight velocity v_{los} is measured, and corrected to the Galactic rest frame. Now, v_r may be calculated from v_{los} only if proper motion data exists. Alternatively, a statistical correction can be applied to estimate v_r from v_{los}

$$\langle v_r^2 \rangle = \frac{\langle v_{\mathrm{los}}^2 \rangle}{1 - \beta \sin^2 \varphi} \tag{17}$$

where φ is the angle between the unit vector from the Galactic Centre to the satellite and the unit vector from the Sun to the satellite.

Note too that in the important isothermal case ($\alpha = 0$), the galaxy rotation curve is flat with amplitude v_0. Then, for members of a population with density falling like $\rho \sim r^{-3}$, such as the Galactic globular clusters, eqn (16) reduces to

$$v_0^2 = (3 - 2\beta)\langle v_r^2 \rangle. \tag{18}$$

This is a generalization of the estimator of Lynden-Bell & Frenk (1981) to the case of anisotropy. When the population is isotropic ($\beta = 0$), it reduces to the appealing simple statement that the circular speed is the rms velocity of the tracers multiplied by $\sqrt{3} \approx 1.732$.

Even if three dimensional distance r is replaced by projected distance R or v_r by some other projections of the velocity, the basic scaling result of eqn (16) remains valid. Different projections simply result in distinct constants C, as we now show.

2.2   A Family of Estimators

Now, suppose that we have actual distances r from the centre of the host galaxy, but only projected or line of sight velocities v_{los}. This is the case for many of M31's satellite galaxies, for which distances have been measured by using the tip of the red giant branch method (see e.g., McConnachie et al. 2005) and for which projected velocities are known from spectroscopy. The calculation proceeds by considering

$$\langle v_{\mathrm{los}}^2 r^\alpha \rangle = \frac{1}{M_{\mathrm{t}}} \int \mathrm{d}^3 \boldsymbol{r} \, \mathrm{d}^3 \boldsymbol{v} \, f v_{\mathrm{los}}^2 r^\alpha = \frac{2\pi}{M_{\mathrm{t}}} \int_{r_{\mathrm{in}}}^{r_{\mathrm{out}}} \mathrm{d}r \int_0^\pi \mathrm{d}\theta \, \rho \sigma_{\mathrm{los}}^2 r^{\alpha+2} \sin\theta \tag{19}$$

We now need the relationship between line-of-sight velocity dispersion σ_{los} and the radial velocity dispersion σ_{r}, namely

$$\sigma_{\mathrm{los}}^2 = \sigma_{\mathrm{r}}^2 \left(1 - \beta \sin^2 \varphi\right), \tag{20}$$

which is similar to equation (17) but here the angle φ is the angle between the line of sight and the position vector of the satellite with respect to the centre of the *host galaxy* (see e.g., Binney & Tremaine 1987, Section 4.2). If the polar z-axis of the coordinate system is chosen such that the sun (that is, the observer) lies on the negative z-axis (i.e., $\theta = \pi$) at a distance d from the centre of the host galaxy, we find that

$$\sin^2 \varphi = \frac{\sin^2 \theta}{1 + 2\frac{r}{d}\cos\theta + \left(\frac{r}{d}\right)^2}. \tag{21}$$

However, for most external galaxies, it is reasonable to assume $d \gg r_{\mathrm{out}}$, and therefore, we can safely approximate[2] that $\sin^2 \varphi \approx \sin^2 \theta$.

Then,

$$\langle v_{\mathrm{los}}^2 r^\alpha \rangle = \langle v_r^2 r^\alpha \rangle \int_0^{\pi/2} \mathrm{d}\theta \, \sin\theta \left(1 - \beta \sin^2 \theta\right), \tag{22}$$

and thus we find that

$$M = \frac{C}{G}\left\langle v_{\mathrm{los}}^2 r^\alpha \right\rangle, \qquad C = \frac{3\,(\alpha + \gamma - 2\beta)}{3 - 2\beta} r_{\mathrm{out}}^{1-\alpha}. \tag{23}$$

Next, we consider the case in which we have full velocity information for the satellites, i.e., both radial velocities and proper motions. For example, this is the case for a subset of the satellites of the Milky Way (see e.g., Piatek et al. 2002). In this case, we can utilize $\sigma^2 = \sigma_{\mathrm{r}}^2 + \sigma_{\mathrm{t}}^2 = (3 - 2\beta)\sigma_{\mathrm{r}}^2$, and therefore the estimator becomes

$$M = \frac{C}{G}\left\langle v^2 r^\alpha \right\rangle, \qquad C = \frac{\alpha + \gamma - 2\beta}{3 - 2\beta} r_{\mathrm{out}}^{1-\alpha}. \tag{24}$$

Finally, we can assume a worst-case scenario in which the only data available are projected distances R and line-of-sight velocities v_{los} for the tracers. Outside of the galaxies of the Local Group, this is the usual state of affairs. So, this would be the form of the estimator to find the dark matter mass of nearby giant ellipticals like M87 from positions and velocities of the globular clusters. The estimator is derived following the same procedure with $R = r \sin\theta$, which results in the relation

$$\langle v_{\mathrm{los}}^2 R^\alpha \rangle = \langle v_r^2 r^\alpha \rangle \int_0^{\pi/2} \mathrm{d}\theta \, \sin^{\alpha+1}\theta \left(1 - \beta \sin^2 \theta\right). \tag{25}$$

Consequently, the corresponding estimator is found to be [3]

$$M = \frac{C}{G}\left\langle v_{\mathrm{los}}^2 R^\alpha \right\rangle, \qquad C = \frac{(\alpha + \gamma - 2\beta)}{I_{\alpha,\beta}} r_{\mathrm{out}}^{1-\alpha} \tag{26}$$

where

$$I_{\alpha,\beta} = \frac{\pi^{1/2}\Gamma(\frac{\alpha}{2} + 1)}{4\Gamma(\frac{\alpha}{2} + \frac{5}{2})} \left[\alpha + 3 - \beta(\alpha + 2)\right] \tag{27}$$

and $\Gamma(x)$ is the gamma function. This case is related to work by Bahcall & Tremaine (1981). So, for example, in the Keplerian case ($\alpha = 1$), a distribution of test particles with $\gamma = 3$ gives

$$C = \frac{32}{\pi} \frac{2 - \beta}{4 - 3\beta} \tag{28}$$

When $\beta = 0$, this implies that $C = 16/\pi$; whilst when $\beta = 1$, $C = 32/\pi$.

Some of these estimators are implicit in other work. In particular, some are equivalent to those introduced by White (1981), who had a different focus on the dynamics of binary galaxies but who made the same scale-free assumptions to obtain robust mass estimators. Very recently, An & Evans (2010, in preparation) found a related family of estimators that are independent of parameters derived from the tracer density (like γ).

3   CHECKS WITH MONTE CARLO SIMULATIONS

In order to verify the correctness of our mass estimators, we generate synthetic data-sets of anisotropic spherical tracer populations.

---

[2] On the other hand, for the satellites of the Milky Way, it is often assumed that $d \ll r_{\mathrm{in}}$, which leads to $\sin\varphi \approx 0$ and consequently $\langle v_{\mathrm{los}}^2 r^\alpha \rangle \approx \langle v_r^2 r^\alpha \rangle$.

[3] The result is valid provided that the integral is limited to spherical shells. However, given the lack of depth information, it might seem more logical to perform the integration over cylindrical shells. Unfortunately, the result is more complicated, as it involves the integrals of incomplete beta functions.
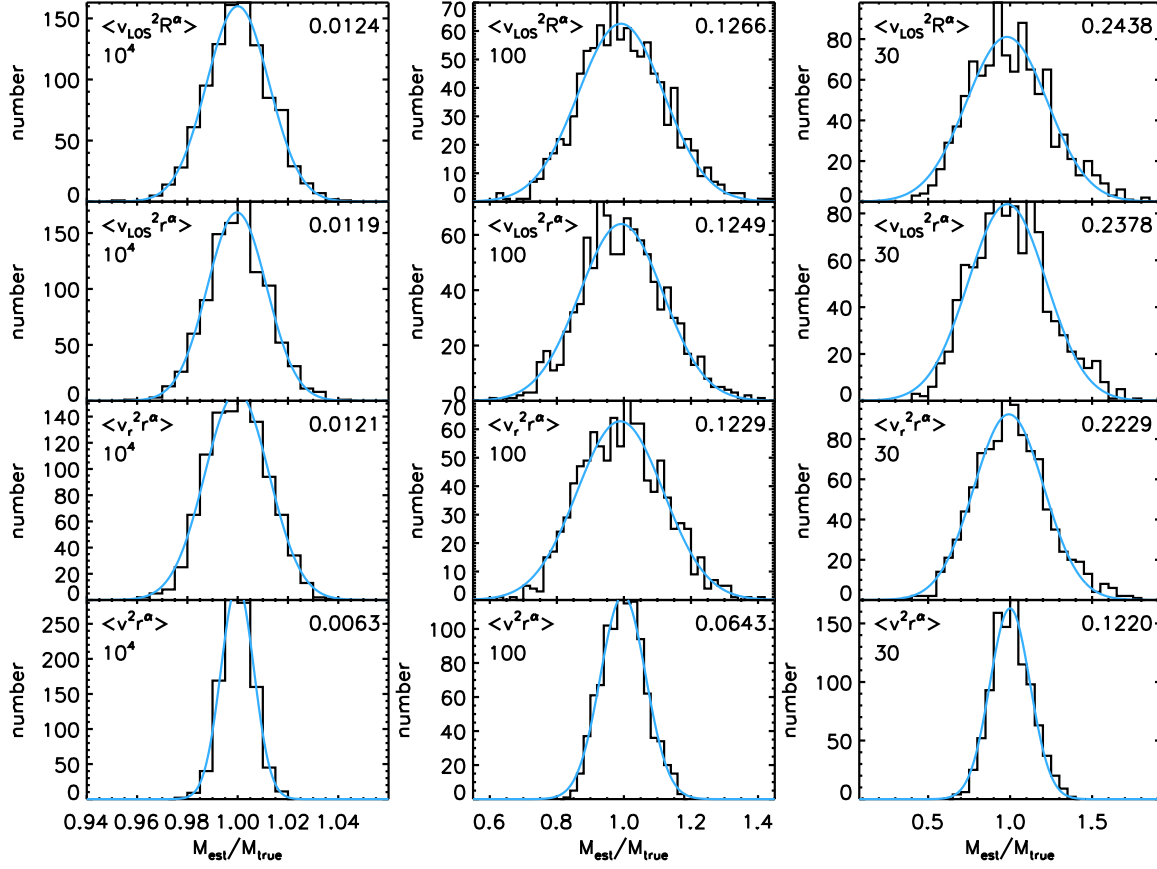
Figure 1. Distribution of mass estimate as a fraction of the true mass for 1000 Monte Carlo realisations, assuming that parameters α, β and γ are known exactly. Left: $N = 10,000$. Middle: $N = 100$. Right: $N = 30$. The number of satellites in the simulation and the form of the estimator used to recover the mass is shown in the top left corner of each panel. A best-fit Gaussian is plotted for each distribution and the standard deviation of the distribution is shown in the top right corner of each panel. On average, the tracer mass estimator recovers the true mass of the host. [The cases shown correspond to $\alpha = 0.55$, $\beta = 0.0$ and $\gamma = 2.7$].

Distances r are selected in $[r_{\rm in}, r_{\rm out}]$ assuming the power-law density profile in eqn (5). Projection directions are determined by the position angles: $\cos\theta$ is generated uniformly in $[-1, 1]$ and ϕ is generated uniformly in $[0, 2\pi]$. If R lies outside of the allowed range, the projection direction is regenerated until R is within $[R_{\rm in}, R_{\rm out}]$.

The phase-space distribution functions that give rise to such density profiles are given in Evans et al. (1997). Tracer velocities are picked from the distributions

$$f(v) \propto \begin{cases} v^{2-2\beta} \left| \psi(r) - \frac{1}{2}v^2 \right|^{[2\gamma - 3\alpha - 2\beta(2-\alpha)]/(2\alpha)} & (\alpha \neq 0) \\ v^{2-2\beta} \exp\left(-\dfrac{v^2}{2\sigma^2}\right) & (\alpha = 0) \end{cases} . \quad (29)$$

For $\alpha > 0$, the maximum velocity at any position is $\sqrt{2\psi(r)}$; for $\alpha \leq 0$, the velocities can become arbitrarily large. Following Binney & Tremaine (1987), we introduce spherical polar coordinates in velocity space (v, ξ, η) so that the velocities resolved in spherical polar coordinates with respect to the centre are then

$$v_r = v\cos\eta \quad v_\theta = v\sin\eta\cos\xi \quad v_\phi = v\sin\eta\sin\xi \quad (30)$$

To generate velocities with the correct anisotropy, ξ is generated uniformly in $[0, 2\pi]$ and η is picked in $[0, \pi]$ from the distribution

$$F(\eta) \propto |\sin\eta|^{1-2\beta} \quad (31)$$

where β is the Binney anisotropy parameter. Finally, the line-of-sight velocities are calculated and used in the tracer mass estimator.

Figure 1 shows the distribution of mass estimates as fractions of the true mass for 1000 realisations, assuming that parameters α, β and γ are known exactly; the left panels show simulations with 10,000 tracers, the middle panels for 100 tracers and the right panels for 30 tracers. The panels use the different forms of the estimator given in eqns (16), (23), (24) and (26) respectively. A Gaussian with the same standard deviation as each distribution is also plotted for each panel. The standard deviation is included in the top-right corner of each plot and gives an estimate of the error in each case.

We see that our mass estimators are unbiased – that is, on average, the true mass is recovered in all cases. The benefit of using three dimensional distances r instead of projected distances R is modest, as is the improvement gained by using v_r in place of $v_{\rm los}$. However, if proper motion data are available, then using v instead of v_r gives a more accurate mass estimate.

So far, we have assumed that we know α, β and γ exactly, which is, of course, not the case. Our estimates for α, β and γ have errors associated with them, not least because the notion of a scale-free density profile in a scale-free potential is an idealization. As these parameters enter the estimator through the prefactor C, it is straightforward to obtain the additional uncertainty in the fi-

nal answer using propagation of errors. As we will show in the next section α and γ are constrained either by cosmological arguments or by the data. The right-most column in Figure 1 (a host with 30 satellites) is the most applicable to our data-sets at present as the Milky Way has 26 satellites and M31 23 satellites with a recorded line-of-sight velocity. The error on the mass estimate obtained in this case is $\sim 25\%$. This is much larger than that the effects of errors on α and γ and so the latter will be ignored for the rest of the discussion.

However, the case of the velocity anisotropy β is different as it is poorly constrained, with theory and data pointing in rather different directions. Changes in β can therefore make a substantial difference to the mass estimate.

Note that these simulations yield no insight into systematic errors, because the mock data are drawn from the same distribution functions used to derive the form of the mass estimators. This is a concern as there are a number of causes of systematic error – for example, dark halos may not be spherical, or infall may continue to the present day so that the observed satellites may not necessarily be virialized. Deason et al. (2010, in preparation) have tested the estimators derived in this paper, as well as a number of other commonly used mass estimators, against simulations. Specifically, they extracted samples of Milky Way-like galaxies and their satellites from the *Galaxies Intergalactic Medium Interaction Calculation* (Crain et al. 2009), a recent high resolution hydrodynamical simulation of a large volume of the Universe. They find that the estimators in this paper significantly out-perform the projected mass estimator of Bahcall & Tremaine (1981) and the tracer mass estimator of Evans et al. (2003).

4   MASS ESTIMATES FOR ANDROMEDA AND THE MILKY WAY

4.1   Choice of Power-Law Index Parameters

We now apply the mass estimators to the Milky Way and M31, the two largest galaxies in the Local Group. In converting heliocentric quantities to Galactocentric ones, we assume a circular speed of 220 km s^{-1} at the Galactocentric radius of the sun ($R_\odot = 8.0$ kpc) and a solar peculiar velocity of $(U, V, W) = (10.00, 5.25, 7.17)$ km s^{-1}, where U is directed inward to the Galactic Centre, V is positive in the direction of Galactic rotation at the position of the sun, and W is positive towards the North Galactic Pole (see e.g., Dehnen & Binney 1998).

For the M31 satellites, positional and velocity data must be computed relative to M31 itself. We take the position of M31 to be $(\ell, b) = (121.2°, -21.6°)$ at a distance of 785 kpc and its line-of-sight velocity to be -123 km s^{-1} in the Galactic rest frame (see e.g., McConnachie et al. 2005; McConnachie & Irwin 2006).

In order to apply our estimators to these systems, we need to compute the power-law index of the host potential α, the velocity anisotropy β and the power-law index of the satellite density distribution γ. There are cosmological arguments suggesting that the potentials of dark haloes are well-approximated by Navarro-Frenk-White (NFW) profiles (Navarro et al. 1996). Figure 2 shows the best-fit power-law to the NFW potential for a wide range of concentrations and virial radii. The fitting is performed in the region $10 < r/\mathrm{kpc} < 300$, which is where the majority of the satellites lie. Now, Klypin et al. (2002) argued that the concentrations of the Milky Way and M31 are $c \approx 12$, whilst the virial radii r_{vir} are in the range 250-300 kpc. In other words, for the range of concentrations and virial radii appropriate to galaxies like the Milky Way and

Table 1. Data table for the satellites of the Milky Way. Listed are Galactic coordinates (l, b) in degrees, Galactocentric distance r in kpc and corrected line-of-sight velocity in km s^{-1}.

Name	l (deg)	b (deg)	r (kpc)	v_{los} (km s^{-1})	Source
Bootes I	358.1	69.6	57	106.6	1,2
Bootes II	353.8	68.8	43	-115.6	3,4
Canes Venatici I	74.3	79.8	219	76.8	5,6
Canes Venatici II	113.6	82.7	150	-96.1	6,7
Carina	260.1	-22.2	102	14.3	8,9
Coma Bernices	241.9	83.6	45	82.6	6,7
Draco	86.4	34.7	92	-104.0	8,10,11
Fornax	237.3	-65.6	140	-33.6	8,12,13
Hercules	28.7	36.9	141	142.9	6,7
LMC	280.5	-32.9	49	73.8	8,14,15
Leo I	226.0	49.1	257	179.0	8,16,17
Leo II	220.2	67.2	235	26.5	8,18,19
Leo IV	265.4	56.5	154	13.9	6,7
Leo T	214.9	43.7	422	-56.0	6,20
Leo V	261.9	58.5	175	62.3	21
SMC	302.8	-44.3	60	9.0	8,22,23
Sagittarius	5.6	-14.1	16	166.3	8,24
Sculptor	287.5	-83.2	87	77.6	8,25,26
Segue 1	220.5	50.4	28	113.5	3,27
Segue 2	149.4	-38.1	41	39.7	28
Sextans	243.5	42.3	89	78.2	8,9,29
Ursa Major I	159.4	54.4	101	-8.8	3,6
Ursa Major II	152.5	37.4	36	-36.5	6,30
Ursa Minor	104.9	44.8	77	-89.8	8,10,11
Willman 1	158.6	56.8	42	33.7	2,3

Sources: 1 - Belokurov et al. (2006), 2 - Martin et al. (2007), 3 - Martin et al. (2008), 4 - Koch et al. (2009), 5 - Zucker et al. (2006), 6 - Simon & Geha (2007), 7 - Belokurov et al. (2007), 8 - Karachentsev et al. (2004), 9 - Mateo (1998), 10 - Bellazzini et al. (2002), 11 - Armandroff et al. (1995), 12 - Saviane et al. (2000), 13 - Walker et al. (2006), 14 - Freedman et al. (2001), 15 - van der Marel et al. (2002), 16 - Bellazzini et al. (2004), 17 - Koch et al. (2007), 18 - Bellazzini et al. (2005), 19 - Koch et al. (2007), 20 - Irwin et al. (2007), 21 - Belokurov et al. (2008), 22 - Cioni et al. (2000), 23 - Harris & Zaritsky (2006), 24 - Ibata et al. (1997), 25 - Kaluzny et al. (1995), 26 - Queloz et al. (1995), 27 - Geha et al. (2009), 28 - Belokurov et al. (2009), 29 - Walker et al. (2006), 30 - Zucker et al. (2006)

M31, we see – fortunately – that the surface in Figure 2 is slowly-changing and flattish with $\alpha \approx 0.55$.

If the satellite number density distribution $n(r)$ follows a power law with index γ, then the number of satellites within any radius, $N(< r)$, also follows a power-law with index $3 - \gamma$. We fit power-laws to the Milky Way and M31 satellite cumulative distributions in order to estimate γ. We restrict ourselves to the inner regions of the satellite distributions, $r \leq 300$ kpc; beyond this range, the satellite population is likely to be seriously incomplete. The distributions and the best-fitting power laws are shown in Figure 3; the Milky Way data is shown in the upper panel and M31 data is shown in the lower panel. We find $\gamma = 2.6$ for the Milky Way and $\gamma = 2.1$ for M31. Note that data from the Sloan Digital Sky Survey (SDSS; York et al. 2000) has been instrumental in the identification of many of the recently-discovered Milky Way dwarfs. The SDSS coverage includes only the region around the North Galactic Cap, and, as such, the distribution of known Milky Way satellites is concentrated in that area of the sky. However, given our underlying assumption that the distribution of satellites is spherically symmetric,
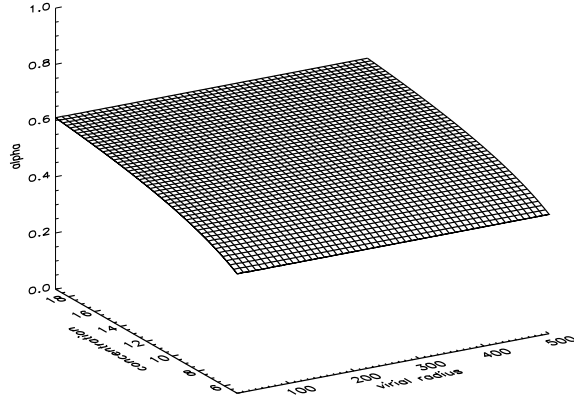
Figure 2. The best-fit value of the power-law index α to an NFW profile as a function of the concentration and virial radius. Note that for plausible values of the concentration c and the virial radius $r_{\rm vir}$ for galaxies like the Milky Way and M31, α lies in the range 0.5-0.6. The surface is smooth and flattish, implying that α is reasonably insensitive to the details of the NFW potential.
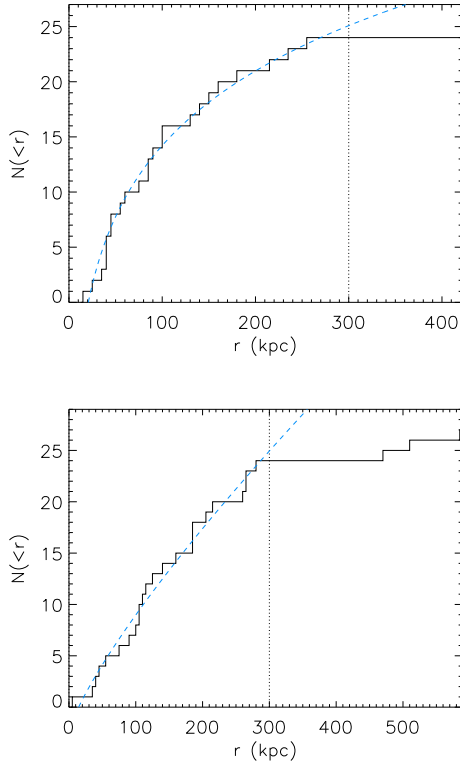

Figure 3. Cumulative numbers of satellite $N(< r)$ for the Milky Way (upper) and M31 (lower). The best-fit power laws in the range $r \leq 300$ kpc are also plotted. The index of these power-law fits may be used to estimate the power-law index of the satellite density distribution $n(r)$.

Table 2. Data table for the satellites of M31. Listed are Galactic coordinates (l, b) in degrees, actual distance r from the centre of M31 in kpc, projected distance R from the centre of M31 in kpc and corrected line-of-sight velocity in km s^{-1}.

Name	l (deg)	b (deg)	r (kpc)	R (kpc)	$v_{\rm los}$ (km s^{-1})	Source
M33	133.6	-31.3	809	206	74	1,2
M32	121.1	-22.0	785	5	95	2,3
IC 10	119.0	-3.3	660	261	-29	2,3,4
NGC 205	120.7	-21.1	824	39	58	1,2
NGC 185	120.8	-14.5	616	189	106	1,2
IC 1613	129.8	-60.6	715	510	-56	2,3,5
NGC 147	119.8	-14.2	675	144	117	1,2
Pegasus	94.8	-43.6	919	473	85	1,2
Pisces	126.7	-40.9	769	268	-37	1,2
And I	121.7	-24.8	745	59	-84	1,2
And II	128.9	-29.2	652	185	83	1,2
And III	119.4	-26.3	749	75	-57	1,2
And V	126.2	-15.1	774	109	-107	1,2
And VI	106.0	-36.3	775	267	-64	1,2
And VII	109.5	-9.9	763	218	21	1,2
And IX	123.2	-19.7	765	41	94	1,6,7
And X	125.8	-18.0	702	110	130	8,9
And XI	121.7	-29.1	785	102	-140	7,10
And XII	122.0	-28.5	830	107	-268	7,10,11
And XIII	123.0	-29.9	785	115	64	7,10
And XIV	123.0	-33.2	740	161	-204	12
And XV	127.9	-24.5	770	94	-57	13,14
And XVI	124.9	-30.5	525	280	-106	13,14
And XVII	120.2	-18.5	794	45		15
And XVIII	113.9	-16.9	1355	589		16
And XIX	115.6	-27.4	933	187		16
And XX	112.9	-26.9	802	128		16
And XXI	111.9	-19.2	859	148		17
And XXII	132.6	-34.1	794	220		17

Sources: 1 - McConnachie et al. (2005), 2 - McConnachie & Irwin (2006), 3 - Karachentsev et al. (2004), 4 - Sakai et al. (1999), 5 - Cole et al. (1999), 6 - Zucker et al. (2004), 7 - Collins et al. (2009, in prep), 8 - Zucker et al. (2007), 9 - Kalirai et al. (2009, in prep), 10 - Martin et al. (2006), 11 - Chapman et al. (2007), 12 - Majewski et al. (2007), 13 - Ibata et al. (2007), 14 - Letarte et al. (2009), 15 - Irwin et al. (2008), 16 - McConnachie et al. (2008), 17 - Martin et al. (2009)

this directional bias does not impair our mass estimators. A bigger worry may be the incompleteness in the satellite distribution, which could affect the power index for the tracer number density if the directional incompleteness varies in different distances.

Finally, there are a number of possibilities for the velocity anisotropy for the satellite galaxies. Previous studies often assumed isotropy, arguing that there is no compelling evidence to the contrary. However, Diemand et al. (2007) found that the velocity anisotropy of satellites in simulations behaves like $\beta(r) \simeq 0.55 \, (r/r_{\rm vir})^{1/3}$ for $0.2r_{\rm vir} \leq r \leq r_{\rm vir}$. To estimate β for the Milky Way and M31 satellites, we calculate the weighted mean of this distribution

$$\bar{\beta} = \frac{\int_{0.2r_{\rm vir}}^{r_{\rm vir}} \beta(r) n(r) r^2 dr}{\int_{0.2r_{\rm vir}}^{r_{\rm vir}} n(r) r^2 dr} \tag{32}$$

where the weighting function $n(r)$ is the satellite number density distribution. This gives $\bar{\beta} = 0.44$ for the Milky Way and $\bar{\beta} = 0.45$ for M31. This is similar to the anisotropy of halo stars ($\beta = 0.37$)
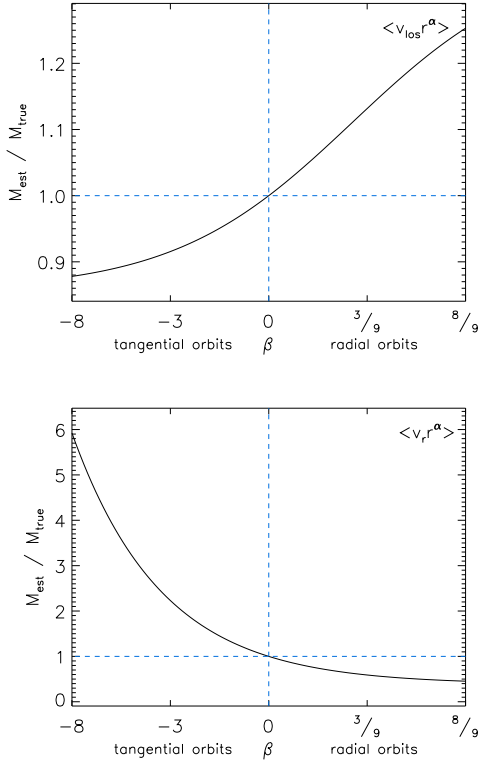
Figure 4. The sensitivity of the estimated mass on the anisotropy parameter β for a satellite population with $\alpha = 0.55$, $\beta = 0$ and $\gamma = 2.7$. The figure shows the mass recovered using the input values of α and γ and varying the value of β. The functional form of the curve is easy to deduce. It is a rational function of β for the upper panel, which uses the estimator of eqn (23), and a linear function of β for the lower panel, which uses eqn (16).

in simulations reported by Xue et al. (2008). Even though these numbers have the backing of simulations, they are somewhat surprising. Most of the Milky Way satellites with measured proper motions are moving on polar or tangential orbits. Using the sample of the 7 Milky Way satellites with proper motions, we can compute the radial and tangential components of the Galactocentric velocity. From these, the observed anisotropy $\beta \sim -4.5$, which favours tangential orbits. This is consistent with the earlier, though indirect, estimate of Wilkinson & Evans (1999), who found $\beta \sim -1$, again favouring tangential orbits. The origin of this discrepancy between simulations and data is not well understood. Perhaps there is considerable cosmic scatter in the anisotropy of the satellites, as it may depend on the details of the accretion history of the host galaxy. Figure 4 plots brings both good news and bad. The upper panel shows that the mass estimates for external galaxies using the line of sight estimator of eqn (23) are reasonably insensitive to the precise value of β. This make sense, as for a galaxy like M31, the line of sight velocity encodes information on both the radial and tangential velocity components referred to the M31's centre. However, in the case of the Milky Way, the situation is very different. The measured velocities provide information almost wholly on the radial component referred to the Galactic Center. In the absences of proper motions, the velocity anisotropy is largely unconstrained by the data. This is the classical mass-anisotropy degeneracy, and so – as the lower panel shows – there is considerable uncertainty in the mass estimates inferred using eqn (16).
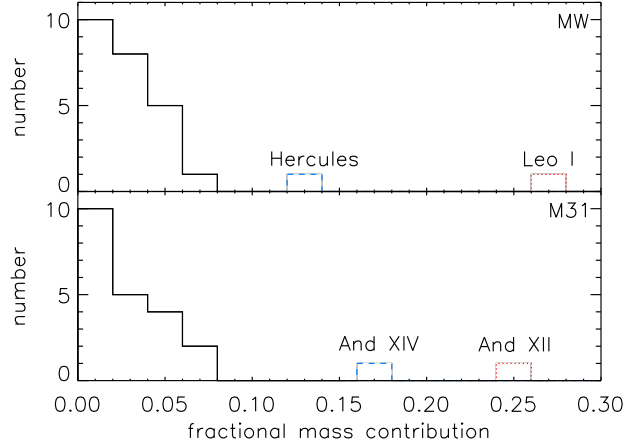

Figure 5. The fractional contribution each satellite makes to the mean mass estimator for the Milky Way (top) and M31 (bottom). For both galaxies, the mass budget is dominated by two satellites. For the Milky Way these are Leo I (red, dotted) and Hercules (blue, dashed). For M31, these are And XII (red, dotted) and And XIV (blue, dashed).

In what follows, we typically quote mass estimates for the anisotropies derived both from observations $\beta_{\rm data}$ and from simulations $\beta_{\rm sim}$, as well as for the case of isotropy ($\beta = 0$). In the absence of consistent indications to the contrary, our preference is to assume isotropy and to give greatest credence to the mass estimates obtained with this assumption.

4.2 Radial Velocity Datasets

Armed with values for α, β and γ, we now set the mass estimators to work. Data for the satellites of the Milky Way and M31 are given in Tables 1 and 2 respectively. Objects for which no line-of-sight velocity has been measured (And XVII, And XVIII, And XIX, And XX, And XXI and And XXII) are included in the tables, but excluded from the analysis.

Using eqn (16) and recalling that the Monte Carlo simulations gave errors of $\sim 25\%$, we give estimates of the mass with 100, 200 and 300 kpc for the Milky Way Galaxy in Table 3. Assuming velocity isotropy, we obtain for the mass of the Milky Way $M_{300} = 0.9 \pm 0.3 \times 10^{12} M_{\odot}$. The cussedness of the mass-anisotropy degeneracy is well illustrated by the fact that using the observationally derived $\beta_{\rm data}$ gives $M_{300} = 3.4 \pm 0.9 \times 10^{12} M_{\odot}$, whilst using that from simulations gives $M_{300} = 0.6 \pm 0.2 \times 10^{12} M_{\odot}$. The huge spread in mass estimates is due to the fact that the line-of-sight velocities for the satellites are almost entirely providing information on the radial velocities as judged from the Galactic Centre. There is almost no information on the tangential motions in our dataset. However, there are other astrophysical reasons why masses higher than $\sim 2 \times 10^{12} M_{\odot}$ are disfavoured.

Using eqn (23), we obtain the mass of M31 within 300 kpc as $M_{300} = 1.4 \pm 0.4 \times 10^{12} M_{\odot}$. Here, though, in sharp distinction to the case of the Milky Way, plausible changes in the velocity anisotropy generate modest changes of the order of 10 per cent in the mass estimate, as shown in Table 3. Of course, this is understandable, as the line-of-sight velocity now has information on both the radial and tangential components, albeit tangled up in the projection.

Taking the masses derived using velocity isotropy ($\beta = 0$), we note that this work hints at the removal of a long-standing puzzle,

Table 3. Enclosed mass within 100, 200 and 300 kpc for the Milky Way and Andromeda galaxies. We offer three estimates: one using the anisotropy inferred from data ($\beta \sim -4.5$), one assuming isotropy ($\beta = 0$) and the third with the anisotropy derived from simulations ($\beta \sim 0.45$).

Galaxy	M_{300} ($\times 10^{11} M_\odot$)			M_{200} ($\times 10^{11} M_\odot$)			M_{100} ($\times 10^{11} M_\odot$)		
	β_{data}	isotropic	β_{sim}	β_{data}	isotropic	β_{sim}	β_{data}	isotropic	β_{sim}
Milky Way	34.2 ± 9.3	9.2 ± 2.5	6.6 ± 1.8	21.1 ± 5.7	5.5 ± 1.6	3.8 ± 1.0	13.9 ± 4.9	3.3 ± 1.1	2.1 ± 0.7
... excl Leo I	25.2 ± 7.5	6.9 ± 1.8	5.0 ± 1.2	...	...	...	...	...	...
... excl Leo I, Her	21.1 ± 6.3	5.8 ± 1.5	4.2 ± 1.1	17.5 ± 5.3	4.6 ± 1.4	3.2 ± 0.8	...	...	...
MW with PMs	38.6 ± 7.0	24.1 ± 5.3	22.1 ± 5.3	28.3 ± 5.5	18.5 ± 4.2	17.0 ± 4.2	22.2 ± 5.0	13.8 ± 3.9	11.4 ± 3.1
... excl Draco	27.1 ± 4.9	14.1 ± 3.1	12.2 ± 2.7	18.1 ± 3.4	10.0 ± 2.3	8.7 ± 2.2	12.9 ± 3.1	6.9 ± 1.9	5.4 ± 1.5
... excl LMC/SMC	38.8 ± 6.8	24.6 ± 5.8	21.7 ± 4.9	28.3 ± 5.7	18.4 ± 4.3	16.3 ± 4.1	22.7 ± 5.6	13.9 ± 4.2	11.2 ± 3.4
... excl Draco, LMC/SMC	25.9 ± 5.1	12.4 ± 2.9	10.6 ± 2.5	17.0 ± 3.5	8.5 ± 2.2	7.2 ± 1.8	11.3 ± 2.9	5.4 ± 1.7	4.2 ± 1.3
M31	15.8 ± 3.3	14.1 ± 4.1	13.1 ± 3.8	15.4 ± 4.1	12.4 ± 3.8	10.6 ± 3.5	2.6 ± 1.0	2.1 ± 1.0	1.8 ± 1.0
... excl AndXII	12.2 ± 2.7	10.9 ± 3.1	10.1 ± 3.2	11.4 ± 3.2	9.2 ± 3.1	7.9 ± 2.8	...	...	...
... excl AndXII, AndXIV	9.6 ± 2.1	8.5 ± 2.4	8.0 ± 2.4	8.6 ± 2.6	6.9 ± 2.4	5.9 ± 2.2	...	...	...
M31 with PMs	15.1 ± 3.8	13.9 ± 3.5	13.1 ± 3.5	...	...	...	...	...	...

namely that the kinematic data on the satellite galaxies suggested that M31 was less massive than the Milky Way, whereas other indicators (such as the total numbers of globular clusters or the amplitude of the gas rotation curve) suggested the reverse. In fact, with the new datasets, the ratio of the masses of M31 to the Milky Way (~ 1.5) is close to that which would be inferred using the Tully-Fisher relationship and the assumption that the luminosity is proportional to the total mass ($250^4/220^4 \approx 1.67$). If instead the radial anisotropies derived from simulations are preferred, then the ratio is ~ 1.98.

However, it may be imprudent to include all the satellites. For example, Leo I has long been known to dominate mass estimates of the Milky Way, on account of its large distance (~ 260 kpc) and high line-of-sight velocity (see e.g. Kulessa & Lynden-Bell 1992; Kochanek 1996; Wilkinson & Evans 1999). It is unclear that Leo I is actually on a bound orbit, as opposed to a hyperbolic one. Hence, many attempts at determining the mass of the Milky Way quote estimates both including and excluding Leo I. In fact, recent photometric and spectroscopic evidence presented by Sohn et al. (2007) favours the picture in which Leo I is bound on an orbit with high eccentricity (~ 0.95) and small perigalacticon (10-15 kpc). In particular, such models give good matches to the surface density and radial velocity dispersion profiles of Leo I, and imply high mass estimates for the Milky Way. However, Sales et al. (2007) using simulations found a population of satellite galaxies on extreme orbits ejected from haloes as a result of three-body slingshot effects, and suggested that Leo I might be an example of such an object. So, although the present evidence favours a bound orbit, a definitive verdict must await the measurement of Leo I's proper motion by the *Gaia* satellite, which should resolve the issue.

Given that there is one satellite that is known to inflate the Milky Way's mass, it is interesting to investigate whether any of the other satellites, particularly the recent discoveries, play similar rôles. The upper panel of Figure 5 shows the fractional contributions each satellite makes to the Milky Way's mass ($C v_{\text{los}}^2 r^\alpha /(GN)$) – it is the total of these values that we take to be the mass estimate. There are two clear outliers; the outermost satellite in this distribution is Leo I, the less extreme satellite is Hercules. Like Leo I, Hercules has a substantial radial velocity and a relatively large Galactocentric distance (~ 130 kpc). Hercules has a highly elongated, irregular and flattened structure (Belokurov et al. 2007; Coleman et al. 2007). This is consistent with tidal disruption during pericentric passages on a highly eccentric orbit ($e > 0.9$). This seems good evidence that Hercules is truly bound to the Milky Way.

We repeat the same analysis for M31 and the results are shown in the bottom panel of Figure 5. Interestingly, we see that there are two outliers in the distribution, namely two of the recent discoveries, And XII and And XIV. Notice that though both objects have a substantial effect on M31's mass estimate, neither are as extreme as Leo I. It is the inclusion of these two new objects in the satellite dataset that has augmented the mass of M31, so that it is now somewhat greater than that of the Milky Way.

But, this begs the question: should these satellites be included? And XIV was discovered by Majewski et al. (2007) in a survey of the outer M31 stellar halo. They recognized its extreme dynamical properties and suggested that it may either be falling into M31 for the first time or that M31's mass must be larger than hitherto estimated by virial arguments. In fact, And XIV's lack of gas and its elongated structure suggest that ram pressure stripping and tidal effects may have been important in its evolution. This is consistent with And XIV being a true satellite of M31 that has already suffered a pericentric passage, a conclusion that could be strengthened with deeper imaging, which might reveal the presence of tidal tails around And XIV.

And XII is a still more ambiguous object – it was discovered as a stellar overdensity by Martin et al. (2006). Spectroscopic observations were subsequently taken by Chapman et al. (2007), who conjectured that the satellite might be falling into the Local Group for the first time. The evidence for this is its large velocity and its likely location behind M31. However, it remains unclear whether this evolutionary track is consistent with the absence of detection of HI gas in the object. Pristine, infalling dwarfs, which have not yet experienced a pericentric passage of 50 kpc or less, should retain sizeable amounts of neutral HI gas, whereas Chapman et al. (2007) constrain the mass in HI to be less than $3 \times 10^3 M_\odot$.

In light of this, we provide more mass estimates, after removing possible ambiguous objects (and re-computing the parameter γ where necessary). For the Milky Way, we exclude Leo I only and then Leo I and Hercules. For M31, we exclude And XII only and then And XII and And XIV. These mass estimates are also shown in Table 3. Note that, for example, the exclusion of Leo I does not change the mass estimate within 100 or 200 kpc, as Leo I is outside of this range. Similarly, And XII and And XIV lie outside of 100 kpc from the center of M31, so the mass estimates without them do not change the final column of the table.
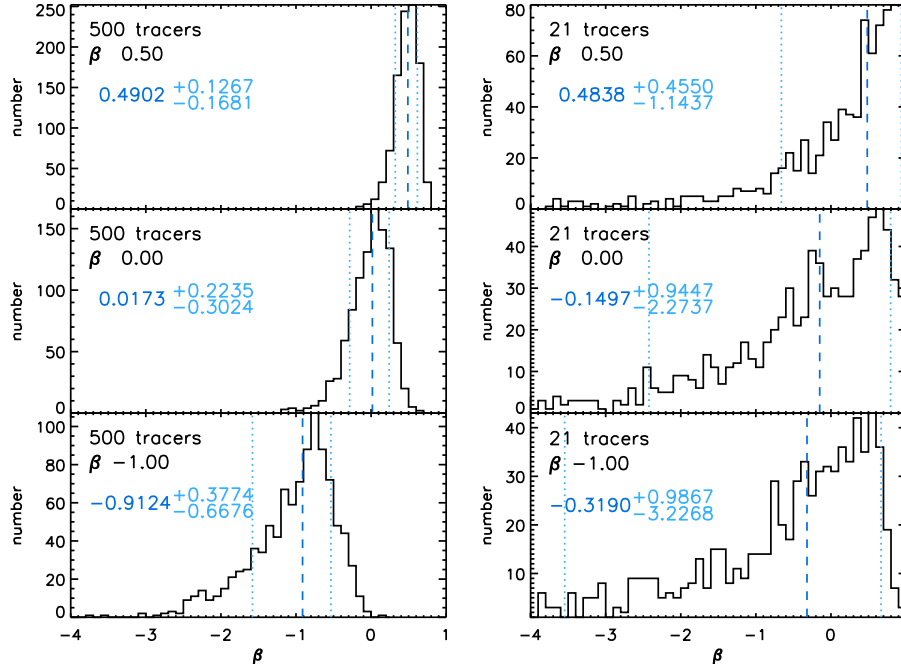
Figure 6. Distribution of β obtained from simultaneous mass and velocity anisotropy fitting. The median of each distribution is shown as a blue dashed line and the 68 % confidence limits as cyan dotted lines. These values are also given in blue and cyan on the plot. The left panel shows the idealised case for 500 tracers, the right panel the case tailored to our M31 data where only 21 tracers are used. The value of β used to generate the tracer population is also shown.

In the case of velocity isotropy ($\beta = 0$), it requires the excision of both And XII and And XIV from the datasets for the mass estimate of M31 to become comparable to or smaller than the Milky Way. For example, the mass of M31 with And XII and And XIV both removed is $0.85 \pm 0.24 \times 10^{12} M_\odot$, as compared to the mass of the Milky Way with Leo I retained of $0.92 \pm 0.25 \times 10^{12} M_\odot$. However, we have argued that And XIV is most likely bound, whilst And XII is a more ambiguous case. In other words, the problem pointed to by Evans & Wilkinson (2000) – namely that the mass of M31 inferred from the kinematics of the satellites is less than the mass of the Milky Way – has indeed been ameliorated by the discovery of more fast-moving M31 satellites.

It seems particularly intriguing that such satellites exist for both the Milky Way and M31. Wilkinson & Evans (1999) used virialized models to estimate that the probability that, in a sample of 30 satellites, there is an object like Leo I, which changes the mass estimate by a factor of a few. They found that the probability is minute, only $\sim$ 0.5%. Prior expectation does not favour the existence of objects like Leo I or And XII, yet in fact, both big galaxies in the Local Group possess such satellites. The clear conclusion is that the satellites in the outer parts of these galaxies cannot all be virialized. This is a point in favour of processes such as those advocated by Sales et al. (2007) to populate such orbits.

4.3 Simultaneous Solution for Mass and Anisotropy

There is one further way in which the estimators can be set to work with the line-of-sight velocities. When three dimensional positions and projected positions are simultaneously available – as for example in the case of M31's satellites – it is possible to use the estimators based on both the $\langle v_{los}^2 r^\alpha \rangle$ and the $\langle v_{los}^2 R^\alpha \rangle$ moments to solve simultaneously for both the total mass and the anisotropy parame-

Table 4. Table of proper motion data for the satellites of the Milky Way and M31. Listed are equatorial proper motions in mas century^{-1}.

Name	$\mu_\alpha \cos\delta$ (mas/century)	μ_δ (mas/century)	Source
Carina	22 ± 9	15 ± 9	1
Draco	60 ± 40	110 ± 30	2
Fornax	48 ± 5	-36 ± 4	3
LMC/SMC	198 ± 5	25 ± 5	4
Sculptor	9 ± 13	2 ± 13	5
Sextans	-26 ± 41	10 ± 44	6
Ursa Minor	-50 ± 17	22 ± 16	7
M33	2.1 ± 0.7	2.5 ± 1.2	8
IC10	-0.2 ± 0.8	2.0 ± 0.8	9
M31	2.1 ± 1.1	-1.0 ± 0.9	10

Sources: 1 - Piatek et al. (2003), 2 - Scholz & Irwin (1994), 3 - Piatek et al. (2007), 4 - Piatek et al. (2008), 5 - Piatek et al. (2006), 6 - Walker et al. (2008), 7 - Piatek et al. (2005), 8 - Brunthaler et al. (2007a), 9 - Brunthaler et al. (2007b), 10 - van der Marel & Guhathakurta (2008), though unlike the other proper motions, this not a measurement but inferred from indirect evidence.

ter. There is however no guarantee that the solution for β is in the physical range $-\infty \leq \beta \leq 1$.

The success of this procedure of course rests on the accuracy of the data. The distances of the M31 satellites are determined by the tip of the red giant branch method and have errors of ± 30 kpc (see e.g., McConnachie et al. 2005). If we use eqns (23) and (26), and simultaneously solve for the unknowns, we obtain

$$M_{300} = 1.5 \pm 0.4 \times 10^{12} M_\odot, \qquad \beta = -0.55^{+1.1}_{-3.2} \qquad (33)$$
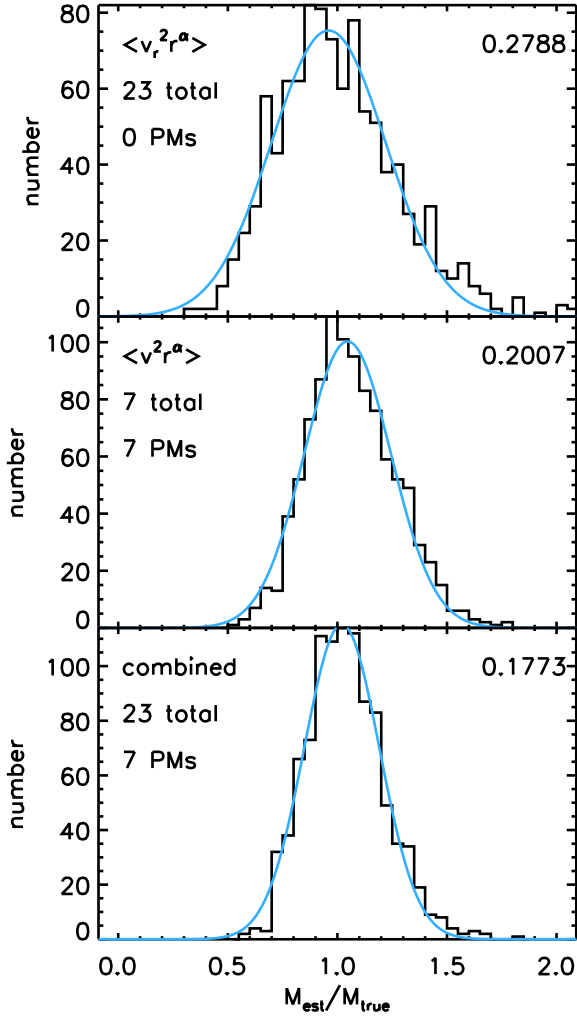
which corresponds to mild tangential anisotropy. These are surprisingly sensible answers given the distance errors.

Fig. 6 is inferred from Monte Carlo simulations and shows the distributions of anisotropy parameters derived from simultaneous mass and anisotropy fitting for mock datasets. Also given in the panels are the median and 68 per cent confidence limits for the anisotropy parameter, in the case of 21 satellite galaxies (comparable to the present dataset for M31) and the case of 500 satellites. Although with 21 tracers, the errors on the anisotropy parameter are substantial, matters improve significantly with larger numbers of tracers. A dataset of 500 halo satellites (dwarf galaxies, globular clusters and planetary nebulae) is not unreasonable for a galaxy like M31 in the near future. This raises the possibility that the method of simultaneous fitting may prove more compelling in the future. In fact, given 500 tracers, it is reasonable to use the estimators based on both the $\langle v_{\rm los}^2 r^\alpha \rangle$ and the $\langle v_{\rm los}^2 R^\alpha \rangle$ moments to fit simultaneously at each distance, thus giving the run of anisotropy parameter and mass with radius.

4.4   Radial and Proper Motion Datasets

Thus far, we have used only the line of sight velocities to make mass estimates. In this section, we add in the proper motions of satellites, where available. Thus, for the Milky Way galaxy, we combine results from eqn (16) for satellites without proper motions and from eqn (24) for those with proper motions, weighting each estimate by the reciprocal of the standard deviation to give the final answer.

Proper motions, albeit with large error bars, have been measured for a total of 9 of the Milky Way satellite galaxies. It seems prudent to exclude Sagittarius, as it is in the process of merging with the Milky Way. Additionally, the interacting Magellanic Clouds are treated as a single system by computing the proper motion of their mass centroid, taking the masses of the LMC and SMC as $\sim 2 \times 10^{10} M_\odot$ and $2 \times 10^9 M_\odot$ respectively (Kroupa & Bastian 1997). This leaves us with a set of 7 satellites with proper motion data, summarized in Table 4. In most cases, errors on proper motions are large and, where multiple studies exist, the measurements are sometimes in disagreement. The proper motions inferred by ground based methods are in reasonable agreement with those derived from the *Hubble Space Telescope* (*HST*) in the cases of Fornax (Piatek et al. 2007; Walker et al. 2008), Carina (Piatek et al. 2003; Walker et al. 2008) and the Magellanic Clouds (Piatek et al. 2008; Costa et al. 2009). But, for Ursa Minor (Scholz & Irwin 1994; Piatek et al. 2005) and for Sculptor (Piatek et al. 2006; Walker et al. 2008), agreement between different investigators is not good, and we have preferred to use the estimates derived from *HST* data. Nonetheless, it is important to include the proper motion data, especially for mass estimates of the Milky Way Galaxy. We use these proper motions along with distance and line-of-sight velocity data to calculate full space velocities for these satellites, as described in Piatek et al. (2002).

In addition, there are two satellites of M31 with measured proper motions, namely M33 and IC 10. This astonishing feat has exploited the *Very Long Baseline Array* to measure the positions of water masers relative to background quasars at multiple epochs (Brunthaler et al. 2005, 2007b). Unfortunately, the technique cannot be extended to M31 itself, as it does not contain any water masers, and so its proper motion is much less securely known. However, van der Marel & Guhathakurta (2008) reviewed the evidence from a number of sources – including kinematics of the M31 satellites, the motions of the satellites at the edge of the Local Group, and the constraints imposed by the tidal distortion of

Figure 7. Distribution of mass estimates as a fraction of true mass for Monte Carlo simulations using (top) 23 satellites with radial velocities, (middle) 7 satellites with proper motions and (bottom) 23 satellites, 7 of which have proper motions. The standard deviation of the best fitting Gaussian is shown in the top-right hand corner of each panel. [These plots assume $\beta = -4.51$, as estimated from the data].

Figure 8. The fractional contribution each satellite with proper motions makes to the mean mass estimate for the Milky Way Galaxy. Notice the extreme effect of Draco's proper motion.

M33's disk – to provide a best estimate. These data are also listed in Table 4.

The Milky Way satellites are so remote that their line-of-sight velocities in the Galactic rest frame are almost identical to their radial velocities, as judged form the Galactic Centre. The proper motion data provide genuinely new information on the tangential motions and this is the only way to break the mass-anisotropy degeneracy. The same argument does not hold with equal force for M31, as the line of sight velocities incorporate contributions from both the radial and tangential components as reckoned from the centre of M31. Nonetheless, it is good practice to use all the data available, even though the proper motions of M33 and IC 10 with respect to the M31 reference frame must be inferred using an estimate of M31's proper motion (rather than a measurement).

For the satellites without proper motions, we use the form of the estimator given in eqns (16) or (23) for the Milky Way and M31 respectively; for those with proper motions, we use eqn (24). We combine results from the two estimators, weighting each estimate by the reciprocal of the standard deviation to give the final answer. To infer the standard deviation, we perform Monte Carlo simulations. So, for the case of the Milky Way, we generate mock datasets of 25 satellites, for which only 7 have proper motions. The errors on radial velocities are dwarfed by the uncertainty caused by the small number statistics and so are neglected. But, the errors on the proper motions are not negligible and they are incorporated into the simulations by adding a value selected at random from the range [-0.5 μ, 0.5 μ], where μ is the proper motion. The flat distribution has been chosen as systematic errors are as important as random Gaussian error in the determination of proper motions. However, we have tested alternatives in which we use the relative observational errors, or the relative observational errors multiplied by 2.5, and find that our results are robust against changes to the error law. The standard deviations of the fractional mass distribution the satellites with and without proper motions are separately computed, as illustrated in the panels of Figure 7. We linearly combine the mass estimates, weighting with the reciprocal of the standard deviation, to give the final values reported in Table 3.

Given that the Milky Way satellites with measured proper motions are moving on polar orbits, it is no surprise that the mass estimate of the Milky Way has now increased. Adopting the value of β we estimate from the data, we find $M_{300} = 3.9 \pm 0.7 \times 10^{12} M_\odot$ for the Milky Way Galaxy and $M_{300} = 1.5 \pm 0.4 \times 10^{12} M_\odot$ for M31. Assuming isotropy, we find $M_{300} = 2.5 \pm 0.5 \times 10^{12} M_\odot$ for the Milky Way Galaxy and $M_{300} = 1.4 \pm 0.4 \times 10^{12} M_\odot$ for M31. Notice however, the mass estimate for M31 has barely changed from the value inferred from the full radial velocity dataset.

Again, we calculate the contribution that each satellite makes to the mass estimate to investigate whether any are dominating the final answer. First, this procedure guards against the possibility of a completely rogue proper motion measurement. Second, there are some suggestions that the Magellanic Clouds may not be bound, or even if bound may only be on its second passage and so may not be part of the relaxed distribution of satellite galaxies (Besla et al. 2007). So, it is helpful to check that our results are not unduly sensitive to its inclusion. As Figure 8 shows, we find that Draco is a clear outlier and nearly doubles the Milky Way mass estimate. If we remove the Draco proper motion from the sample, we instead recover a mass $M_{300} = 2.7 \pm 0.5 \times 10^{12} M_\odot$ (assuming β_{data}) or $M_{300} = 1.4 \pm 0.3 \times 10^{12} M_\odot$ (assuming isotropy). It is particularly concerning that the proper motion of Draco has such a substantial effect, because – as judged from the size of the error bars in Table 4 – it is one of the noisier measurements. By contrast, the exclusion of the

Magellanic Clouds has only a minor effect, as is evident from the results listed in Table 3.

We have covered a number of possibilities, so it is probably useful for us to give our best estimates. On balance, we think the case for including at least And XIV among the satellite sample for Andromeda is strong. Whilst And XII is a more ambiguous case, the lack of any HI gas suggests to us that it should also be included. Among the satellites of the Milky Way, we favour including Leo I based on the work of Sohn et al. (2007), whilst we are inclined to discard the proper motion of Draco reported in Scholz & Irwin (1994) until corroborated. Until the discrepancy between the velocity anisotropies reported in simulations and in data is explained, we prefer to use the data as our guide

So, our best estimate for the mass of the Milky Way within 300 kpc is

$$M_{300} \sim 2.7 \pm 0.5 \times 10^{12} M_\odot \tag{34}$$

whilst for M31, it is

$$M_{300} \sim 1.5 \pm 0.4 \times 10^{12} M_\odot. \tag{35}$$

These estimates are obtained using the combined radial velocity and proper motion datasets. The error bars only incorporate the statistical uncertainty. As we have emphasised, there are much greater uncertainties induced by selection of satellite members and velocity anisotropy. In particular, when these uncertainties are considered, it is not possible to decide which of the Milky Way or M31 is more massive based on satellite kinematic data alone.

5   DISCUSSION

It is instructive to compare our results with a number of recent estimates of the masses of the Local Group and its component galaxies. Xue et al. (2008) extracted a sample of ~ 2400 blue horizontal branch stars from the SDSS. These are all resident in the inner halo within 60 kpc of the Galactic centre. This has the advantage that the BHBs are surely virialized, but the disadvantage that no inference can be made about the mass exterior to 60 kpc. Hence, any estimate as to the total mass is driven wholly by prior assumptions rather than the data. In fact, Xue et al. (2008) assumed an NFW halo with a canonical concentration holds, and then estimated the virial mass of the Milky Way's dark matter halo as $M = 1.0^{+0.3}_{-0.2} \times 10^{12} M_\odot$, using Jeans modelling with an anisotropy parameter inferred from numerical simulations. This is lower than our preferred value, but in good agreement with our comparable calculations using line of sight velocity datasets alone.

A somewhat similar calculation for M31 has been reported by Seigar et al. (2008). The mass of the baryonic material is estimated using a Spitzer 3.6 μm image of the galaxy, together with a mass-to-light ratio gradient based on the galaxy's $B - R$ colour. This is combined with an adiabatically-contracted NFW halo profile to reproduce the observed HI rotation curve data. They find a total virial mass of M31's dark halo as $8.2 \pm 0.2 \times 10^{11} M_\odot$. This is lower than most of our estimates, with the exception of those based on samples excluding both And XII and And XIV.

Although these calculations are interesting, it is worth remarking that the final masses are not wholly controlled by the data. We know that, from Newton's theorem, any mass distribution outside the limiting radius of our data has no observational effect in a spherical or elliptical system. To estimate the virial mass from data confined to the inner parts (such as BHBs or the optical disk) requires an understanding of the structure of the pristine dark halo initially,

as well as how it responds to the formation of the luminous baryonic components. It is this that controls the final answer.

Li & White (2008) used the Millennium Simulation to extract mock analogues of the Local Group and calibrate the bias and error distribution of the Timing Argument estimators (see e.g., Kahn & Woltjer 1959; Raychaudhury & Lynden-Bell 1989). From this, they obtain a total mass of the two large galaxies in the Local Group of $5.3 \times 10^{12} M_\odot$ with an inter-quartile range of $[3.8 \times 10^{12}, 6.8 \times 10^{12}] \, M_\odot$ and a 95 % confidence lower limit of $1.8 \times 10^{12} M_\odot$. Importantly, Li & White (2008) showed that the mass estimate from the timing argument is both unbiased and reasonably robust. This is a considerable advance, as there have long been worries that the gross simplification of two-body dynamics implicit in the original formulation of the Timing Argument may undermine its conclusions.

It therefore seems reasonable to assume that the combined mass of the Milky Way Galaxy and M31 is at least $3.8 \times 10^{12} M_\odot$, and perhaps more like $5.3 \times 10^{12} M_\odot$. The low estimates of the Milky Way and M31 masses of Xue et al. (2008) and Seigar et al. (2008) are not compatible with this, and barely compatible with Li & White's 95 % lower limit. Using our preferred values in eqns (34) and (35), the combined mass in the Milky Way and M31 galaxies is $4.2 \pm 0.6 \times 10^{12} M_\odot$. This is comparable to the $3.8 \times 10^{12} M_\odot$ of Li & White.

Li & White (2008) also estimated a virial mass for the Milky Way of $2.4 \times 10^{12} M_\odot$ with a range of $[1.1 \times 10^{12}, 3.1 \times 10^{12}] \, M_\odot$, based on timing arguments for Leo I. Given all the uncertainties, this is in remarkable accord with our best estimate.

6 CONCLUSIONS

We have derived a set of robust tracer mass estimators, and discussed the conditions under which they converge. Given the positions and velocities of a set of tracers – such as globular clusters, dwarf galaxies or stars – the estimators compute the enclosed mass within the outermost datapoints. The accuracy of the estimator has been quantified with Monte Carlo simulations. The estimators are applicable to a wide range of problems in contemporary astrophysics, including measuring the masses of elliptical galaxies, the haloes of spiral galaxies and galaxy clusters from tracer populations. They are considerably simpler to use than distribution function based methods (see e.g. Little & Tremaine 1987; Kulessa & Lynden-Bell 1992; Wilkinson & Evans 1999), and involve no more calculation than taking weighted averages of combinations of the positional and kinematical data. They should find widespread applications.

The mass estimators are applied to the satellite populations of the Milky Way and M31 to find the masses of both galaxies within 300 kpc. These estimates are the first to make use of the recent burst of satellite discoveries around both galaxies. Both satellite populations have nearly doubled in size since previous estimates were made. We summarise our results by answering the questions; What are (1) the minimum, (2) the maximum and (3) the most likely masses of the Milky Way and M31 galaxies?

(1) The mass of the Milky Way Galaxy within 300 kpc could be as low as $0.4 \pm 0.1 \times 10^{12} M_\odot$. This would imply that Leo I is gravitationally unbound, contrary to the recent evidence provided by by Sohn et al. (2007). Leo I would then be either an interloper or an object being ejected from the Milky Way by an encounter. It would also require that the proper motion of Draco (Scholz & Irwin 1994) is incorrect, which is not inconceivable given the difficulty of the

measurements. It implies that the satellite galaxies are moving on radial orbits and so the velocity anisotropy is radial.

The mass of M31 within 300 kpc could plausibly be as low as $0.8 \pm 0.2 \times 10^{12} M_\odot$. This would be the case if both And XII and And XIV are not gravitationally bound, which is possible if mechanisms such as those proposed by Sales et al. (2007) are ubiquitous. It would also require that the proper motion data on M33 and IC10 or – perhaps more likely – the indirectly inferred proper motion of M31 is in error. Again, such a low estimate for the mass occurs only if the satellites are moving predominantly radially.

Although it is interesting to ask how low the masses of the Milky Way and M31 could be, it does produce a mystery in the context of the Timing Argument, which typically yields larger combined masses. It is possible that some of the mass of the Local Group is unassociated with the large galaxies. Although not the conventional picture, this is probably not ruled out and there have been suggestions that $\sim 10^{12} M_\odot$ may be present in the Local Group in the form of baryons in the warm-hot intergalactic medium (Nicastro et al. 2003). There are few constraints on the possible existence of dark matter smeared out through the Local Group, and unassociated with the large galaxies. However, the clustering of the dwarf galaxies around the Milky Way and M31 does suggest that the gravity of the dark matter is centered on the prominent galaxies.

(2) The largest mass estimate we obtained for the Milky Way Galaxy is $3.9 \pm 0.7 \times 10^{12} M_\odot$. This extreme values is driven by the assumption of tangential anisotropy for the satellites, so that the measured line of sight velocities also imply substantial tangential motions as well. The estimate assumes all the satellites including Leo I to be bound, and the anomalously high proper motion measurement of Draco to be valid.

Note that the present data allow considerably more scope to increase the mass of the Milky Way Galaxy than M31. Our largest mass estimate for M31 is a much more modest $1.6 \pm 0.4 \times 10^{12} M_\odot$, which occurs when we analyse the whole sample incorporating And XII and And XIV and assume tangentially anisotropic velocity distributions.

The current concensus is that the two galaxies are of a roughly similar mass, with M31 probably the slightly more massive of the two. This though is inferred from indirect properties, such as the numbers of globular clusters, which correlates with total mass albeit with scatter, or the amplitude of the inner gas rotation curve. The stellar halo of M31 is certainly more massive than that of the Milky Way, although this may not be a good guide to the dark halo. Of course, it could be that the current concensus is wrong, and that the Milky Way halo is more massive than that of Andromeda. There is also some indirect evidence in favour of this – for example, the typical sizes of the M31 dwarf spheroidals are large than those of the Milky Way, which is explicable if the Milky Way halo is denser. However, it does not seem reasonable to postulate that the mass of the Milky Way is substantially larger than that of M31. Hence, the very large estimate of $3.9 \pm 0.7 \times 10^{12} M_\odot$ is best understood as a manifestation of the degeneracy in the problem of mass estimation with only primarily radial velocity data.

(3) Our preferred estimates come from accepting Leo I, And XII and And XIV as bound satellites, whilst discarding the Draco proper motion as inaccurate. This gives an estimate for the mass of the Milky Way within 300 kpc as $2.7 \pm 0.5 \times 10^{12} M_\odot$ and for M31 as $1.5 \pm 0.4 \times 10^{12} M_\odot$, assuming the anisotropy implied by the data ($\beta \approx -4.5$). The error bars are just the statistical uncertainty and do not incorporate the uncertainty in anisotropy or sample member-

ship. In view of this, it is not possible to decide which of the Milky Way galaxy or M31 is the more massive based on the kinematic data alone.

These values for the masses are attractive for a number of reasons. First, the mass ratio between the Milky Way and M31 is of roughly of order unity, which accords with a number of other lines of evidence. Second, the values allow most of the dark matter in the Local Group implied by the Timing Argument to be clustered around the two most luminous galaxies. Third, they are within the range found for cosmologically motivated models of the Milky Way and M31 (Li & White 2008).

We prefer to assume the anisotropy implied by the admittedly scanty data on the proper motions of the satellites. However, for completeness, we quickly sketch the effects of dropping this assumption. If the velocity distribution is isotropic, or even radially anisotropic as suggested by the simulations, then the mass of the Milky Way becomes $1.4 \pm 0.3 \times 10^{12} M_\odot$ or $1.2 \pm 0.3 \times 10^{12} M_\odot$ respectively. Similarly for M31, the values are $1.4 \pm 0.4 \times 10^{12} M_\odot$ (isotropy) or $1.3 \pm 0.4 \times 10^{12} M_\odot$ (radially anisotropic).

The greatest sources of uncertainty on the masses remain the role of possibly anomalous satellites like Leo I and the velocity anisotropy of the satellite populations. There is reason to be optimistic, as the *Gaia* satellite will provide proper motion data on all the dwarf galaxies that surround the Milky Way and M31, as well as many hundreds of thousands of halo stars. The analysis that we have carried out here indicates that proper motions are important if we wish to increase the accuracy of our estimates, as well as understand the dynamical nature of objects like Leo I. While we are not yet able to exploit the proper motions, *Gaia* will allow us to do so.

REFERENCES

Armandroff T. E., Olszewski E. W., Pryor C., 1995, AJ, 110, 2131
Bahcall J. N., Tremaine S., 1981, ApJ, 244, 805
Bellazzini M., Ferraro F. R., Origlia L., Pancino E., Monaco L., Oliva E., 2002, AJ, 124, 3222
Bellazzini M., Gennari N., Ferraro F. R., 2005, MNRAS, 360, 185
Bellazzini M., Gennari N., Ferraro F. R., Sollima A., 2004, MNRAS, 354, 708
Belokurov V. et al., 2008, ApJ, 686, L83
Belokurov V. et al., 2009, MNRAS, p. 903
Belokurov V. et al., 2007, ApJ, 654, 897
Belokurov V. et al., 2006, ApJ, 647, L111
Besla G., Kallivayalil N., Hernquist L., Robertson B., Cox T. J., van der Marel R. P., Alcock C., 2007, ApJ, 668, 949
Binney J., Tremaine S., 1987, Galactic dynamics.Princeton, NJ, Princeton University Press, 1987
Braun R., 1991, ApJ, 372, 54
Brunthaler A., Reid M. J., Falcke H., Greenhill L. J., Henkel C., 2005, Science, 307, 1440
Brunthaler A., Reid M. J., Falcke H., Henkel C., Menten K. M., 2007a, arXiv:0708.1704
Brunthaler A., Reid M. J., Falcke H., Henkel C., Menten K. M., 2007b, A&A, 462, 101
Chapman S. C. et al., 2007, ApJ, 662, L79
Cioni M.-R. L., van der Marel R. P., Loup C., Habing H. J., 2000, A&A, 359, 601
Cole A. A. et al., 1999, AJ, 118, 1657
Coleman M. G. et al., 2007, ApJ, 668, L43
Costa E., Méndez R. A., Pedreros M. H., Moyano M., Gallart C., Noël N., Baume G., Carraro G., 2009, AJ, 137, 4339
Crain R. A. et al., 2009, MNRAS, 399, 1773
Dehnen W., Binney J. J., 1998, MNRAS, 298, 387
Diemand J., Kuhlen M., Madau P., 2007, ApJ, 667, 859
Evans N. W., Hafner R. M., de Zeeuw P. T., 1997, MNRAS, 286, 315
Evans N. W., Wilkinson M. I., 2000, MNRAS, 316, 929
Evans N. W., Wilkinson M. I., Guhathakurta P., Grebel E. K., Vogt S. S., 2000, ApJ, 540, L9
Evans N. W., Wilkinson M. I., Perrett K. M., Bridges T. J., 2003, ApJ, 583, 752
Fellhauer M. et al., 2006, ApJ, 651, 167
Freedman W. L. et al., 2001, ApJ, 553, 47
Geha M., Willman B., Simon J. D., Strigari L. E., Kirby E. N., Law D. R., Strader J., 2009, ApJ, 692, 1464
Gottesman S. T., Hunter J. H., Boonyasait V., 2002, MNRAS, 337, 34
Harris J., Zaritsky D., 2006, AJ, 131, 2514
Heisler J., Tremaine S., Bahcall J. N., 1985, ApJ, 298, 8
Helmi A., 2004, ApJ, 610, L97
Ibata R., Martin N. F., Irwin M., Chapman S., Ferguson A. M. N., Lewis G. F., McConnachie A. W., 2007, ApJ, 671, 1591
Ibata R. A., Wyse R. F. G., Gilmore G., Irwin M. J., Suntzeff N. B., 1997, AJ, 113, 634
Irwin M. J. et al., 2007, ApJ, 656, L13
Irwin M. J., Ferguson A. M. N., Huxor A. P., Tanvir N. R., Ibata R. A., Lewis G. F., 2008, ApJ, 676, L17
Johnston K. V., Law D. R., Majewski S. R., 2005, ApJ, 619, 800
Kahn F. D., Woltjer L., 1959, ApJ, 130, 705
Kaluzny J., Kubiak M., Szymanski M., Udalski A., Krzeminski W., Mateo M., 1995, A&AS, 112, 407
Karachentsev I. D., Karachentseva V. E., Huchtmeier W. K., Makarov D. I., 2004, AJ, 127, 2031
Karachentsev I. D., Kashibadze O. G., Makarov D. I., Tully R. B., 2009, MNRAS, 393, 1265
Klypin A., Zhao H., Somerville R. S., 2002, ApJ, 573, 597
Koch A., Kleyna J. T., Wilkinson M. I., Grebel E. K., Gilmore G. F., Evans N. W., Wyse R. F. G., Harbeck D. R., 2007, AJ, 134, 566
Koch A., Wilkinson M. I., Kleyna J. T., Gilmore G. F., Grebel E. K., Mackey A. D., Evans N. W., Wyse R. F. G., 2007, ApJ, 657, 241
Koch A. et al., 2009, ApJ, 690, 453
Kochanek C. S., 1996, ApJ, 457, 228
Kroupa P., Bastian U., 1997, New Astronomy, 2, 77
Kulessa A. S., Lynden-Bell D., 1992, MNRAS, 255, 105
Law D. R., Majewski S. R., Johnston K. V., 2009, ApJ, 703, L67
Letarte B. et al., 2009, arXiv:0901.0820
Li Y., White S. D. M., 2008, MNRAS, 384, 1459
Little B., Tremaine S., 1987, ApJ, 320, 493
Lynden-Bell D., Frenk C. S., 1981, The Observatory, 101, 200
Majewski S. R. et al., 2007, ApJ, 670, L9
Martin N. F., de Jong J. T. A., Rix H.-W., 2008, ApJ, 684, 1075
Martin N. F., Ibata R. A., Chapman S. C., Irwin M., Lewis G. F., 2007, MNRAS, 380, 281
Martin N. F., Ibata R. A., Irwin M. J., Chapman S., Lewis G. F., Ferguson A. M. N., Tanvir N., McConnachie A. W., 2006, MNRAS, 371, 1983
Martin N. F. et al., 2009, arXiv:0909.0399
Mateo M. L., 1998, ARA&A, 36, 435
McConnachie A. W. et al., 2008, ApJ, 688, 1009

McConnachie A. W., Irwin M. J., 2006, MNRAS, 365, 902

McConnachie A. W., Irwin M. J., Ferguson A. M. N., Ibata R. A., Lewis G. F., Tanvir N., 2005, MNRAS, 356, 979

Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563

Nicastro F. et al., 2003, Nature, 421, 719

Piatek S., Pryor C., Bristow P., Olszewski E. W., Harris H. C., Mateo M., Minniti D., Tinney C. G., 2005, AJ, 130, 95

Piatek S., Pryor C., Bristow P., Olszewski E. W., Harris H. C., Mateo M., Minniti D., Tinney C. G., 2006, AJ, 131, 1445

Piatek S., Pryor C., Bristow P., Olszewski E. W., Harris H. C., Mateo M., Minniti D., Tinney C. G., 2007, AJ, 133, 818

Piatek S., Pryor C., Olszewski E. W., 2008, AJ, 135, 1024

Piatek S. et al., 2002, AJ, 124, 3198

Piatek S., Pryor C., Olszewski E. W., Harris H. C., Mateo M., Minniti D., Tinney C. G., 2003, AJ, 126, 2346

Queloz D., Dubath P., Pasquini L., 1995, A&A, 300, 31

Raychaudhury S., Lynden-Bell D., 1989, MNRAS, 240, 195

Rohlfs K., Kreitschmann J., 1988, A&A, 201, 51

Sakai S., Madore B. F., Freedman W. L., 1999, ApJ, 511, 671

Sales L. V., Navarro J. F., Abadi M. G., Steinmetz M., 2007, MNRAS, 379, 1475

Saviane I., Held E. V., Bertelli G., 2000, A&A, 355, 56

Scholz R.-D., Irwin M. J., 1994, in MacGillivray H. T., ed., IAU Symposium Vol. 161, Astronomy from Wide-Field Imaging. p. 535

Seigar M. S., Barth A. J., Bullock J. S., 2008, MNRAS, 389, 1911

Simon J. D., Geha M., 2007, ApJ, 670, 313

Sohn S. T. et al., 2007, ApJ, 663, 960

van der Marel R. P., Alves D. R., Hardy E., Suntzeff N. B., 2002, AJ, 124, 2639

van der Marel R. P., Guhathakurta P., 2008, ApJ, 678, 187

Walker M. G., Mateo M., Olszewski E. W., 2008, ApJ, 688, L75

Walker M. G., Mateo M., Olszewski E. W., Bernstein R., Wang X., Woodroofe M., 2006, AJ, 131, 2114

Walker M. G., Mateo M., Olszewski E. W., Pal J. K., Sen B., Woodroofe M., 2006, ApJ, 642, L41

White S. D. M., 1981, MNRAS, 195, 1037

Wilkinson M. I., Evans N. W., 1999, MNRAS, 310, 645

Xue X. X. et al., 2008, ApJ, 684, 1143

York D. G. et al., 2000, AJ, 120, 1579

Zucker D. B. et al., 2006, ApJ, 650, L41

Zucker D. B. et al., 2006, ApJ, 643, L103

Zucker D. B. et al., 2004, ApJ, 612, L121

Zucker D. B. et al., 2007, ApJ, 659, L21