

The Need for Speed: Escape velocity and dynamical mass measurements of the Andromeda galaxy

Prajwal R. Kafle^{1*}, Sanjib Sharma², Geraint F. Lewis², Aaron S. G. Robotham¹, and Simon P. Driver¹

¹ ICRAR, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia

² Sydney Institute for Astronomy (SfA), School of Physics A28, The University of Sydney, NSW 2006, Australia

10 January 2018

ABSTRACT

Our nearest large cosmological neighbour, the Andromeda galaxy (M31), is a dynamical system, and an accurate measurement of its total mass is central to our understanding of its assembly history, the life-cycles of its satellite galaxies, and its role in shaping the Local Group environment. Here, we apply a novel approach to determine the dynamical mass of M31 using high velocity Planetary Nebulae (PNe), establishing a hierarchical Bayesian model united with a scheme to capture potential outliers and marginalize over tracers unknown distances. With this, we derive the escape velocity run of M31 as a function of galacto-centric distance, with both parametric and non-parametric approaches. We determine the escape velocity of M31 to be $470 \pm 40 \text{ km s}^{-1}$ at a galacto-centric distance of 15 kpc, and also, derive the total potential of M31, estimating the virial mass and radius of the galaxy to be $0.8 \pm 0.1 \times 10^{12} M_{\odot}$ and $240 \pm 10 \text{ kpc}$, respectively. Our M31 mass is on the low-side of the measured range, this supports the lower expected mass of the M31-Milky Way system from the timing and momentum arguments, satisfying the H I constraint on circular velocity between $10 \lesssim R/\text{kpc} < 35$, and agreeing with the stellar mass Tully-Fisher relation. To place these results in a broader context, we compare them to the key predictions of the Λ CDM cosmological paradigm, including the stellar-mass-halo-mass and the dark matter halo concentration-virial mass correlation, and finding it to be an outlier to this relation.

Key words: stars: individual: Planetary Nebulae – galaxies: individual: M31 – galaxies: kinematics and dynamics – methods: statistical

1 INTRODUCTION

Of the many billions of galaxies visible in the “Observable Universe”, only those within the local few megaparsecs are amenable to detailed dissection by our telescopes. Within this regards, our own Milky Way (MW) and the neighbouring Andromeda (M31), hold pride of place, providing the opportunity to reveal an unparalleled view of the formation and evolution of large galaxies, directly confronting our cosmological paradigm. Unfortunately, we are buried deep within the MW and thus M31 represents the only large spiral galaxy that can be observed in detail over its entirety, and hence is the focus of numerous observational programs (Ibata et al. 2004; Merrett et al. 2006; Kalirai et al. 2006; Halliday et al. 2006; McConnachie et al. 2009; Gilbert et al. 2009; Dalcanton et al. 2012).

An accurate measurement of the mass profile of the M31 galaxy is fundamental in understanding structural and dy-

namical properties as well as the formation history of the galaxy. So far several attempts have been made to acquire M31 dynamical mass, which can be classified into five broad categories:

- (i) the rotation curve method (Rubin & Ford 1970; Roberts & Whitehurst 1975; Sofue & Rubin 2001; Carignan et al. 2006; Chemin et al. 2009),
- (ii) velocity dispersion, virial theorem and tracer mass formalism (Rood 1979; Courteau & van den Bergh 1999; Côté et al. 2000; Galleti et al. 2006; Lee et al. 2008; Watkins et al. 2010; Tollerud et al. 2012; Veljanoski et al. 2014),
- (iii) dynamical model or distribution function fitting (Evans et al. 2000; Evans & Wilkinson 2000),
- (iv) tidal stream orbit modelling (Ibata et al. 2004; Fardal et al. 2013), and
- (v) the timing and angular momentum arguments (van der Marel et al. 2012; Diaz et al. 2014; González et al. 2014).

For an exhaustive review of the topic we refer the reader

* E-mail: prajwal.kafle@uwa.edu.au, pkafauthor@gmail.com

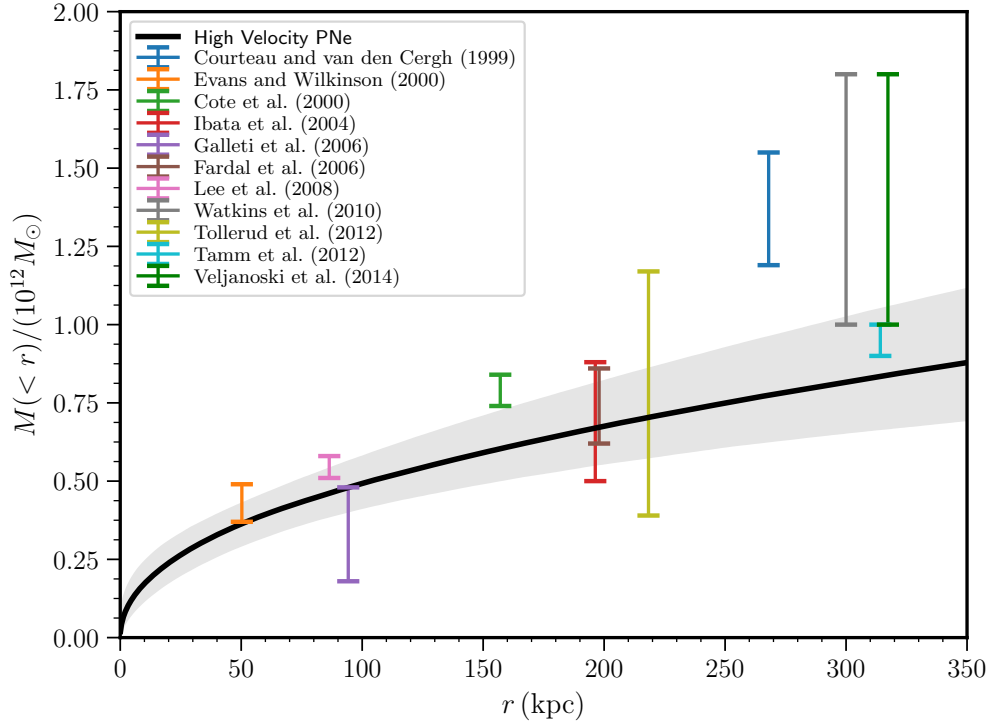


Figure 1. Summary of the recent mass measurements of the M31 galaxy obtained from the different literature sources (as labelled) that use varieties of techniques. The black solid line with grey shade (showing uncertainties) is the mass profile of the galaxy calculated in this paper by modelling the high velocity Planetary Nebulae (PNe). To convert the projected radius (R) to spherical radius (r), an approximate relation $R = 2r/\pi$ is assumed (Schneider 2006).

to Fardal et al. (2013) while in Fig. 1 we provide a crude visual summary of the range of the galaxy masses taken from recent literature sources. The use of different mass tracers, measurements at range of radii, and differences in the approaches to infer the masses make a fair comparison between the reported measurements a daunting task, nonetheless, a convenient summary of all this work is that the total mass of M31 is still uncertain, estimated to be as low as $\sim 0.7 \times 10^{12} M_{\odot}$ (Evans et al. 2000; Tamm et al. 2012) and as high as $\sim 2.5 \times 10^{12} M_{\odot}$ (Evans et al. 2000; Watkins et al. 2010) with plethora of measurements in the intermediate range (e.g. Lee et al. 2008; Fardal et al. 2013; Veljanoski et al. 2014, etc).

In the light of the huge scatter in the quoted mass of the galaxy, we seek for an alternative way to improve the measurement. In this, we attempt an independent measurement of the mass of M31 using the escape velocity inferred from the high velocity tracers, the method first proposed by Leonard & Tremaine (1990). The method has remained successful in inferring the escape speed and dynamical mass of the MW, of which the studies by Smith et al. (2007) and Piffi et al. (2014) have remained influential. Both the studies use the same Radial Velocity Experiment (RAVE, Steinmetz et al. 2006) survey, albeit different version of the data, of the Galactic disc. While Smith et al. (2007) study is only limited to the solar neighbourhood, Piffi et al. (2014) improves the method further to explore the radial dependence of the Galactic escape velocity. More recently, Williams et al. (2017) use the halo sample from the Sloan

Digital Sky Survey (SDSS) and were able to extend the method out to the MW centric distance of 50 kpc.

The key reasons for the success of this method are that it is relatively simple and it is empirically powerful as it can estimate the escape velocity from the line-of-sight velocities alone with a similar level of accuracy that can be achieved even when the full phase-space motions are used (Leonard & Tremaine 1990; Fich & Tremaine 1991). The method can be easily adopted for the case of the M31 galaxy or for that matter to any galaxy provided we have enough tracers with high velocity residing in a galaxy. In this paper, we first improve on the original method by adding additional features such as providing a Bayesian framework, a proper model to capture potential outliers, provisions for the marginalization of the unknown variables and propagation of uncertainties in the observables, a parametric and non-parametric radial fitting for the escape velocities, and eventually use the formalism to make independent predictions for the mass model of M31.

The paper is arranged as follows: In Section 2, we introduce our sample of the dynamical tracers of M31, namely Planetary Nebulae (PNe). Section 3 presents Bayesian framework of our modelling scheme. Our results and some model predictions are given in Section 4. In Section 5 we discuss our result to provide it a proper cosmological context. Finally, we draw our main conclusions in Section 6. Throughout the paper we denote the normal distribution by \mathcal{N} and the uniform distribution by \mathcal{U} . The values for the Hubble constant $H_0 = 67.8 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and the matter

density of the universe $\Omega_m = 0.308$ are assumed from the Planck Collaboration et al. (2016).

2 DATA

2.1 Planetary Nebulae (PNe)

With the development of wide-field multi-object spectrographs (Douglas et al. 2002) there has been a surge in the number of identified PNe, that happens to be one of the brightest and ubiquitous kinematic tracer of nearby galaxies. Currently, the PNe catalogue, constructed by Merrett et al. (2006) and Halliday et al. (2006), comprises the largest publicly available sample of the dynamical tracers of M31, providing sufficiently ample sample with accurate enough line-of-sight velocity information to allow the dynamical modelling of the galaxy. The data were observed using the Planetary Nebula Spectrograph on the William Herschel Telescope in La Palma and were made publicly available on-line¹. In total, the published catalogue provides 3300 emission-line objects of which only 2730 are identified as the PNe, and among the likely-PNe sample roughly 6% of the objects are associated with external galaxies and satellite galaxies within M31. We cull the contaminants and only consider the remaining 2637 PNe that belong to the M31 as our starting sample. A large fraction of these sources lie in the disk of the galaxy, and are close to the dynamical centre of the galaxy and therefore, they have short orbital time meaning they are relatively better phase-mixed in compare to the tracers in the outskirts of the galaxy where the sub-structures are known to dominate (Ibata et al. 2007; Richardson et al. 2009; Carlberg et al. 2011; Gilbert et al. 2012; Ibata et al. 2014; Kafle et al. 2017).

2.2 Frame of references and coordinate transformations

To begin with, we assume the central properties of the M31 galaxy as the followings:

Right Ascension	00 ^h 42 ^m 44.33 ^{ss}
Declination	+41°16′07.5″
Position angle ¹	37.7°
Inclination angle ²	77.5°
Distance from the Sun (d_{M31}) ³	780 kpc
Heliocentric radial velocity ($v_{M31,h}$) ⁴	-301 km s ⁻¹

¹ de Vaucouleurs (1958), ² Hodge (1992), ³ Holland (1998); Stanek & Garnavich (1998); McConnachie et al. (2005); Conn et al. (2012), and ⁴ de Vaucouleurs et al. (1991); van der Marel & Guhathakurta (2008). Additionally, we assume the distance of the Sun from the centre of the MW to be $R_0 = 8.2$ kpc (Bland-Hawthorn & Gerhard 2016). In our modelling exercise we require peculiar velocity (v_{\parallel}) of dynamical tracers of the M31 galaxy, that is the observed heliocentric line-of-sight velocity of the tracers (v_h) with contributions from the solar-reflex ($U_{\odot}, V_{\odot}, W_{\odot}$), local-standard

of rest (V_{LSR}) and a heliocentric radial motion of the M31 ($v_{M31,h}$) deducted. This can be done using the following transformation

$$v_{\parallel} = (v_h + T) - (v_{M31,h} + T) \cos \Omega, \quad (1)$$

with

$$T = U_{\odot} \cos l \cos b + (V_{\odot} + V_{LSR}) \sin l \cos b + W_{\odot} \sin b, \quad (2)$$

where l and b are the Galactic longitude and latitude, and Ω is the angular separation of the tracer from the centre of the M31. Consistent with the previous dynamical studies (e.g. Evans & Wilkinson 2000; Evans et al. 2000; Veljanoski et al. 2014, etc.) we neglect the tangential motion of M31 in our calculation. This approximation is admissible mainly because the tangential motion of M31 with respect to the MW is known to be less than about 17 km s⁻¹ (van der Marel et al. 2012) and such small transverse motion would only induce a side-wise motion of a few km s⁻¹ even for PNe as remote as $\Omega \approx 10^\circ$, which is negligible compare to our measurement uncertainties. Furthermore, to complete the transformation we assume the values for $U_{\odot} = 11.1$ km s⁻¹, $V_{\odot} = 12.24$ km s⁻¹, $W_{\odot} = 7.25$ km s⁻¹ and $V_{LSR} = 239.3$ km s⁻¹ (Schönrich et al. 2010). Similarly, we desire position vector of the tracer in M31 centric coordinate frame. Given equatorial coordinates and the heliocentric distances, we can derive respective Cartesian vectors of the M31 as well as its tracers, using say equations B1-B3 of Metz et al. (2007). Then to transform the Cartesian position vectors of the M31 tracers from the heliocentric frame of reference (denoted by $x_{*,h}$) to the M31 centric frame ($x_{*,M31}$), we can shift the position vectors to the centre of M31 and rotate them to the M31 centric coordinate system using

$$\mathbf{x}_{*,M31} = \mathcal{R}_{h \rightarrow M31}(\mathbf{x}_{*,h} - \mathbf{x}_{M31,h}), \quad (3)$$

where $\mathbf{x}_{M31,h}$ is the heliocentric Cartesian vector of the M31 galaxy. Here, $\mathcal{R}_{h \rightarrow M31}$ is the appropriate rotation matrix and is given by

$$\mathcal{R}_{h \rightarrow M31} = \begin{pmatrix} 0.770 & 0.324 & 0.549 \\ -0.632 & 0.502 & 0.591 \\ -0.084 & -0.802 & 0.592 \end{pmatrix}, \quad (4)$$

which is a function of R_0 and central coordinates of the M31 galaxy and is calculated using equation B12 of Metz et al. (2007).

In Fig. 2(a) we show our starting sample of 2637 PNe on-sky projection colour coded by the peculiar velocity (v_{\parallel}) whereas in (b) we present distribution of v_{\parallel} . Here, ~ 85 stars are seen to be high velocity candidates, depending on the kinematic restriction we apply to classify high velocity stars. The red vertical dashed lines in panel (b) put at ± 300 km s⁻¹ show our default choice of the minimum velocity, discussed later in Section 3. The heliocentric distances to these PNe are unfortunately unknown and typical uncertainty in their line-of-sight velocity $\sigma_v \sim 14$ km s⁻¹ (Merrett et al. 2006).

3 METHOD: HIERARCHICAL BAYESIAN MODELLING AND ACCESSORIES

Below we present an outline of the method that comprises three major steps of our approach: determination of the es-

¹ http://www.strw.leidenuniv.nl/pns/PNS_public_web/PN.S_data.html

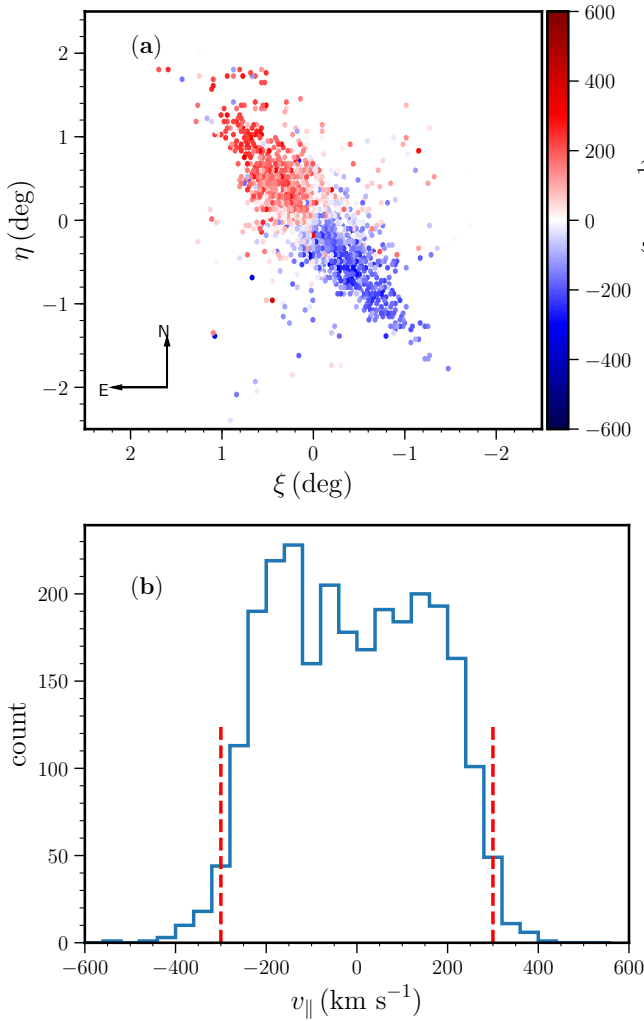


Figure 2. Peculiar velocity distributions of our starting M31 PNe sample, corrected for the M31 systematic motion. (a) Velocity distribution in the conventional on-sky coordinates ξ and η in degrees obtained from the gnomonic projection of the tracers Right Ascension and Declination. (b) Distribution of the velocities with red dashed lines showing threshold velocity $v_{\min} = \pm 300 \text{ km s}^{-1}$ used to classify the high velocity stars ($|v_{\parallel}| > |v_{\min}|$).

cape velocity, simple model predictions (derivation of cumulative mass, potential and circular velocity runs) and inference of the virial properties.

3.1 Escape velocity curve modelling

3.1.1 Setting up the inference

Similar to [Leonard & Tremaine \(1990\)](#), we start by assuming that the high velocity $\{v_{\parallel} : |v_{\parallel}| > v_{\min}\}$ wings of the distribution function of the galaxy follow a power-law distribution and can be expressed as

$$f(v_{\parallel}|r) = f_0 (v_{\text{esc}}(r) - |v_{\parallel}|)^{k+1}, \quad (5)$$

where the normalization, f_0 , is given by

$$f_0 = \int_{v_{\min}}^{v_{\text{esc}}} f(v_{\parallel}|r) dv_{\parallel} \quad (6)$$

$$= \frac{k+2}{(v_{\text{esc}}(r) - v_{\min})^{k+2}}.$$

The v_{\min} is a threshold velocity, that is the tracer with the absolute peculiar velocity ($|v_{\parallel}|$ in equation 1) larger than v_{\min} is classified as a high velocity tracer and makes it to our final sample. Generally, in the literature v_{\min} is assumed to be 300 km s^{-1} ([Leonard & Tremaine 1990](#); [Smith et al. 2007](#); [Piffl et al. 2014](#)) and we also make it our default choice, however later we investigate the effect of our choice in final results. Similar to [Piffl et al. \(2014\)](#) we express the escape speed as a function of the tracer galacto-centric distance r , this is to allow us to infer the spatial run of the speed. The specific details about the exact expressions for the $v_{\text{esc}}(r)$ are provided in Section 3.1.2.

Some tracers in our data could probably be outliers (possibly unbound stars) and to capture this we adopt an outlier, or background model, g , which is given by

$$g(v_{\parallel}) = g_0 \mathcal{T}(v_{\parallel}|0, \sigma_g, \nu_g), \quad (7)$$

and the model is normalized between $[v_{\min}, \infty]$ such that

$$g_0 = \frac{1}{2} - \frac{1}{\pi} \tan^{-1}(v_{\min}/\sigma_g).$$

Here, \mathcal{T} represents the Student t-distribution, defined by three parameters mean, variance (σ_g^2) and degree of freedom (ν_g), and is known to approximate the \mathcal{N} distribution as $\nu_g \rightarrow \infty$. The \mathcal{T} distribution is preferable over the \mathcal{N} distribution as the former has broader wings. We fix $\sigma_g = 1000 \text{ km s}^{-1}$ and $\nu_g = 1$, but investigate the effects of these assumptions in our analysis in Section 4. Additionally, we like to put a note here that instead of introducing an outlier model and thereby increasing the modelling complexity, one could simply use 3σ or 5σ velocity clippings to get rid of outliers at the earlier data curating step. However, given the potential future application of the modelling scheme, as new data arrive, it is therefore sensible to adopt the former, a comparatively superior approach.

Equipped with the above models, now we can express a joint probability distribution for the ‘complete-data’ (high velocity and outlier stars) as

$$p(v_{\parallel}^t|r) = \eta f(v_{\parallel}^t|r) + (1 - \eta) g(v_{\parallel}^t), \quad (8)$$

where η indicates the outlier fraction, v_{\parallel}^t represents the clean, error-free ‘true’ version of the observed peculiar velocity (v_{\parallel}) and σ_v is the error in v_{\parallel} . Since v_{\parallel}^t is not directly observed, we treat it as a latent variable and assume the errors σ_v are Gaussian (\mathcal{N}). Finally, we can construct the full likelihood function as

$$\mathcal{L}(\Theta) = p(v_{\parallel}^t|\Theta, r) p(r|\alpha) \mathcal{N}(v_{\parallel}|v_{\parallel}^t, \sigma_v), \quad (9)$$

Here, $p(r|\alpha)$ is the tracer density profile, which for simplicity we assume to be a single power-law of the galacto-centric distance r , that is

$$p(r|\alpha) \propto r(\mathbf{x}_{\star, \text{M31}})^{-\alpha}. \quad (10)$$

The distance, r , is a function of the position vector, $\mathbf{x}_{\star, \text{M31}}$, that is related to the observables RA and Dec, and unknown heliocentric distances, s , of the tracers via equation 3. As

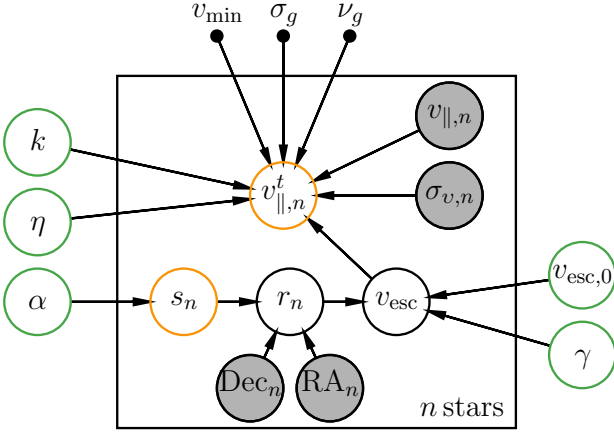


Figure 3. A plate diagram for escape velocity modelling. It shows a graphical illustration of the dependencies between the observed data (grey shaded circle) and model parameters that are kept free (green circle) or fixed (black dot) in the form of a probabilistic graphical model. Orange circles represent the latent parameters.

we do not know the distance s to the tracers we treat it as a latent parameter. Furthermore, we keep the logarithmic power-law slope α of the tracer density distribution in the range between 2 and 3.5, in line with recent literature (e.g. Irwin et al. 2005; Ibata et al. 2007; Courteau et al. 2011; Gilbert et al. 2012; Ibata et al. 2014, etc).

3.1.2 Expressions of Escape Velocity

We can determine v_{esc} , required for equation 5, either parametrically, or non-parametrically. In the first approach, we express v_{esc} in a simple parametric form given by

$$v_{\text{esc}}(r) = v_{\text{esc},0} (r/r_s)^{-\gamma}, \quad (11)$$

with the scale-length, r_s , fixed at 15 kpc.

In the other approach, we assume that the $v_{\text{esc}}(r)$ profile over the radial extent of the data is a linear interpolation of $v_{\text{esc},j}$ measured at distance r_j , where the number of nodes j and corresponding radius r_j are pre-specified. We can express this form for v_{esc} as the following

$$v_{\text{esc}}(r) = \text{Interpolate}(r_j, v_{\text{esc},j}). \quad (12)$$

Here, the number of nodes j can be viewed as a parameter that provides the radial resolution of $v_{\text{esc}}(r)$. Ideally, j can be kept infinitely large to fully capture the v_{esc} range of the galaxy. However, for practical purposes we keep $j = 6$ and calculate $\{v_{\text{esc},1}, \dots, v_{\text{esc},6}\}$ at $r/\text{kpc} = 5, 10, 15, 20, 25$ and 30; essentially $v_{\text{esc}}(r)$ that enters into equation 5 is linear interpolation of $v_{\text{esc},j}$ at the specified radii r_j .

The probabilistic graphical model indicating the interdependences of our model parameters and observables are shown in Fig. 3. Syntax-wise, in the case of the non-parametric approach we replace $\{\gamma, v_{\text{esc},0}\}$ in Θ with $\{v_{\text{esc},1}, \dots, v_{\text{esc},j}\}$ while the rest of the steps remain exactly same.

3.1.3 Prior and posterior distributions

Ultimately, in the Bayesian format the posterior distribution that we like to sample can be written as the product of priors and likelihood function as the following

$$p(\Theta, s_t, v_{\parallel}^t | v_{\parallel}) \propto p(k, v_{\text{esc},0}) \prod_n \mathcal{L}(\Theta), \quad (13)$$

where product is over all n stars. Here, Θ represents the set of model parameters $\{k, v_{\text{esc},0}, \gamma, \alpha, \eta\}$ that we aim to constrain, which forms the higher level of our hierarchical model whereas the heliocentric distance s and the error free peculiar velocity v_{\parallel}^t of the tracers make the lower level of our model. Here $p(k, v_{\text{esc},0})$ represent priors on k and $v_{\text{esc},0}$ parameters. For different Milky-Way like simulated galaxies obtained from cosmological simulations the maximum value of k is found to cover a range between 2.3-4.7 (Piffi et al. 2014). Following this we consider $p(k) \in \mathcal{U}[0, 4.7]$ as a prior on k . Similarly, for $v_{\text{esc},0}$ also we adopt the (less) informed uniform prior, $p(v_{\text{esc},0} | \text{km s}^{-1}) \in \mathcal{U}[v_{\text{min}}, 1000]$. Also, we assume flat priors for the remaining parameters, namely, $\eta \in \mathcal{U}[0, 1]$ and $\gamma \in \mathcal{U}[0, 1]$. The heliocentric distance s/kpc of each tracer is proposed between $\mathcal{U}[600, 1100]$ constrained to follow the density distribution given by equation 10 with hyper-prior $\alpha \in \mathcal{U}[2, 3.5]$ and the proposal distribution for $|v_{\parallel}^t|/\text{km s}^{-1}$ is assumed to follow $\mathcal{U}[v_{\text{min}}, |v_{\parallel}| + 5\sigma_v]$.

3.2 Mass modelling

With the escape velocity curve next we can derive the mass profile and the virial properties of the galaxy. For this we interpret the escape velocity, v_{esc} , as the minimum speed required to reach some limiting radius, r_{max} , and relate it to the galactic potential using the relation

$$v_{\text{esc}} = \sqrt{2|\Phi(r) - \Phi(r_{\text{max}})|}, \quad (14)$$

where Φ represents the total potential of the galaxy at a given radius r . Similar to Smith et al. (2007) and Piffi et al. (2014), we also assume r_{max} to be thrice that of the virial radius, a sufficiently large distance. Since the gravitational potential is a weak function of radius at such large distances, we expect results are insensitive to small changes in the limiting radius, which is demonstrated to be a valid assumption in Piffi et al. (2014). The above equation can be solved to express $\Phi(r)$ as a function of v_{esc} , which can then be fitted to a mass model of the M31 and the model can be finally used to infer key dynamical properties of the galaxy. But first, we must construct a realistic model of the galaxy.

We assume that the two most dominant stellar components of the M31 that is bulge and disk follow Hernquist (1990) and Miyamoto & Nagai (1975) type potentials given by

$$\Phi_b = -\frac{G M_b}{r+q} \quad (15)$$

and

$$\Phi_d = -\frac{G M_d}{\sqrt{R^2 + (a + \sqrt{z^2 + b^2})^2}}, \quad (16)$$

respectively, where M_b and q represent the bulge mass and scale-length, whereas M_d , a and b are for the disk mass, scale-length and scale-height respectively.

In addition to the bulge and disk models, we assume the potential of the dark matter halo to have a Navarro-Frenk-White (NFW, Navarro et al. 1996) profile given by

$$\Phi_h = \frac{-G M_{\text{vir}} \ln(1 + r c / r_{\text{vir}})}{g(c) r}, \quad (17)$$

with $g(c) = \ln(1 + c) - c/(1 + c)$, and

$$M_{\text{vir}} = \frac{4\pi}{3} r_{\text{vir}}^3 \Delta \rho_c \quad (18)$$

where M_{vir} , r_{vir} and c denote the virial mass, virial radius and concentration parameter, whereas Δ is the virial over-density parameter and $\rho_c = 3H_0^2/(8\pi G)$ is the critical density of the Universe. Generally, two contrasting choices for Δ are preferred in the literature. First, for the spherical collapse definition $\Delta = \Omega_m \delta_{\text{th}}$, where $\delta_{\text{th}} = 340$ is an over-density of the dark matter compared to the average matter density (Bryan & Norman 1998), that is $\Delta \simeq 100$. We denote virial mass and radius for this cases as M_{vir} and r_{vir} respectively. Second, some literature assume $\Delta = 200$, and in this case we label the corresponding mass and radius as M_{200} and r_{200} respectively. The different choices for Δ is just a matter of convention and there is no correct or incorrect choice. The outer radius and mass enclosed are inversely proportional to the Δ parameter as for a system with monotonically decreasing density distribution 200 times over-density would occur at smaller radius than 100 times over-density compare to the average matter density of the universe. We report final measurements for both the cases.

Finally, we combine the contribution of each component of the galaxy and express the total potential of the galaxy as

$$\Phi_t(r) = \Phi_b(r) + \langle \Phi_d(R, z) \rangle_\theta + \Phi_h(r). \quad (19)$$

Clearly, the bulge and halo models are spherically symmetric but the disk model is a function of both cylindrical radius ($R = r \sin \theta$) and height above the plane of the disk ($z = r \cos \theta$), and is azimuthally averaged.

There are total of seven parameters (M_b , q , M_d , a , b , M_{vir} , c) that completes our bulge-disk-halo model, of which the parameters of the baryonic components are assumed as the following

Bulge scale-length (q)	0.7 kpc,	
Bulge mass (M_b)	$3.4 \times 10^{10} M_\odot$,	
Disk scale-length (a)	6.5 kpc,	(20)
Disk scale-height (b)	0.26 kpc, and	
Disk mass (M_d)	$6.9 \times 10^{10} M_\odot$.	

The values for scale-lengths and scale-height are adopted from Bekki et al. (2001) and Font et al. (2006) whereas the masses (in the units of $10^{10} M_\odot$) are taken to be a straight average of the most recent estimates obtained from Geehan et al. (2006, $M_b = 3.2$, $M_d = 7.2$, a case with the disk mass-to-light ratio of 3.3), Seigar et al. (2008, $M_b = 3.5$, $M_d = 5.8$, inferred from the *Spitzer* 3.6-micron imaging data and $B - R$ color profile), Chemin et al. (2009, $M_b = 2.32$, $M_d = 7.1$, bulge mass is derived from stellar velocity dispersions and the disk mass is determined from stellar population models), Corbelli et al. (2010, $M_b = 3.8$, $M_d = 8.8$, estimated by fitting bulge, disc, gas and dark matter halo

combined model to H I kinematics data) and Tamm et al. (2012, $M_b = 4.4$, $M_d = 5.7$, derived from fitting the rotation curve). The remaining two parameters, c and M_{vir} , that defines the dark matter halo are kept free and constrained via the maximum likelihood analysis. The likelihood function or the total probability of obtaining the data (total potential (Φ_t) given the model parameters (c and M_{vir}) that we aim to maximise is given by

$$p(c, M_{\text{vir}} | \Phi(r)) = \prod_j \mathcal{N}(\Phi_t(r_j) | \Phi(r_j), \sigma_{\Phi(r_j)}). \quad (21)$$

Having established the Bayesian framework for our problem, we like to ultimately sample the posterior distribution $p(\Theta, s_t, v_{\parallel}^t | v_{\parallel})$ (equation 13) and $p(c, M_{\text{vir}} | \Phi(r))$ (equation 21). Below we briefly discuss the sampling technique.

3.3 Inference: Metropolis within Gibbs (MWG)

The posterior distribution $p(\Theta, s_t, v_{\parallel}^t | v_{\parallel})$ is a function of the set of model parameters $\Theta = \{v_{\text{esc},0}, \gamma, \eta, \alpha, k\}$ or $\Theta = \{\{v_{\text{esc},0}, \dots, v_{\text{esc},j}\}, \gamma, \eta, \alpha, k\}$ depending on whether we adopt the parametric (equation 11) or the non-parametric (equation 12) form for v_{esc} . To sample the posterior we make use of the Metropolis within Gibbs sampler from BMCMC software² (Sharma 2017a,b). Firstly, using the general Metropolis-Hastings algorithm, a set of values for the model parameters, Θ , is proposed from a proposal distribution. Note the posterior distribution function also explicitly depends on the known unknowns or the latent variables such as the heliocentric distance and true peculiar velocity, v_{\parallel}^t , of the tracers. We must marginalize the distribution function over these latent variables before we can infer anything about the model parameters. Unfortunately, the expression for the likelihood function given in Equation 9 is non-analytical meaning marginalization must be done numerically. Therefore, we set-up a multi-level (hierarchical) inference where we could directly consider the latent variables as hyper-parameters and sample them for each proposal of the model parameters. For this we employ the BMCMC that uses Gibbs sampling method for the marginal inference. In short, given a set of observables, $(\alpha, \delta, \mathcal{R}, v_{\parallel}, \sigma_v)$, for each tracer along with the proposed values for (v_{\parallel}^t, s) and assuming a galactocentric number density distribution (equation 10), BMCMC yields the posterior distributions in the form of walks/chains of the model parameters Θ and ancillary data such as the log-likelihood value at each walk. We use these chains for the final inference, and also where needed we propagate them to derive physical quantities of the interest.

Similarly, in the separate case of fitting the mass model via equation 21, we again make use of the BMCMC. In this case there is no latent variable and hence, BMCMC only uses the general Metropolis Hastings algorithm to sample the distribution for the model parameters $\{M_{\text{vir}}, c\}$. For all the undertaken experiments, we run MWG for sufficient autocorrelation time to ensure that the distributions of parameters are stabilized around certain values. We report the median of the probability distribution of a model parameter as its best estimate and 16th and 84th percentiles of the distribution as measure of its uncertainty.

² <https://github.com/sanjibs/bmcmc>

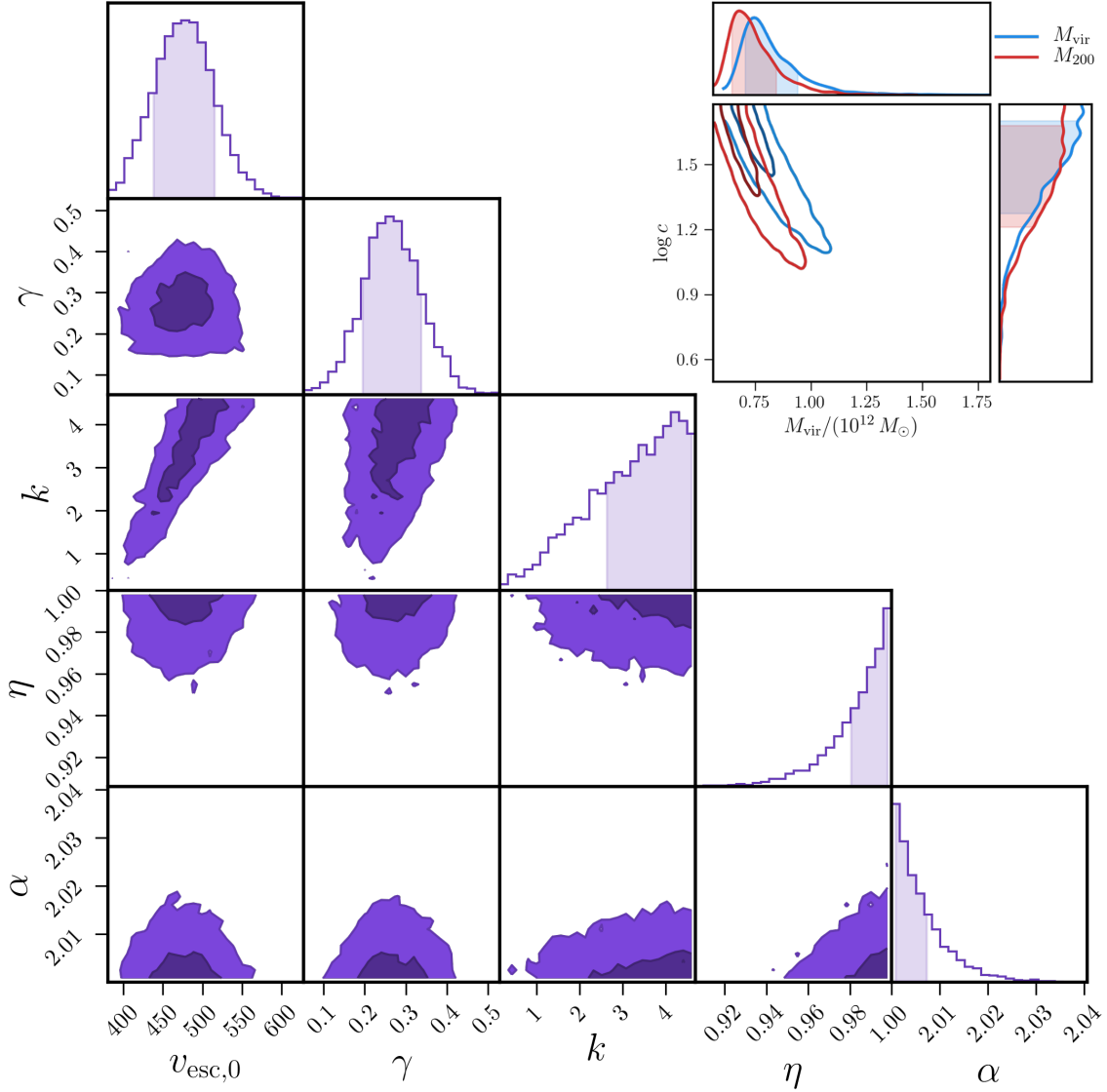


Figure 4. Joint posterior distributions of the model parameters obtained from the escape velocity and mass modellings, showing the default case, where $v_{\min} = 300 \text{ km s}^{-1}$, $r_s = 15 \text{ kpc}$, $d_{\text{M31}} = 780 \text{ kpc}$, $v_{\text{M31,h}} = -301 \text{ km s}^{-1}$, $\sigma_g = 1000 \text{ km s}^{-1}$ and $\nu_g = 1$ are assumed. Panel in the inset shows the joint distributions of concentration and virial mass obtained from the mass modelling for the corresponding case, where blue and red distributions show results for two different definitions of the halo masses as labelled.

4 RESULTS

In this work, we present new measurements of the escape velocity curve $v_{\text{esc}}(r)$ and the virial properties of the M31 galaxy. Below, we first present these findings, and later discuss the possible systematics that may have crept into our final results due to some assumptions that we had to make. Table 1 summarizes the inferred values of our prime model parameters for multitude of cases.

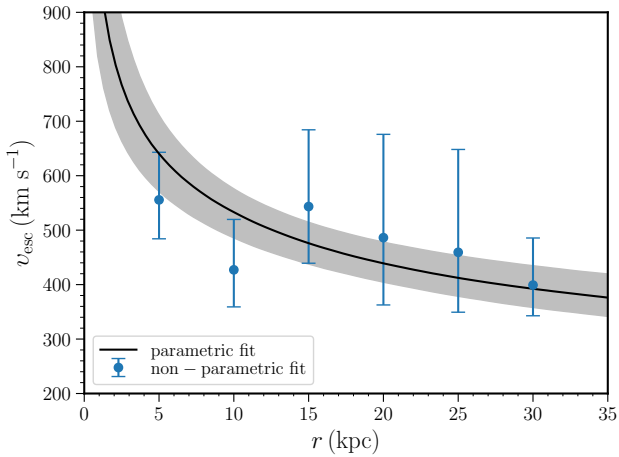
4.1 Escape velocity: inference

Fig. 4 shows the joint probability distributions of the model parameters obtained from a parametric fit of v_{esc} (equation 11) to the high velocity PNe sample assuming ‘default’ values of the parameters and key physical properties of the M31, that is, the threshold velocity $v_{\min} = 300 \text{ km s}^{-1}$,

$r_s = 15 \text{ kpc}$, $d_{\text{M31}} = 780 \text{ kpc}$, $v_{\text{M31,h}} = -301 \text{ km s}^{-1}$, $\sigma_g = 1000 \text{ km s}^{-1}$, $\nu_g = 1$ and uniform priors on k and $v_{\text{esc},0}$; the first row of Table 1 presents the measurements for this case. Using our scheme, we derive an escape velocity curve of the M31 and its logarithmic power-law slope $\gamma = 0.26 \pm 0.07$, and measure the escape velocity $v_{\text{esc},0} = 470 \pm 40 \text{ km s}^{-1}$ at the galacto-centric radius of $r_s = 15 \text{ kpc}$. The parametric run of the v_{esc} is shown with the black solid line in Fig. 5, where the grey band around the line shows the associated uncertainties inferred from the posterior distribution. Complementing the parametric approach, we also present the v_{esc} measurements at the given radii obtained from a non-parametric fit (Equation 12), which is shown in the figure with blue dots and vertical error bars. We see that the non-parametric run has larger uncertainties in comparison to the parametric case, which is expected as the non-parametric case has larger degree of freedom due to the increased num-

Table 1. Estimated model parameters from different analysis. When unspecified, the default values for the model parameters are assumed, i.e., $v_{\min} = 300 \text{ km s}^{-1}$, $r_s = 15 \text{ kpc}$, $d_{M31} = 780 \text{ kpc}$, $v_{M31,h} = -301 \text{ km s}^{-1}$, $\sigma_g = 1000 \text{ km s}^{-1}$ and $\nu_g = 1$.

Cases	Number of PNe	$v_{\text{esc},0}$ (km s^{-1})	γ	η	k	$M_{\text{vir}}(M_{200})$ ($10^{12} M_{\odot}$)	$r_{\text{vir}}(r_{200})$ (kpc)
Default	86	470^{+40}_{-40}	$0.26^{+0.07}_{-0.07}$	$0.99^{+0.02}_{-0.01}$	3^{+1}_{-1}	$0.8^{+0.1}_{-0.1}(0.7^{+0.1}_{-0.1})$	$240^{+10}_{-10}(188^{+7}_{-11})$
$d_{M31} = 790 \text{ kpc}$	86	480^{+40}_{-40}	$0.27^{+0.07}_{-0.07}$	$0.99^{+0.02}_{-0.01}$	3^{+1}_{-1}	$0.8^{+0.1}_{-0.1}(0.7^{+0.1}_{-0.1})$	$240^{+10}_{-10}(188^{+7}_{-11})$
$d_{M31} = 770 \text{ kpc}$	86	470^{+40}_{-40}	$0.27^{+0.07}_{-0.07}$	$0.99^{+0.02}_{-0.01}$	3^{+1}_{-1}	$0.8^{+0.1}_{-0.1}(0.7^{+0.07}_{-0.12})$	$240^{+10}_{-10}(187^{+7}_{-10})$
$v_{M31,h} = -296 \text{ km s}^{-1}$	87	490^{+40}_{-40}	$0.25^{+0.07}_{-0.07}$	$0.99^{+0.02}_{-0.01}$	3^{+1}_{-1}	$0.9^{+0.1}_{-0.2}(0.8^{+0.1}_{-0.2})$	$250^{+10}_{-20}(190^{+10}_{-10})$
$v_{M31,h} = -306 \text{ km s}^{-1}$	90	470^{+40}_{-40}	$0.27^{+0.07}_{-0.07}$	$0.99^{+0.02}_{-0.01}$	3^{+1}_{-1}	$0.73^{+0.07}_{-0.13}(0.66^{+0.06}_{-0.11})$	$240^{+10}_{-10}(184^{+6}_{-9})$
$v_{\min} = 290 \text{ km s}^{-1}$	111	470^{+40}_{-30}	$0.27^{+0.06}_{-0.06}$	$0.99^{+0.01}_{-0.01}$	3^{+1}_{-1}	$0.76^{+0.07}_{-0.12}(0.69^{+0.06}_{-0.11})$	$239^{+7}_{-12}(186^{+6}_{-9})$
$v_{\min} = 310 \text{ km s}^{-1}$	65	480^{+40}_{-40}	$0.24^{+0.08}_{-0.08}$	$0.98^{+0.03}_{-0.01}$	3^{+1}_{-1}	$0.9^{+0.1}_{-0.2}(0.8^{+0.1}_{-0.2})$	$250^{+10}_{-20}(190^{+10}_{-10})$
$r_s = 14 \text{ kpc}$	86	480^{+40}_{-40}	$0.26^{+0.07}_{-0.08}$	$0.99^{+0.02}_{-0.01}$	3^{+1}_{-1}	$0.8^{+0.1}_{-0.1}(0.7^{+0.1}_{-0.1})$	$240^{+10}_{-10}(189^{+7}_{-11})$
$r_s = 16 \text{ kpc}$	86	470^{+40}_{-40}	$0.27^{+0.07}_{-0.07}$	$0.99^{+0.02}_{-0.01}$	3^{+1}_{-1}	$0.8^{+0.1}_{-0.1}(0.7^{+0.07}_{-0.12})$	$240^{+10}_{-10}(187^{+7}_{-10})$
$\sigma_g = 800 \text{ km s}^{-1}$	86	470^{+40}_{-40}	$0.25^{+0.08}_{-0.07}$	$0.98^{+0.03}_{-0.01}$	3^{+1}_{-1}	$0.8^{+0.1}_{-0.2}(0.7^{+0.1}_{-0.1})$	$240^{+10}_{-20}(189^{+8}_{-11})$
$V_{\odot} + V_{\text{LSR}} = 240 \text{ km s}^{-1}$	86	480^{+40}_{-40}	$0.27^{+0.07}_{-0.07}$	$0.99^{+0.02}_{-0.01}$	3^{+1}_{-1}	$0.8^{+0.1}_{-0.1}(0.71^{+0.07}_{-0.13})$	$240^{+10}_{-10}(188^{+7}_{-11})$
$V_{\odot} + V_{\text{LSR}} = 260 \text{ km s}^{-1}$	86	470^{+40}_{-40}	$0.27^{+0.07}_{-0.08}$	$0.99^{+0.02}_{-0.01}$	3^{+1}_{-1}	$0.8^{+0.1}_{-0.1}(0.7^{+0.07}_{-0.12})$	$240^{+10}_{-10}(187^{+6}_{-10})$

**Figure 5.** Estimated escape velocity curve using parametric (the black solid line with grey shades) and non-parametric (the blue dot with error bars) approaches with default assumptions (results corresponding to the first row in Table 1).

number of free parameters compared to the other case³. Comparing the results from the two different fits, in overall, we see that the results agree within the range of uncertainties and thus, validates our choice for the parametric expression of $v_{\text{esc}}(r)$ (i.e., equation 11). However, there are some small

³ known as the bias-variance trade-off (Hastie et al. 2009)

systematic differences. For $R < 15 \text{ kpc}$, the non-parametric estimate of v_{esc} is less than the parametric one, while for $R > 15 \text{ kpc}$, opposite is true. The observed HI rotation curve also shows a similar behaviour with respect to our derived parametric rotation curve (Fig. 6). This suggests that our parametric model is not flexible enough to capture the full information available in the data. Since the systematic deviations are small, we only consider the parametric form so as to take advantage of the reduced variance and modelling complexity.

Unfortunately, we are unable to constrain some model parameters that were kept free. Importantly, we find that although a large value is preferred for k , it remains unconstrained. Unconstrained k parameters due to scant number of high velocity stars have remained the main limitation of the undertaken approach, the fate that has been also shared by the MW high velocity works (Leonard & Tremaine 1990; Smith et al. 2007, refer to their figures 1 and 7 respectively). Also, we find that the logarithmic number density slope favour shallow ($\alpha = 2$) over steep model ($\alpha = 3$) and the outlier fraction $\eta \rightarrow 1$ suggesting that effectively only a tiny fraction of our tracers are actually drawn from the background model.

4.2 Virial properties: inference

Here, we first substitute the measured $v_{\text{esc}}(r)$ run into Equation 14 to derive the total potential of the galaxy. Then we fit the three component galaxy model to the derived potential and ultimately estimate the virial properties (M_{vir} and c)

of the dark matter halo, keeping the disk and bulge models fixed with Equation 20. This exercise is repeated separately for both the definitions of halo virial over-density. In Fig. 4 inset we show the joint distributions of the $M_{\text{vir}}(M_{200})$ and c resulted from this exercise in blue(red) contours for a case with ‘default’ values of the model parameters. The marginalized distributions of the halo parameters are shown in the corresponding colours at the top and the right sides of the inset-figure. Unfortunately, the marginalized distributions for the c hint that within a large range of parameter space that we explore i.e. [1,60], we are unable to constrain the c parameter, but we are able to measure the virial masses regardless; our final estimates for the $M_{\text{vir}}(M_{200})$ (also shown in the first row in Table 1) are 0.8 ± 0.1 (0.7 ± 0.1) $\times 10^{12} M_{\odot}$. Substituting these values for the masses in equation 18, we determine $r_{\text{vir}}(r_{200}) = 240 \pm 10$ (188^{+7}_{-11}) kpc. As discussed earlier in Section 3.2, we note that the M_{vir} or r_{vir} values are always greater than for M_{200} or r_{200} .

4.3 Relaxing the assumptions

Finally, it is instructive to investigate repercussions of relaxing some of the key assumptions on our final results, which we discuss below.

Effects of change in M31 central properties

Arguably, the heliocentric distance and the line-of-sight velocity of the M31 galaxy are known fairly accurately, with the scatter of roughly 1 – 2% amongst the literature (e.g., de Vaucouleurs et al. 1991; Stanek & Garnavich 1998; McConnachie et al. 2005; van der Marel & Guhathakurta 2008; Conn et al. 2012, etc.). Nevertheless, we systematically shift the adopted value of $d_{\text{M31}} = 780$ kpc and $v_{\text{M31}} = -301 \text{ km s}^{-1}$ (Section 2.2) by ± 10 kpc and $\pm 5 \text{ km s}^{-1}$ respectively and re-run our analysis to understand their implications on our final results. Rows 2-5 in Table 1 provide the estimates of the main model parameters for these cases. Importantly, we see that while the introduced systematic shifts in the d_{M31} have as such negligible effect on our final results, the biases introduced in the v_{M31} alters the $|M_{\text{vir}}|$ estimates by $\sim 8\%$. Furthermore, one of our assumption is that the Sun’s total tangential velocity relative to the Galactic centre, $V_{\odot} + V_{\text{LSR}}$ sums to 251.5 km s^{-1} noting that it is approximately equal to the literature averaged value of 248 km s^{-1} (for $R_{\odot} = 8.2$ kpc) from Bland-Hawthorn & Gerhard (2016). Since it is apparent from equations 1 and 2 that $v_{\parallel} \propto (V_{\odot} + V_{\text{LSR}})$, we test our choice. For this we adopt two different values 240 km s^{-1} and 260 km s^{-1} for the Sun’s total tangential velocity and re-run the analysis. The measurements are reported in the bottom two rows of the Table 1, and it is evident that the systematic of $\pm 10 \text{ km s}^{-1}$ in $V_{\odot} + V_{\text{LSR}}$ does not alter our final estimates, and likewise when we adopt slightly different values of $U_{\odot} = 10 \text{ km s}^{-1}$ and $W_{\odot} = 7 \text{ km s}^{-1}$ from Bland-Hawthorn & Gerhard (2016). The null effect is expected because the T term defined in equation 2 that contains these constants, for a small angle Ω , largely cancels out in equation 1.

Effects of change in the threshold velocity (v_{min})

As already described, v_{min} is the lower bound on the magnitude of the v_{\parallel} that is used to discriminate the high velocity stars. In our final results, we adopt $v_{\text{min}} = 300 \text{ km s}^{-1}$, which is consistent with the value adopted in the literature Leonard & Tremaine (1990); Smith et al. (2007); Piffl et al. (2014). However, here we desire to investigate the effect of change in v_{min} . For this first we lower the limit and set $v_{\text{min}} = 290 \text{ km s}^{-1}$, this increases sample size to 111 as more stars are now classified as high velocity. With the new set of data we find that the $M_{\text{vir}}/10^{12} M_{\odot}$ reduces marginally from 0.8 to 0.76. Second, when we increase the limit and set $v_{\text{min}} = 310 \text{ km s}^{-1}$ we find that the sample size reduces to 65 and the $M_{\text{vir}}/10^{12} M_{\odot}$ slightly increases from 0.8 to 0.9. Moreover, we find that relaxing the assumption on v_{min} by $\pm 10 \text{ km s}^{-1}$ leads to insignificant changes in the estimates of other parameters, e.g., $v_{\text{esc},0}$, γ etc as well. Rows 6-7 in Table 1 provide the estimates of the model parameters for these cases.

Indiscernible effects due to change in some parameters

There are also some model dependencies that we tested but which led to negligible changes in our final results, we briefly discuss them here. Throughout the paper, to model the outliers (Equation 7), we assume that the $\sigma_g = 1000 \text{ km s}^{-1}$ and $\nu_g = 1$. We test that adopting 800 km s^{-1} or 1200 km s^{-1} for σ_g and/or varying $\nu_g \in [0.5, 2]$ effectively do not result any changes in the final M_{vir} measurements. Similarly, given the large distance between M31-MW ($d_{\text{M31}} \gg R_0$) the assumed position of the Sun on the MW disk ($R_{\odot} = 8.2$ kpc versus 8.5 kpc) has negligible effect in our analysis. Moreover, we check that our final results are robust to the change of at least 5% in the upper limits of the prior ranges assumed for v_{esc} , γ , v_{\parallel}^t and α . Similarly, reducing the lower limit of prior range for α parameter from 2.0 to 1.0 do not modify our final M_{vir} measurement. A slight change of ± 1 kpc in the adopted value of escape velocity scale-length $r_s = 15$ kpc has no discernible effect in our mass measurement. Results discussed in this Section are again provided in Table 1.

5 DISCUSSION

5.1 Independent validation

It is important to affirm our mass model with existing independent measurements, which are based on completely different physics and hence, should have different systematics. In summary;

With the rotation curve

Rotation curves have been derived for the M31 galaxy and the history of the measurement dates back to the seminal works by Rubin & Ford (1970); Rubin et al. (1980) that use the optical data from H α emission lines. Similarly, the curves have also been derived from the radio data using HI or CO observations (e.g. Roberts & Whitehurst 1975; Loinard et al. 1995; Carignan et al. 2006; Chemin et al. 2009; Corbelli et al. 2010, etc.). It is crucial that we check

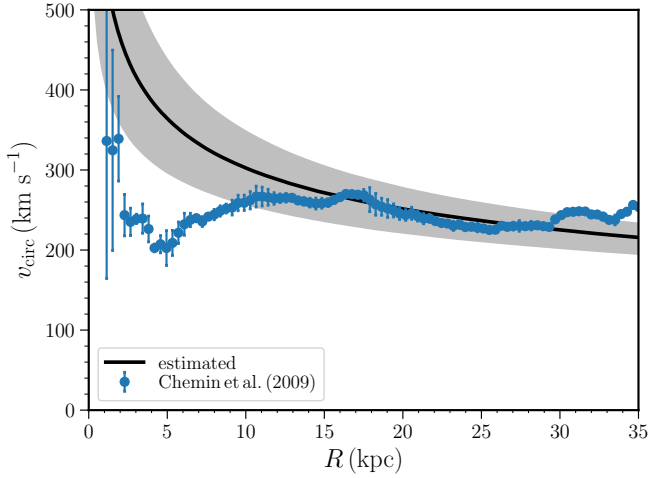


Figure 6. Circular velocity curve of the M31. Blue dots with error bars are measured values by Chemin et al. (2009) using H I emission line observations whereas grey shade with black solid line is our best estimate.

our model prediction for the rotation curve with these independent measurements. In the spherical approximation and assuming that the gravitational and the centripetal acceleration are equal, we can derive the circular speed v_{circ} from the enclosed mass $M(< r)$ within radius r using

$$v_{\text{circ}}(r) = \frac{GM(< r)}{r}, \quad (22)$$

where the enclosed mass is derived from the escape velocity curve using the relation

$$M(< r) = -\frac{r^2 v_{\text{esc}}}{G} \frac{dv_{\text{esc}}}{dr},$$

that with Equation 11 it reduces to

$$M(< r) = \frac{\gamma r}{G} v_{\text{esc}}^2. \quad (23)$$

The black line and the shaded grey regions in Fig. 6 show the derived v_{circ} and its associated uncertainty. Additionally, the blue dots with error bars show an independent measurement of the M31 rotation curve, which is constructed using H I emission line observations of five fields toward the galaxy; for details refer to Chemin et al. (2009). Some notable features in the H I rotation curve are that the velocities assume large values (350 km s^{-1}) in the innermost regions, which dips at $R = 4 \text{ kpc}$, gradually increases up to $\sim 270 \text{ km s}^{-1}$ at $\simeq 15 \text{ kpc}$ and remains flat (230 km s^{-1}) out to 35 kpc . The v_{circ} we derive from the modelling of the high speed stars is largely in an agreement with the H I rotation curve at $R \gtrsim 10 \text{ kpc}$, considering the large uncertainty around the v_{circ} estimate that we have. Clearly, the two curves are not in agreement in the inner region, which is expected as the assumption of spherical symmetry of the disk used in deriving our v_{circ} profile.

We note that at $r_s = 15 \text{ kpc}$ the derived v_{esc} is a factor of ~ 1.5 times larger than $\sqrt{2}v_{\text{circ}}$. If we look this result in conjunction with equation 14 of Smith et al. (2007) or reference therein, interestingly, we can infer that there is a significant contribution to the v_{esc} by mass exterior to r_s , demonstrating the presence of a dark halo in M31.

Finally, in Figure 1 we compare our derived cumulative mass profile of the M31 with different literature values and find that they are largely consistent. Exception to these are the M31 mass measurements using the tracer mass formalism (Watkins et al. 2010; Veljanoski et al. 2014), which could be systematically off as the results are degenerate to the unknown velocity anisotropy parameter of the stellar halo.

With the prediction of the baryonic Tully-Fisher relation (bTFR)

A fundamental empirical relationship between the total baryonic or stellar mass in galaxies and their maximum rotation velocity exists, which is known as the bTFR and it states that the baryonic mass of the spiral galaxy is proportional to the circular speed to the power of roughly ~ 4 (McGaugh et al. 2000; Gurovich et al. 2004; Trachternach et al. 2009; McGaugh 2012; Dutton 2012; McGaugh & Schombert 2015). For our assumed value for the total stellar mass of the M31 of $13 - 20 \times 10^{10} M_{\odot}$ and adopting the proportionality constant of $41 - 53 M_{\odot} / (\text{km s}^{-1})^4$ from McGaugh (2012), it is predicted that the range of maximum $v_{\text{circ}} / \text{km s}^{-1} \in [222, 264]$. This estimate can be directly compared with the v_{circ} at the flat part of the measured rotation curve. Consistent with the bTFR, we also see from Fig. 6 that our v_{circ} is equal to 250 km s^{-1} say at $R \approx 15 \text{ kpc}$.

With the timing and momentum mass arguments

As a final consistency check, we turn to the timing and mass arguments that utilize the relative motion of M31 and the MW to independently estimate the intergalactic mass or effectively the total mass of the Local Group. From this, one can subtract off the MW mass, assuming it is known accurately enough, and directly predict the M31 mass.

The fact that M31 is moving radially toward the MW indicates that the mass of these galaxies must be sufficient to overcome cosmic expansion; this is known as the timing mass argument and was first proposed by Kahn & Woltjer (1959). Similarly, the argument that M31 and the MW should have equal and opposite momenta in the barycentric frame of reference of the Local Group is known as the momentum argument (Einasto & Lynden-Bell 1982). Early works based on the timing mass argument, such as Li & White (2008); van der Marel et al. (2012), have advocated for consistently high mass, i.e., $\sim 5 \times 10^{12} M_{\odot}$, for the Local Group, suggesting the M31 mass should be roughly 4-6 times that of the MW. However, more recently improved timing (González et al. 2014) and momentum (Diaz et al. 2014) mass arguments have downward revised the M31-MW combined mass to $\sim 2.5 \pm (0.4 - 1.5) \times 10^{12} M_{\odot}$. Our current estimate for the mass of M31, $0.8 \pm 0.1 \times 10^{12} M_{\odot}$, coupled with the MW mass of $0.8 - 1.2 \times 10^{12} M_{\odot}$ (Kafle et al. 2012, 2014; Bland-Hawthorn & Gerhard 2016), implies a total mass of $\sim 2 \times 10^{12} M_{\odot}$, which considering the uncertainties in both measurements lean supports to the downward revised mass of the MW-M31 system.

5.2 Cosmological context

It is essential to understand the properties of Local Group galaxies within a cosmological context. Moreover, observations of the local universe have revealed that the existence of the ‘MW-Magellanic Clouds-M31’-like trio is cosmologically rare (Robotham et al. 2012), which is even more scarce if we also seek for the planar configuration of the member satellites (Ibata et al. 2015). As our current theory of structure formation the Λ CDM is getting ever more sophisticated mainly, due to increase in particle resolution, refinement of “gastro-physics” and hydrodynamics, it is useful to compare its core predictions relevant to our observed results as it would allow us to establish the uniqueness or generality of the Local Group. In the following, we compare our results against two profound theoretical predictions, that is, the concentration–virial mass scaling and the stellar-mass–halo-mass relations.

Comparison with concentration–virial mass relation

A generic prediction of simulated virialized halos in the Λ CDM cosmology is that its mass and concentration are inversely proportional to each other, and assume a typical relation of $c \propto M_{\text{vir}}^{-0.12}$ for $\log(M_{\text{vir}}/M_{\odot}) \in [11, 13]$ (e.g. Bullock et al. 2001; Macciò et al. 2007; Duffy et al. 2008; King & Mead 2011; Klypin et al. 2016; Ludlow et al. 2016, etc.). In Fig. 7(a), we show the similar $c - M_{\text{vir}}$ relations from two representative fits from simulations Macciò et al. (2007, in black dashed line) and Duffy et al. (2008, in black dash-dotted line) to which we overlap the green contours estimated of the M31 from Section 4. Similarly, the heat map shows the measured $c - M_{\text{vir}}$ relation for the MW from Kafle et al. (2014). Interestingly, we see that both the galaxies show similar joint distributions although they are derived using different approaches and tracers, and barely overlaps with the Λ CDM predictions at 2σ level.

Comparison with Stellar-Mass–Halo-Mass relation

Another prediction of the Λ CDM is that the stellar mass of the central galaxy embedded within a dark matter halo is correlated with its host halo mass, and is known as the stellar-mass–halo-mass relation (SHM; e.g. Moster et al. 2010; Behroozi et al. 2010, 2013, etc). The SHM can be well defined by a double power law where on the either side of a peak-value the trends drop, and the peak-value corresponds to the knee of the stellar mass function (e.g. Baldry et al. 2012; Moffett et al. 2016, etc.). In Fig. 7b, the black dashed line shows a representative SHM adapted from Behroozi et al. (2010) near our range of interest, that is $M_{\text{vir}} \sim 10^{11-12.5}$. The grey shaded regions represent a typical scatter of ~ 0.3 dex that the SHM relations in different literature sources are considered to have. The blue dot with error bar marks our measurement for the halo mass of the M31 galaxy, given the disk and bulge combined stellar mass of $1.03 \times 10^{11} M_{\odot}$ taken from equation 20. Clearly, the estimate for the M31 galaxy is higher than the generic prediction of the Λ CDM. Similarly, for a comparison we also show the green dot with error bar that represents the estimated values of dark matter halo ($1.3_{-0.3}^{+0.4} \times 10^{12} M_{\odot}$) and combined baryonic ($6.5 \times 10^{10} M_{\odot}$) masses of the MW adopted from

Piffi et al. (2014), who uses identical modelling approach to this work. Additionally, in orange dot we show the yet another measurement of the MW mass, but from Kafle et al. (2014) that uses the Jeans formalism.

Here, we like to leave a cautionary note that the above discussed theoretical predictions are for pure dark matter simulations, and are prone to serious systematics as they do not include baryonic processes such as cooling, star formation, and feedback. For example, the collapse of gas due to cooling leads to adiabatic contraction of the dark matter halo, which increases its concentration. Feedback, on the other hand, can have the reverse effect. Similarly, the relationship between dark matter halos and galaxy stellar masses from the halo abundance matching technique rely on the accuracies of observed stellar mass function, the theoretical halo mass function and techniques of abundance matching.

Before we summarize our result, we point out some of the limitations of the escape velocity modelling technique. As highlighted in Piffi et al. (2014) and Smith et al. (2007) the conceptual underpinning of the technique that the density of stars in the velocity space tend to zero as velocity $\rightarrow v_{\text{esc}}$ (equation 5) is fairly weak. While in some analytic equilibrium models of stellar systems such as Jaffe (1983) and Hernquist (1990) there is a non-zero probability density of finding a star all the way to the v_{esc} and zero beyond that boundary; in other models, for example, the King-Michie model (King 1966), the density is zero even at speed $< v_{\text{esc}}$. Moreover, it is unknown whether the phase space distribution of the galactic tracers extend to v_{esc} and also whether it depends on the stellar types (for example, PNe versus regular stars). Any truncation in the velocity distribution will result an underestimation of the true v_{esc} , meaning in such case our estimates of v_{esc} and thus M_{vir} would only provide the lower limits of the true values. As such this will naturally solve the disagreement we note earlier between the theoretical stellar-mass–halo-mass relation and the observed value of the M31. Being said that Smith et al. (2007) notes that for a set of galaxies obtained from cosmological simulations (Abadi et al. 2006) the level of truncation in the velocity distributions of the stellar component are found to be less than 10 per cent, and it seems not to have any systematic effect in their recovered v_{esc} .

6 SUMMARY

In this work, we establish a Hierarchical Bayesian framework in which we improve the method by Leonard & Tremaine (1990), and for the first time report an independent measurement of the radial dependence of the escape velocity curve, $v_{\text{esc}}(r)$, and a new estimate of the dynamical mass of the M31 galaxy.

We employ the Planetary Nebulae (PNe) data by Merrett et al. (2006) and Halliday et al. (2006) as it is currently the only publicly available tracer of the M31 disk with ample sample size and accurate enough line-of-sight velocity measurements. After removing the marked extragalactic contaminants and satellite members, and stars with $|v_{\parallel}| \leq v_{\text{min}}$, the final sample we analyse in our main results comprises 86 PNe. From this high velocity sample of the PNe along-with our modelling scheme, we first estimate the

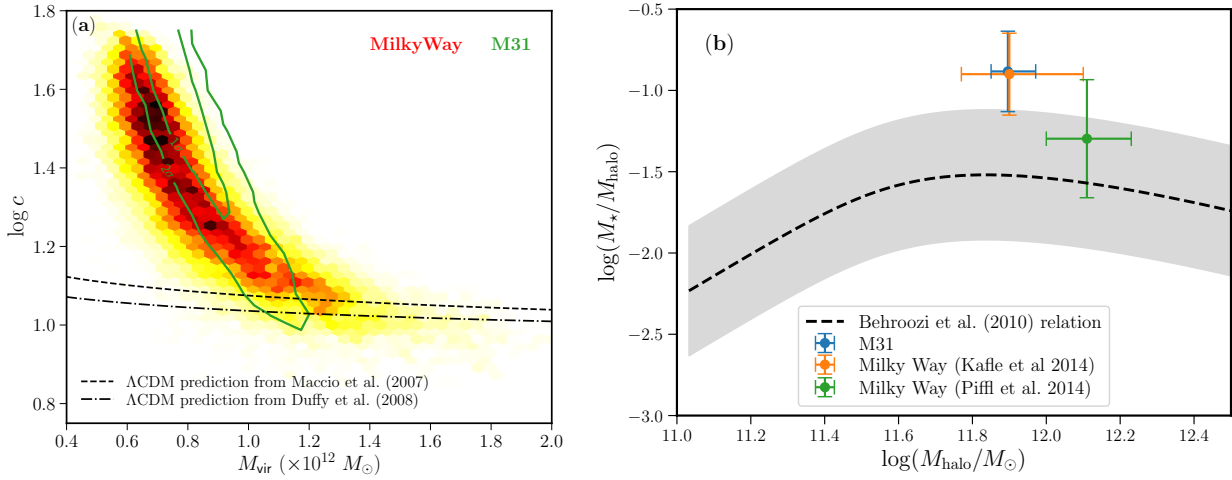


Figure 7. Comparison with the prediction of the Λ CDM. Left panel (a) showing concentration (c)–virial mass (M_{vir}) joint probability distributions where green contour is our best estimate for the M31 whereas the heat map shows similar relation for the Milky Way taken from Kafle et al. (2014). The black dashed and dash-dotted lines demonstrate a typical $c - M_{\text{vir}}$ relations predicted by Λ CDM dark matter simulations of Macciò et al. (2007); Duffy et al. (2008) respectively. Right panel (b): Stellar mass–halo mass relation expected from the abundance matching in the dark matter simulation, shown with the black dashed lines and grey shaded regions, where blue marker show the observed positions of the M31, and orange (from Kafle et al. 2014) and green (from Piffl et al. 2014) markers show the Milky Way galaxies respectively.

$v_{\text{esc}}(r)$ and then derive the total galactic potential to which finally we fit a three component (bulge, disk and dark matter halo) mass model and infer the virial properties of the galaxy. In the end, to provide a proper cosmological context, we discuss how the newly estimated mass of the M31 compares to the mass of our own the MW and also, to some generic predictions of the Λ CDM such as the concentration–virial mass relation, the stellar mass–halo mass relation and the stellar Tully–Fisher relation.

Following are the main conclusions of the paper:

(i) We present both the parametric and non-parametric $v_{\text{esc}}(r)$ profiles of the M31. Assuming the minimum threshold velocity $v_{\text{min}} = 300 \text{ km s}^{-1}$, the criteria set to classify high velocity stars, we measure $v_{\text{esc}} = 470 \pm 40 \text{ km s}^{-1}$ at the galacto-centric radius of $r = 15 \text{ kpc}$. Additionally, we are also able to constrain the logarithmic power-law slope of the profile $\gamma = 0.26 \pm 0.07$.

(ii) Using the derived $v_{\text{esc}}(r)$ profile and assuming spherical symmetry we are able to further derive the cumulative mass profile, the circular velocity profile as well as the total potential of the galaxy. To the derived potential we then fit a three component model of the galaxy — Hernquist bulge, Miyamoto–Nagai disk and NFW dark matter halo model, of which we adopt the bulge and disk structural models from the literature (Bekki et al. 2001; Tamm et al. 2012) and keep them fixed. It is important to keep these components fixed because we only restrict our fitting to the derived potentials at $r \gtrsim 10 \text{ kpc}$ as we observe that only in this regime the derived circular velocity profile is in a good agreement with the rotation curve constructed from the H I measurements. However, we keep the two defining parameters that is the concentration and the virial mass of the dark matter halo free. We find that assuming literature averaged bulge mass of $3.4 \times 10^{10} M_{\odot}$ and disk mass of $6.9 \times 10^{10} M_{\odot}$, the derived potential of the galaxy is best fit by a halo of the

virial mass $M_{\text{vir}}(M_{200}) = 0.8 \pm 0.1 (0.7 \pm 0.1) \times 10^{12} M_{\odot}$ that corresponds to the virial radius of $240 \pm 10 (188_{-11}^{+7}) \text{ kpc}$.

(iii) We find that the circular velocity curve (v_{circ}) between $10 \lesssim R/\text{kpc} < 35$ estimated by us is in good agreement with an independent prediction from the H I observation. Similarly, the value of $v_{\text{circ}} = 250 \text{ km s}^{-1}$ we obtain at the flat part of the curve at $R \simeq 15 \text{ kpc}$ is consistent with the prediction from the baryonic Tully–Fisher relation. We note that the measured $c - M_{\text{vir}}$ joint distributions and the observed locus of the stellar and dark matter halo masses of the M31 barely agree to the theoretical predictions at only 2σ levels.

ACKNOWLEDGEMENTS

We like to thank the referee Prof. Mathias Steinmetz for a constructive and insightful report.

Software credit: ASTROPY (Astropy Collaboration et al. 2013), BMC MC (Sharma 2017a), CHAINCONSUMER (Hinton 2016), DAFT <https://github.com/dfm/daft>, IPYTHON (Pérez & Granger 2007), MATPLOTLIB (Hunter 2007), NUMPY (van der Walt et al. 2011), PANDAS (McKinney 2012) and SCIPY (Jones et al. 2001).

REFERENCES

- Abadi M. G., Navarro J. F., Steinmetz M., 2006, *MNRAS*, **365**, 747
 Astropy Collaboration et al., 2013, *A&A*, **558**, A33
 Baldry I. K., et al., 2012, *MNRAS*, **421**, 621
 Behroozi P. S., Conroy C., Wechsler R. H., 2010, *ApJ*, **717**, 379
 Behroozi P. S., Wechsler R. H., Conroy C., 2013, *ApJ*, **770**, 57
 Bekki K., Couch W. J., Drinkwater M. J., Gregg M. D., 2001, *ApJ*, **557**, L39
 Bland-Hawthorn J., Gerhard O., 2016, *ARA&A*, **54**, 529

- Bryan G. L., Norman M. L., 1998, *ApJ*, **495**, 80
- Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, *MNRAS*, **321**, 559
- Carignan C., Chemin L., Huchtmeier W. K., Lockman F. J., 2006, *ApJ*, **641**, L191
- Carlberg R. G., et al., 2011, *ApJ*, **731**, 124
- Chemin L., Carignan C., Foster T., 2009, *ApJ*, **705**, 1395
- Conn A. R., et al., 2012, *ApJ*, **758**, 11
- Corbelli E., Lorenzoni S., Walterbos R., Braun R., Thilker D., 2010, *A&A*, **511**, A89
- Côté P., Mateo M., Sargent W. L. W., Olszewski E. W., 2000, *ApJ*, **537**, L91
- Courteau S., van den Bergh S., 1999, *AJ*, **118**, 337
- Courteau S., Widrow L. M., McDonald M., Guhathakurta P., Gilbert K. M., Zhu Y., Beaton R. L., Majewski S. R., 2011, *ApJ*, **739**, 20
- Dalcanton J. J., et al., 2012, *ApJS*, **200**, 18
- Diaz J. D., Koposov S. E., Irwin M., Belokurov V., Evans N. W., 2014, *MNRAS*, **443**, 1688
- Douglas N. G., et al., 2002, *PASP*, **114**, 1234
- Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., 2008, *MNRAS*, **390**, L64
- Dutton A. A., 2012, *MNRAS*, **424**, 3123
- Einasto J., Lynden-Bell D., 1982, *MNRAS*, **199**, 67
- 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
- Fardal M. A., et al., 2013, *MNRAS*, **434**, 2779
- Fich M., Tremaine S., 1991, *ARA&A*, **29**, 409
- Font A. S., Johnston K. V., Guhathakurta P., Majewski S. R., Rich R. M., 2006, *AJ*, **131**, 1436
- Galleti S., Federici L., Bellazzini M., Buzzoni A., Fusi Pecci F., 2006, *A&A*, **456**, 985
- Geehan J. J., Fardal M. A., Babul A., Guhathakurta P., 2006, *MNRAS*, **366**, 996
- Gilbert K. M., et al., 2009, *ApJ*, **705**, 1275
- Gilbert K. M., et al., 2012, *ApJ*, **760**, 76
- González R. E., Kravtsov A. V., Gnedin N. Y., 2014, *ApJ*, **793**, 91
- Gurovich S., McGaugh S. S., Freeman K. C., Jerjen H., Staveley-Smith L., De Blok W. J. G., 2004, *Publ. Astron. Soc. Australia*, **21**, 412
- Halliday C., et al., 2006, *MNRAS*, **369**, 97
- Hastie T., Tibshirani R., Friedman J., 2009, *The Elements of Statistical Learning—Data Mining, Inference, and Prediction*, second edn. Springer Series in Statistics, Springer New York, doi:10.1007/978-0-387-84858-7
- Hernquist L., 1990, *ApJ*, **356**, 359
- Hinton S., 2016, *JOSS*, **1**
- Hodge P. W., ed. 1992, *The Andromeda galaxy Astrophysics and Space Science Library Vol. 176*
- Holland S., 1998, *AJ*, **115**, 1916
- Hunter J. D., 2007, *Computing In Science & Engineering*, **9**, 90
- Ibata R., Chapman S., Ferguson A. M. N., Irwin M., Lewis G., McConnachie A., 2004, *MNRAS*, **351**, 117
- 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., et al., 2014, *ApJ*, **780**, 128
- Ibata R. A., Famaey B., Lewis G. F., Ibata N. G., Martin N., 2015, *ApJ*, **805**, 67
- Irwin M. J., Ferguson A. M. N., Ibata R. A., Lewis G. F., Tanvir N. R., 2005, *ApJ*, **628**, L105
- Jaffe W., 1983, *MNRAS*, **202**, 995
- Jones E., Oliphant T., Peterson P., et al., 2001, *SciPy: Open source scientific tools for Python*, <http://www.scipy.org/>
- Kaffe P. R., Sharma S., Lewis G. F., Bland-Hawthorn J., 2012, *ApJ*, **761**, 98
- Kaffe P. R., Sharma S., Lewis G. F., Bland-Hawthorn J., 2014, *ApJ*, **794**, 59
- Kaffe P. R., Sharma S., Robotham A. S. G., Lewis G. F., Driver S. P., 2017, *MNRAS*, **464**, 4858
- Kahn F. D., Woltjer L., 1959, *ApJ*, **130**, 705
- Kalirai J. S., et al., 2006, *ApJ*, **648**, 389
- King I. R., 1966, *AJ*, **71**, 64
- King L. J., Mead J. M. G., 2011, *MNRAS*, **416**, 2539
- Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, *MNRAS*, **457**, 4340
- Lee M. G., Hwang H. S., Kim S. C., Park H. S., Geisler D., Sarajedini A., Harris W. E., 2008, *ApJ*, **674**, 886
- Leonard P. J. T., Tremaine S., 1990, *ApJ*, **353**, 486
- Li Y.-S., White S. D. M., 2008, *MNRAS*, **384**, 1459
- Loinard L., Allen R. J., Lequeux J., 1995, *A&A*, **301**, 68
- Ludlow A. D., Bose S., Angulo R. E., Wang L., Hellwing W. A., Navarro J. F., Cole S., Frenk C. S., 2016, *MNRAS*, **460**, 1214
- Macciò A. V., Dutton A. A., van den Bosch F. C., Moore B., Potter D., Stadel J., 2007, *MNRAS*, **378**, 55
- McConnachie A. W., Irwin M. J., Ferguson A. M. N., Ibata R. A., Lewis G. F., Tanvir N., 2005, *MNRAS*, **356**, 979
- McConnachie A. W., et al., 2009, *Nature*, **461**, 66
- McGaugh S. S., 2012, *AJ*, **143**, 40
- McGaugh S. S., Schombert J. M., 2015, *ApJ*, **802**, 18
- McGaugh S. S., Schombert J. M., Bothun G. D., de Blok W. J. G., 2000, *ApJ*, **533**, L99
- McKinney W., 2012, *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*. O'Reilly Media, Inc.
- Merrett H. R., et al., 2006, *MNRAS*, **369**, 120
- Metz M., Kroupa P., Jerjen H., 2007, *MNRAS*, **374**, 1125
- Miyamoto M., Nagai R., 1975, *PASJ*, **27**, 533
- Moffett A. J., et al., 2016, *MNRAS*, **457**, 1308
- Moster B. P., Somerville R. S., Maubetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, *ApJ*, **710**, 903
- Navarro J. F., Frenk C. S., White S. D. M., 1996, *ApJ*, **462**, 563
- Pérez F., Granger B. E., 2007, *Comput. Sci. Eng.*, **9**, 21
- Piffl T., et al., 2014, *A&A*, **562**, A91
- Planck Collaboration et al., 2016, *A&A*, **594**, A13
- Richardson J. C., et al., 2009, *MNRAS*, **396**, 1842
- Roberts M. S., Whitehurst R. N., 1975, *ApJ*, **201**, 327
- Robotham A. S. G., et al., 2012, *MNRAS*, **424**, 1448
- Rood H. J., 1979, *ApJ*, **232**, 699
- Rubin V. C., Ford Jr. W. K., 1970, *ApJ*, **159**, 379
- Rubin V. C., Ford W. K. J., Thonnard N., 1980, *ApJ*, **238**, 471
- Schneider P., 2006, *Extragalactic Astronomy and Cosmology*. Springer
- Schönrich R., Binney J., Dehnen W., 2010, *MNRAS*, **403**, 1829
- Seigar M. S., Barth A. J., Bullock J. S., 2008, *MNRAS*, **389**, 1911
- Sharma S., 2017a, *bmcnc: MCMC package for Bayesian data analysis*, Astrophysics Source Code Library (ascl:1709.009)
- Sharma S., 2017b, *ARA&A*, **55**, 213
- Smith M. C., et al., 2007, *MNRAS*, **379**, 755
- Sofue Y., Rubin V., 2001, *ARA&A*, **39**, 137
- Stanek K. Z., Garnavich P. M., 1998, *ApJ*, **503**, L131
- Steinmetz M., et al., 2006, *AJ*, **132**, 1645
- Tamm A., Tempel E., Tenjes P., Tihhonova O., Tuvikene T., 2012, *A&A*, **546**, A4
- Tollerud E. J., et al., 2012, *ApJ*, **752**, 45
- Trachternach C., de Blok W. J. G., McGaugh S. S., van der Hulst J. M., Dettmar R.-J., 2009, *A&A*, **505**, 577
- Veljanoski J., et al., 2014, *MNRAS*, **442**, 2929
- Watkins L. L., Evans N. W., An J. H., 2010, *MNRAS*, **406**, 264
- Williams A. A., Belokurov V., Casey A. R., Evans N. W., 2017, *MNRAS*, **468**, 2359
- de Vaucouleurs G., 1958, *ApJ*, **128**, 465
- de Vaucouleurs G., de Vaucouleurs A., Corwin Jr. H. G., Buta R. J., Paturel G., Fouqué P., 1991, *Third Reference Catalogue of Bright Galaxies. Volume I: Explanations and references*.

- Volume II: Data for galaxies between 0^h and 12^h . Volume III:
Data for galaxies between 12^h and 24^h . Springer
- van der Marel R. P., Guhathakurta P., 2008, [ApJ](#), **678**, 187
- van der Marel R. P., Fardal M., Besla G., Beaton R. L., Sohn
S. T., Anderson J., Brown T., Guhathakurta P., 2012, [ApJ](#),
753, 8
- van der Walt S., Colbert S. C., Varoquaux G., 2011,
[Computing in Science & Engineering](#), 13