Phase Space Models of the Dwarf Spheroidals

N. C. Amorisco and N. W. Evans

Abstract

This paper introduces new phase-space models of dwarf spheroidal galaxies (dSphs). The stellar component has an isotropic, lowered isothermal (or King) distribution function. A physical basis for the isotropization of stellar velocities is given by the theory of tidal stirring, whilst the isothermality of the distribution function guarantees the observed flatness of the velocity dispersion profile in the inner parts. For any analytic dark matter potential – whether of cusped or of cored form – the stellar density and velocity dispersion are analytic.

The origin of the observational correlation between half-light radius $R_h$ and line of sight central velocity dispersion $\sigma_{p,0}$ is investigated. We prove that a power-law correlation $R_h \propto \sigma_{p,0}^D$ can exist if, and only if, the dark halo potential is a power-law of the radius. Although a power-law is a good approximation in the central parts ($D = 2$ for a Navarro-Frenk-White halo, $D = 1$ for cored halos), the theoretical correlation curve between $R_h$ and $\sigma_{p,0}$ dramatically steepens at larger half-light radii. Using our phase space models, we show that different dark halo profiles – whether cored or cusped – lead to very similar mass estimates within one particular radius, $\approx 1.7R_h$. The formula for the enclosed mass is $M(<1.7R_h) \approx 5.8\sigma_{p,0}^2 R_h/G$ and extends out to much larger radii than previous investigations. This is a tight result for models with a flattish projected velocity dispersion profile (out to several half-light radii).

We show that deviations between mass measures due to different density profiles are substantially smaller than the uncertainties propagated by the observational errors on the half-light radius and central velocity dispersion. We produce a mass measure for each of the dSphs and find that the two most massive of the Milky Way dSphs are the most luminous, namely Sgr ($M(<1.7R_h) \approx 2.8 \times 10^8 M_\odot$) and Fornax ($\sim 1.3 \times 10^8 M_\odot$). The least massive of the Milky Way satellites are Willman 1 ($\sim 4 \times 10^7 M_\odot$) and Segue 1 ($\sim 6 \times 10^5 M_\odot$).

Key words: galaxies: dwarf – galaxies: kinematics and dynamics – Local Group

1 INTRODUCTION

The dwarf spheroidal (dSph) galaxies surrounding the Milky Way have been the subject of intensive observational programs in recent years. Thanks to the labours of a number of groups (see e.g., Mateo et al. [1993], Kleyna et al. [2002, 2003], Wilkinson et al. [2004], Walker et al. [2009]), radial velocity surveys with multi-object spectrographs have now provided datasets of thousands of stars for the bright dSphs, like Fornax and Draco. Early indications that the velocity dispersion profiles might be flat (Kleyna et al. [2003]) have now been confirmed for all the bright dSphs out to the radius at which the mean surface brightness falls to the background (Walker et al. [2007]).

The huge interest in the dSphs is of course provoked by the enormous mass-to-light ratios inferred from their stellar kinematics. Ever since the pioneering work of Aaronson (1983), it has been apparent that the dSphs are the most dark matter dominated systems in the Universe. From a theoretical perspective, the substantial dark matter content makes the dSphs relatively simple. They are typically composed of intermediate-age to old stellar populations apparently embedded in massive dark halos. In the bright dSphs, star formation therefore ceased many dynamical times ago, and so the stellar content should be well-mixed in the potential of the dark halo. This therefore offers one of the best opportunities to learn about the structure of dark halos in the local universe.

Since 2005, there has been a succession of discoveries of very faint Local Group dwarfs in data from the Sloan Digital Sky Survey. This includes at least ten new dwarf spheroidals (Willman et al. [2005], Zucker et al. [2006],...).

* E-mail: amorisco@ast.cam.ac.uk, nwe@ast.cam.ac.uk
Belokurov et al. [2006a, 2006b, 2007], one dwarf irregular [Irwin et al. 2007], as well as three satellites that have properties intermediate between dwarf galaxies and globular clusters [Willman et al. 2005b; Belokurov et al. 2007; 2009]. All these objects have surface brightnesses and luminosities lower than any previously known galaxies, and consequently have come to be known collectively as the ultrafaints. It is often assumed, although without much evidence, that the ultrafaints are the low luminosity counterparts of the classical dwarfs.

It is fair to say that the modelling of dSphs has not kept pace with the march of the observational data. A number of authors have inferred properties of the dSphs based on the Jeans equations. Typically, this takes the form of assuming a parametric light profile for the stellar component and inferring the velocity dispersion from the Jeans equations given an assumed law for the dark matter halo, often of cusped or Navarro-Frenk-White form (Strigari et al. 2008; Pönarrubia et al. 2008; Walker et al. 2009c). This is a legitimate procedure, although it does have substantial drawbacks. First, there is no guarantee that a physical distribution function exists for the model. For example, it is not possible to embed an isotropic cored stellar profile in a Navarro-Frenk-White halo, even though the Jeans equations yield a solution – as the configuration falls foul of the Central Velocity Dispersion Theorem (An & Evans 2009); see also Ciotti & Pellegrini (1992) for an earlier related result). Second, the luminous and dark matter profiles are both posited a priori and it is therefore unlikely that this approach will lead to any insight beyond the starting assumptions. There is no physical connection between the luminous and dark matter, other than the fact that the velocity dispersions can support the model against gravitational collapse.

Here, we shall take a different approach based on phase space modelling. This is harder than Jeans modelling and has been pursued less often, but is also more powerful. A significant previous assault on the problem was made by Wilkinson et al. (2002), although with a restricted class of models in which the baryonic and dark matter have the same characteristic scalelength. The flatness of the velocity dispersion profiles of dSphs, as is evident in the impressive data of Walker et al. (2007), suggests that an obvious starting point for the dSph stars is an isotropic, isothermal distribution function. Better still is to use their spatially limited analogues, the lowered or quasi-isothermals, which have tidal radii imposed by the Milky Way potential. These distribution functions are familiar in the modelling of self-gravitating star clusters (Michie 1963; King 1966). Let us again emphasise that this approach is tailored to give models that have flattish projected velocity dispersion profiles out to several half-light radii. The stellar component of the dSphs is of course not self-gravitating, so we will develop the theory of lowered or quasi-isothermal distribution functions embedded in dark matter haloes. Notice that this gives a physically motivated starting point which ensures the flatness of the velocity dispersion profiles in the inner parts of the dSphs. There is also a physical connection between the light profile and the dark matter, as the relaxation of stars towards a quasi-isothermal distribution function takes place in the dark matter potential. The same distribution in energy space gives rise to different light profiles in different dark matter potentials.

Since detailed photometric and kinematic profiles are available only for a few of the dSphs, we do not model each galaxy individually. Rather, we take advantage of the fact that half-light radii and central velocity dispersions are instead available for at least 28 dSphs, orbiting either the Milky Way or the Andromeda galaxies. Observational data show a clear correlation between these two physical quantities, highlighting a connection between the length scale of these systems, their kinematic properties and hence their dark matter content (Walker et al. 2009c). Although the existence of the correlation is clear-cut, its precise form and origin is open to some dispute.

We begin in §2 by re-visiting the data on half-light radii and velocity dispersion. Walker et al. (2009c) interpreted the data as a power-law correlation, and analysed the consequences in terms of a universal halo profile for all the dSphs. We show that, in such a context, a strict power-law correlation necessarily implies that the universal gravitational potential is a power-law of the radius. In other words, for realistic halo models, the correlation always shows deviations from the power-law form at large half-light radii. These residuals contain important physical information on the form of the dark halo potential.

The next two sections sketch the theoretical framework to analyse the correlation. In §3, we derive the behaviour of the correlation in the central parts of cusped and cored universal halos, by assuming an isothermal distribution function for the luminous stellar components. Although these asymptotic results hold good if the luminous material is embedded deep within the dark halo, they eventually became unreliable. Accordingly, §4 develops the theory of quasi-isothermal distribution functions in dark halo potentials, which enable us to extend the correlation into the regime where the power-law breaks down. This gives us families of distribution functions that build dSphs, for which the velocity dispersion profile is flat out to several multiples of the half-light radii and for which the correlation between velocity dispersion and half-light radius is essentially the same, modulo scaling transformations.

We return to the hypothesis of universality, and fit a single dark halo model to the data on half-light radius and velocity dispersion in §5. It is perhaps better though to perform an object by object analysis. This gives predictions on the mass of the dark halo for each dSph in §6. We show that our modelling gives robust mass estimates out to about 1.7 times the half-light radius. This strong result is a direct consequence of using distribution function modelling for the luminous component instead of a parametric density profile. This eliminates, by construction, all non-physical solutions that a Jeans analysis cannot exclude.

2 THE CORRELATION

2.1 Walker’s Ansatz

Walker et al. (2009c) looked for a correlation between the half-light radius and central velocity dispersion of the form:

\[
\frac{R_h}{\text{pc}} \approx C \left( \frac{\sigma_{0,0}}{\text{kms}} \right)^D
\]  

(1)
or, equivalently,
\[
\log \left( \frac{R_h}{\text{pc}} \right) \approx D \log \left( \frac{\sigma_{D,0}}{\text{kms}^{-1}} \right) + \log C , \quad (2)
\]
where the coefficient \(C\) and exponent \(D\) are chosen to give the best fit to the data on the dSphs. Here, \(R_h\) is the projected half-light radius, that is, the radius of the projected cylinder which contains a half of the total luminosity of the system, whilst \(\sigma_{D,0} \equiv \sigma_D(0)\) is the projected or line of sight velocity dispersion at the centre.

We will re-visit the fitting shortly, but it is worth exploring at outset the consequences of a strict power-law correlation like eqn (1). It has been claimed that the dSphs could actually be characterized by some kind of physical universality, for example concerning a common mass scale (see e.g., Mateo et al. 1993; Gilmore et al. 2007; Strigari et al. 2008). The hypothesis of the uniformity of the properties of the dark matter halos embedding the local dSphs constitutes an obvious first step in trying to grasp the physical meaning of any correlations. Let \(\Phi(r)\) be the spherically symmetric gravitational potential characterizing all the dSphs, and let \(\rho_*(r)\) be the stellar density distribution of any dSph (they can be different). Given the overwhelming preponderance of dark matter in dSphs, the potential well \(\Phi\) is supposed to be generated by the dark matter halo only. The luminous components are test or tracer particles only.

Each luminous component may be described by its distribution function \(f_*\), which we assume to have the same functional form for all the dSphs. There is no evidence for anisotropic velocity dispersions, so we make the simplest possible assumption of isotropy. The distribution function \(f_*\) must be a function of the energy only, namely \(f_* = f_*(E/\sigma^2)\), where \(\sigma\) is a constant. As a consequence, the luminous density distribution is a function of radius through the gravitational potential only:
\[
\rho_*(r) = \frac{\Phi(r)}{\sigma^2} . \quad (3)
\]
Here, we do not need to fix the properties of the distribution function \(f_*\), and thus of the density \(\rho_*\), in any greater detail. We just require that the velocity dispersion parameter \(\sigma\) be closely related to the actual physical (central) velocity dispersion \(\langle v^2 \rangle_*\), as naturally happens on dimensional grounds.

The equation which defines the half-light radius \(R_h\) is
\[
M_\ast(R_h) = 4\pi \int_0^{R_h} R \, dR \int_0^\infty \rho_*(\Phi(r)/\sigma^2) \, \frac{r \, dr}{\sqrt{r^2 - R^2}} = 2\pi \int_0^\infty R \, dR \int_0^\infty \rho_*(\Phi(r)/\sigma^2) \, \frac{r \, dr}{\sqrt{r^2 - R^2}} . \quad (4)
\]
A power-law relation \(M_\ast(R_h) \propto \sigma^D\) is equivalent to the invariance of eqn (4) with respect to the transformation \(\sigma \to \beta \sigma, \quad r \to \beta^D r\) for all positive real constants \(\beta\). This is an invariant transformation if and only if
\[
\rho_*(\frac{\Phi(r)}{\sigma^2}) = \rho_*(\frac{\Phi(\beta^D r)}{(\beta \sigma)^2}) . \quad (5)
\]

For reasonable, non-constant luminous densities \(\rho_\ast\), the condition (5) is satisfied if and only if the gravitational potential itself has a power-law dependence on the radius, that is if
\[
\frac{d \log \Phi}{dr} \frac{d \log r}{d \log r} = 0 . \quad (7)
\]
It follows that the gravitational potential itself is a power-law:
\[
\Phi(r) = \Phi_0 \left( \frac{r}{r_0} \right)^\delta , \quad (8)
\]
where \(\delta = 2/D\). Under the present universality hypothesis, the exponent \(D\) is directly determined by the shape of the dark matter potential only, that is, by the constant \(\delta\). In other words, the correlation can be re-written as
\[
\frac{R_h}{r_0} = \theta(\delta, f_*) \left( \frac{\sigma^2}{\Phi_0} \right)^{\frac{1}{\delta}} . \quad (9)
\]

The symbol \(\theta\) indicates a simple coefficient, independent of the velocity dispersion parameter, but directly related to the value of the exponent \(\delta\) and to the properties of the distribution function \(f_*\). Different distribution functions \(f_*\) can only affect the precise value of the coefficient \(\theta\), but cannot modify the exponent \(D\). We have thus proved the general theorem that for a spherically symmetric, universal dark halo, a power-law correlation between half-light radius and central velocity dispersion can exist only if the potential is a power-law of the radius.

Of course, the gravitational potential of any dark halo is surely more complicated than a simple power-law. Nonetheless, many popular dark halo models, such as the Navarro-Frenk-White or cored isothermal profiles, are well-approximated by power-laws in the inner parts. For dSphs in which the luminous material is embedded deep within a dark halo, a power-law approximation may work well. Nonetheless, we expect that the power-law correlation (1) will break down for objects with larger half-light radii.

### 2.2 The Data

We adopt the dataset reported by Walker et al. (2010). In Table 1, we reproduce the entire dataset, correcting a typographical error in the half-light radius of Tucana.

The major awkwardness in fitting the data is incorporating the observational errors, which are of comparable significance on both axes. Walker et al. (2009a) used an iterative fitting method that assigns weights according to measurement uncertainties in both dimensions from Rutledge et al. (1997). They found log \(C \approx -1.5\) and \(D \approx 5\) (see Walker et al. 2010). One consequence of the statistical technique is that the fit is strongly constrained to lie close to datapoints with small error bars, which in this case means, for example, the Sagittarius (Sgr) dSph, amongst others. This may be undesirable here, as the Sgr is undergoing disruption in the Milky Way halo and its halo properties may be somewhat different to the bulk of the sample. It is not even clear that the progenitor of the Sgr was a dSph (Niederste-Ostholt et al. 2010; Peinarrubia et al. 2010).

Here, we prefer to use the alternative statistical technique of Structural Analysis (Kendall & Stuart 1973).
The method is based on the maximization of the likelihood:

$$\ln L = -\frac{1}{2} \sum_{i=1}^{N} \left( \frac{y_i - Y_i}{\sigma_{y_i}} \right)^2 + \left( \frac{z_i - F(Y_i; D, C)}{\sigma_{z_i}} \right)^2,$$  \hspace{1cm} (10)

in which \((y_i, z_i)\) are the \(N\) observed values of the data pairs with standard deviations \((\sigma_{y_i}, \sigma_{z_i})\), while \(Z_i = F(Y_i; D, C)\) is the underlying power-law model, namely eqns \((1)\) or \((2)\). The essence of the method is that the \(N\) coordinates \(Y_i\) are considered unknown. This is in contrast to the classical \(y^2\) evaluation, in which the independent variable has no observational error and so \(Y_i = y_i\). For each pair \((D, C)\) of the free parameters, the \(N\) coordinates \(Y_i\) are chosen to be those that maximize the likelihood.

Let us first consider the case in which the fit is performed in the natural plane, and thus the model function \(F\) is given by eqn \((1)\). In this case, the errors \((\sigma_{y_i}, \sigma_{z_i})\) are straightforwardly identified with the observational errors in Table \(1\) and thus the likelihood \(L\) is a continuous and differentiable function of the coordinates \(Y_i\). As a consequence, the following set of \(N\) equations determines the values of \(Y_i\)

$$\frac{\partial \ln L}{\partial Y_i} = \frac{y_i - Y_i}{\sigma_{y_i}^2} + \frac{\partial F}{\partial Y_i} \left[ \frac{z_i - F(Y_i; D, C)}{\sigma_{z_i}^2} \right] = 0 \hspace{1cm} (11)$$

For general values of \(D\), eqns \((11)\) cannot be solved analytically, but, given the continuity and differentiability of the likelihood, they can be easily solved numerically.

If the fitting is performed in the logarithmic plane, then matters are slightly more complex. Here, the errors \((\sigma_{y_i}, \sigma_{z_i})\) are asymmetric with respect to the central observation point \((y_i, z_i)\). The \((Y_i, Z_i)\) plane is then effectively divided in four quadrants, each characterized by a different \((\sigma_{y_i}, \sigma_{z_i})\) pair, and for this reason eqns \((11)\) must be reconsidered.

Let us consider an observation point \((y_i, z_i)\), which naturally defines four quadrants \(Q_j\), in the \((Y_i, Z_i)\) plane, characterized by well defined pairs of errors \((\sigma_{y_i}^j, \sigma_{z_i}^j)\), as shown in Fig. \(1\). Each of these quadrants could in principle be intersected by the fitting function \(Z_i = F(Y_i; D, C)\). In our case, the fitting function is a straight line, and then for a fixed \((D, C)\) pair, only three quadrants will be intersected. For each of these three quadrants, we replace eqns \((11)\) with

$$\ln L_i = -\inf_{(Y_i, Z_i) \in Q_j} \left[ \frac{y_i - Y_i}{\sigma_{y_i}^j} \right]^2 + \left[ \frac{z_i - F(Y_i; D, C)}{\sigma_{z_i}^j} \right]^2.$$  \hspace{1cm} (12)

For the \(i\)-th observational point, eqns \((12)\) allow us to define three candidate contributions \((\ln L)_{ij}^i\) to the likelihood, one for each of the three intersected quadrants \(Q_j\). These points are those that maximize the likelihood in each single intersected quadrant. Among these possible contributions, only the one with the smallest absolute value is chosen, so that:

$$\ln L = \sum_i (\ln L)_i = -\sum_i \min_j \left| (\ln L)_{ij}^i \right|.$$  \hspace{1cm} (13)

This generalization allows us to solve the problem of fitting in the logarithmic plane, whilst maintaining intact the basic idea of the Structural Analysis method: eqn \((12)\) naturally reduces to eqns \((11)\) when errors become symmetric.
unique potential well only. are embedded in exactly the same dark matter halo. The again supposing that all the luminous components of dSphs We can gain some useful insights into the correlation by 3 ISOTHERMAL DISTRIBUTION SENSE. First, there is a range of pairs (ing a precise fit to the data in Table 1 makes very little it is steeper than both of them. value between the best fit exponents of the two subsets, but This is after all what was expected at outset as only a pure power-law dark matter potential can give a power-law cor-

fits 3\;\text{ISOTHERMAL DISTRIBUTION}

Table 2. The values of D and log C together with the \chi^2, using fitting in the natural plane (upper table) and logarithmic plane (lower table)

| Subset         | D    | log C | \chi^2 |
|----------------|------|-------|--------|
| Entire sample  | 3.7±0.75 | -0.76±0.55 | 76     |
| Classical dSphs| 3.5±2.8 | -0.78±1.5 | 10.5   |
| Faint dSphs    | 3.3±0.6 | -0.22±0.7 | 47     |
| Entire Sample  | 3.9±0.9 | -1.04±0.8 | 84     |
| Classical dSphs| 3.6±2.6 | -0.87±1.5 | 10     |
| Faint dSphs    | 3.5±0.85 | -0.54±0.85 | 62     |

2.4 Results

The fitting analyses have been performed in both the natural and logarithmic plane separately for the entire sample presented in Table 1 (28 objects). It is also interesting to split the sample into the classical dSphs – namely the first eight objects in Table 1 – and the next twenty objects – which are predominantly the ultrafaint dwarfs.

For each of the three samples, Table 2 lists the best-fit values of the exponent D and the logarithm of the coefficient C, together with errors corresponding to the 68% confidence regions, and finally the associated \chi^2 value. As might have been anticipated, the two chosen parameters are highly correlated, which contributes the largish uncertainties on the best fit values. Note that the results obtained in the natural and in the logarithmic plane can be different. This is not evident in the case of the classical dSphs, but this set comprises only 8 objects, which is perhaps too limited to constrain the parameters of the model. It seems possible that the classical dSphs and the ultrafaints are offset in C and D. This means that the confidence regions associated with the two subsets hardly overlap in the (D, log C) plane. As a consequence, the best-fit exponent of the entire sample is not an average value between the best fit exponents of the two subsets, but it is steeper than both of them.

It is a reasonable conclusion from Table 2 that claiming a precise fit to the data in Table 1 makes very little sense. First, there is a range of pairs (D, C) that are reasonably compatible with the data. Second, the best fit parameters still give quite poor fits, especially for the faint dwarfs. This is after all what was expected at outset as only a pure power-law dark matter potential can give a power-law correlation. In particular, the exponent of the correlation D cannot be constrained any better than saying that it satisfies 3.2 \lesssim D \lesssim 4.4. With the present data, establishing such an interval is a more sensible task than providing a single best-fit solution, whatever fitting method has been used.

3 ISOTHERMAL DISTRIBUTION FUNCTIONS

We can gain some useful insights into the correlation by again supposing that all the luminous components of dSphs are embedded in exactly the same dark matter halo. The correlation then necessarily reflects the properties of this unique potential well only.

As a flexible family of dark matter haloes, we use

$$\rho(r) = \frac{\rho_0}{\left(\frac{r}{r_0}\right)^a \left(1 + \left(\frac{r}{r_0}\right)^b\right)^{c/6}}. \quad (14)$$

Throughout the paper, we will concentrate in particular on the following three choices:

$$\begin{align*}
(a, b, c) &= (1, 1, 2) \\
(a, b, c) &= (0, 2, 3) \\
(a, b, c) &= (0, 2, 4)
\end{align*} \quad (15)$$

The first choice of parameters yields the cosmologically-motivated Navarro-Frenk-White (NFW) model. The second two choices are standard examples of cored models, one with the same asymptotic density fall-off as the NFW profile ($\rho \sim r^{-3}$), whilst the second with a faster fall-off ($\rho \sim r^{-4}$).

With an eye to later developments, it is also helpful to introduce a dimensionless half-light radius $\tilde{R}_h$ by scaling the true distance by the characteristic radius $r_0$ of the dark halo

$$\tilde{R}_h \equiv R_h/r_0. \quad (16)$$

Next, we use the characteristic density $\rho_0$ to define the dimensionless velocity dispersions:

$$\tilde{\sigma}^2 \equiv \frac{\sigma^2}{\Phi_0}, \quad \tilde{\sigma}^2_p(x) \equiv \frac{\sigma^2_p(r/r_0)}{\Phi_0}. \quad (17)$$

where $\Phi_0$ is the first non-constant term in the Taylor expansion of the potential.

Taking our inspiration from the observed flatness of dSph velocity dispersion profiles (see e.g., [Kleva et al. 2001] [Walker et al. 2007]), we assume that the stars have an isothermal Maxwellian distribution function [Evans et al. 2009]:

$$f_\ast(E) \propto \exp\left(-\frac{E}{\sigma^2}\right), \quad (18)$$

where $E$ is the energy and $\sigma$ is a constant. For this choice, clearly, any component of the stellar velocity dispersion is just the constant $\sigma$, which is also identical to the projected velocity dispersion $\sigma_p$.

Let us start with an NFW halo, which in its central parts, has the gravitational potential

$$\Phi_{\text{NFW}}(r) = -4\pi G \rho_0 \sigma^2 \left(1 - \frac{r}{2r_0}\right) + \mathcal{O}(r^3), \quad (19)$$

thus $\Phi_0 = 2\pi G \rho_0 r_0^2$. If the stellar component is embedded deep within the halo, the potential is approximately linear. Then $\delta = 1$ in eqns (9) and (10), and so the $\tilde{R}_h(\sigma_p)$ correlation is quadratic ($D = 2$). Specifically, for an isothermal distribution function, we get:

$$\tilde{R}_h \approx 2.027 \tilde{\sigma}^2_p. \quad (20)$$

As an alternative, let us take any cored model. Then, the gravitational potential is harmonic in the inner parts of the roughly constant density core:

$$\Phi_{\text{core}}(r) = -4\pi G \rho_0 \left[1 - \frac{1}{6} \left(\frac{r}{r_0}\right)^2\right] + \mathcal{O}(r^3), \quad (21)$$

so that $\Phi_0 = 4\pi G \rho_0 r_0^2/3$. As a consequence, the associated $\tilde{R}_h(\sigma_p)$ correlation is linear ($D = 1$), namely

$$\tilde{R}_h = \sqrt{2} \tilde{\sigma}_p. \quad (22)$$
In both cases, the exponent \( D \) of the correlation is significantly shallower than that inferred from the observational data \( (3.2 \lesssim D \lesssim 4.4) \).

Of course, whilst a power-law correlation is a fair starting point to perform a fit to the (modest) available data, it does not make a great deal of sense when translated in the context of models. The direct consequence of this assumption is the power-law parametrization of the potential itself. For systems whose luminous scalelength becomes comparable to the scalelength of the dark halo itself, significant deviations from a power-law correlation are expected. We need more complex models to give a proper description of the data, and it is to this subject that we now turn.

## 4 QUASI-ISOTHERMAL DISTRIBUTION FUNCTIONS

### 4.1 The Half-Light Radius and Central Velocity Dispersion

Henceforth, we use the full potentials associated with the dark matter halos \([14]\), and in particular with the three profiles given by \([15]\). As these tend to zero at large radii \( r \gg r_0 \), we use a lowered isothermal distribution function. An isothermal component has a divergent total mass if immersed in a potential in a regular potential at spatial infinity. We cannot allow such a behaviour, since then the half-light radius \( R_h \) is not a well defined quantity. We thus choose the quasi-isothermal or King distribution function \([16,17] \) for the stars:

\[
\rho_*(E) = \frac{\rho_{*0}}{(2\pi\sigma^2)^{3/2}} \exp \left( -E_K/\sigma^2 \right) - 1 ,
\]

where \( E_K = E - \Phi(r_\ast) \) and \( r_\ast \) is the tidal radius \([18] \).

The density and velocity dispersion profile generated by this distribution function \([19] \) are given in terms of the underlying gravitational potential as:

\[
\rho_*(r) = 4\pi \int_0^{\sqrt{2}r} f_\ast v^2 dv
= \rho_{*0} \left[ \exp \left( b^2 \right) \text{erf} \left( b \right) - \frac{2}{\sqrt{\pi}} \left( b + \frac{2b^3}{3} \right) \right] ,
\]

where \( f_\ast = \frac{\rho_{*0} v^2}{\rho_{*0} v^2} \) and \( \rho_{*0} \) is the central density of the King model. \( b \) is a constant, and is given by:

\[
\langle v^2 \rangle_\ast = \frac{4\pi}{\rho_\ast} \int_0^{\sqrt{2}r} f_\ast v^4 dv
= 3\sigma^2 \left[ 1 - \frac{8\theta^5}{15\pi \rho_\ast(r)/\rho_{*00}} \right] ,
\]

in which, for brevity, we have written

\[
b^2 = \left( b(r,\sigma, r_\ast) \right)^2 \equiv -\frac{\Phi(r) - \Phi(r_\ast)}{\sigma^2} .
\]

The equation that defines the half-light radius \( R_h \) is

\[
4\pi \int_{R_0}^{R_h} \rho_\ast \left( r, \sigma, r_\ast \right) \frac{rdr}{\sqrt{r^2 - R^2}} \equiv 2\pi R_h \Sigma_\ast \left( R_h \right) = \pi \int_{R_0}^{R_h} R dR \Sigma_\ast \left( R \right) ,
\]

while the projected velocity dispersion is given by

\[
\sigma_{p0}^2 \left( R \right) = \frac{2}{3\Sigma_\ast \left( R_h \right)} \int_{R_h}^{\infty} \rho_\ast \left( r \right) \langle v^2 \rangle_\ast \frac{rdr}{\sqrt{r^2 - R^2}} .
\]

Our models are determined by four dimensional scales: given a \( (r_\ast, \sigma) \) pair for the luminous component and a \( (r_0, \rho_0) \) pair for the dark halo potential \( \Phi \), the properties of the models are completely fixed. As the luminous components are tracers, the value of the parameter \( \rho_{*0} \), which fixes the central stellar density, does not need to be specified. It does not affect either the half-light radius \( R_h \) or the projected central velocity dispersion \( \sigma_{p0} \). If we instead use dimensionless parameters \([\text{see eqns (16) and (17)}] \), then we can complete the set by introducing, together with \( \tilde{\sigma} \), the dimensionless tidal radius \( \tilde{r}_t = r_t/r_0 \), which is the true tidal radius divided by the characteristic scale of the dark halo.

### 4.2 Theoretical Correlation Curves

Our aim is to describe the behaviour of the half-light radius \( R_h \) in the dimensionless parameter space \((\tilde{r}_t, \tilde{\sigma})\) which characterizes our models. Some of the asymptotic properties are analytic, as discussed in Appendix A.

In order to obtain a unique \( R_h(\tilde{r}_t, \tilde{\sigma}) \) correlation, we need to fix the second free parameter \( \tilde{\sigma} \). For reasons that will shortly become apparent, this we do by setting the structural parameter

\[
\kappa \equiv r_t/R_0 = \tilde{r}_t/R_h
\]

to be a constant. That is, we construct our \( R_h(\tilde{r}_t, \tilde{\sigma}) \) correlation by following one of the contours of \( \kappa \) in the \((\tilde{r}_t, \tilde{\sigma})\) plane \(\text{see for example Fig. A1}\) and characterise the correlations so obtained by \( R_h(\tilde{r}_t, \tilde{\sigma}) \).

By considering the photometric profiles of the eight classical dwarfs in \(\text{Irwin & Hatzidimitriou 1992}\), we can grasp a lower limit for \( \kappa \). The tidal radius \( r_t \) must be greater than the final radius \( r_{\text{last}} \) which has a significant non-zero photometric measure. From this, we conclude that all the classical dSphs have \( \kappa \gg 4 \). A somewhat crude, but nevertheless useful, upper value for \( \kappa \) can be given by arguing that it is highly improbable the tidal radius of a dSph is larger than a quarter or a fifth of its distance from the centre of the Galaxy, giving us the constraint \( \kappa \ll 100 \).

The eight classical dSphs also show projected velocity dispersion profiles which are flat out to the last measured points \(\text{see e.g., Walker et al. 2009}\). For Draco, this means out to approximately nine half-light radii; for Carina, Leo I
and Sculptor, out to approximately five. As a consequence, we restrict attention to models which satisfy the condition

\[ \sigma_p(\gamma R_h) \geq 0.9 \sigma_p(0). \]  

(30)

This ensures that the profile is flattish out to \( \gamma \) multiples of the half-light radius.

For a fixed value of \( \kappa \), the effect of imposing the condition \((30)\) is to define the model with the highest possible \( \sigma_p \) and \( R_h^\kappa \). This model represents the endpoint of the corresponding \( R_h^\kappa(\sigma_p,0) \) relation. Models beyond the endpoint are unacceptable as their velocity dispersion profiles do not resemble those of the dSphs. In other words, the higher the value of \( \gamma \), the shorter the theoretical relation \( R_h^\kappa(\sigma_p,0) \). For a fixed value of \( \kappa \), the curves obtained for large values of \( \gamma \) are exactly contained within those of smaller values of \( \gamma \). For example, the upper left panel of Fig. 2 shows the relation obtained embedding our models with fixed \( \kappa = 10 \) within a NFW halo; the two different displayed endpoints correspond to \( \gamma = 8 \) and \( \gamma = 4 \). We note that the curves follow the asymptotic quadratic relation \((20)\), which is valid in the regime in which \( \sigma_{p,0} \ll 1 \).

The upper right panel of Fig. 2 shows the effects of varying the value of the parameter \( \kappa \). The theoretical relations corresponding to \( \kappa = 10 \) and \( \kappa = 100 \) are plotted together.

Figure 2. The dimensionless half-light radius \( R_h^\kappa \) is plotted against the dimensionless projected central velocity dispersion \( \sigma_{p,0} \). Upper left: The theoretical relation \( R_h^{10} \) with the two endpoints corresponding to \( \gamma = 4 \) and \( \gamma = 8 \) for an NFW profile. The red line shows the analytic and asymptotic result \((20)\). Upper right: The curves \( R_h^{10} \) and \( R_h^{100} \) corresponding to \( \gamma = 6 \) for an NFW profile. Lower left: The curves \( R_h(\sigma_{p,0}) \) for the cored dark matter density distribution \((14)\) with \( (a=0, b=2, c=3) \). The red line shows the analytic and asymptotic result \((22)\). Lower right: The curves \( R_h(\sigma_{p,0}) \) for the cored dark matter density distribution \((14)\) with \( (a=0, b=2, c=4) \).

Figure 3. The 68\% and 95\% confidence regions of the likelihood \((31)\) respectively for the dark matter halo profiles. Full dots indicate the best fit models. The red shaded areas in the \((r_0, \rho_0)\) plane are forbidden by the constraints \((33)\) and \((34)\). Here, \( r_0 \) is measured in pc, and \( \rho_0 \) in \( M_{\odot} pc^{-3} \).
The displayed endpoints are fixed in both cases by $\gamma = 6$. The $\kappa = 100$ relation extends to higher values of $\sigma$ and $R_h$. In this case, the inclusion between the two relations is not exact. However, in their common domain, the two relations are still roughly identical, given that the observational errors on the available $\sigma_{\text{p,0}}$ and $R_h$ data-points are quite large.

The properties of the $R_h(\sigma_{\text{p,0}})$ relations produced by our models allow us to impose a useful simplification. We need consider only the $R_h(\sigma_{\text{p,0}})$ relation produced by the highest plausible value of $\kappa$ and lowest plausible value of $\gamma$. Provided we overestimate $\kappa$ and underestimate $\gamma$, we can be sure that, besides small deviations, all possible $R_h(\sigma_{\text{p,0}})$ curves obtained for other realistic ($\kappa, \gamma$) pairs are in fact contained in ours. As a consequence, we define our final $R_h(\sigma_{\text{p,0}})$ relation as that obtained by the pair $(\kappa, \gamma) = (100, 6)$. With this set, the reader may worry that we are overestimating $\gamma$ and hence overlooking some viable models. However, we have shown that the effect on the $R_h(\sigma_{\text{p,0}})$ relation of a lower $\gamma$ is similar to that of a higher $\kappa$. Since we are surely overestimating $\kappa$, we are not losing models at the end of the $R_h(\sigma_{\text{p,0}})$ relation. The upper right panel of Fig. 2 therefore shows our final $R_h(\sigma_{\text{p,0}})$ relation for the NFW dark halo. Its endpoint for $\gamma = 6$ corresponds to $(\sigma_{\text{p, max}}, R_h, \text{max}) \approx (0.35, 1.2)$. The lower panels of Fig. 2 show the final $R_h(\sigma_{\text{p,0}})$ relations obtained for the two cored profiles of eqn. (13) with $(a = 0, b = 2, c = 3)$ and $(a = 0, b = 2, c = 4)$ respectively. To recapitulate, the former has a density fall-off like $r^{-3}$ at large radii, the latter like $r^{-4}$. Their endpoints are respectively at $(\sigma_{\text{p, max}}, R_h, \text{max}) \approx (0.66, 1.5)$ and $(\sigma_{\text{p, max}}, R_h, \text{max}) \approx (0.54, 1.65)$. In both panels the red profiles show the asymptotic relation for the three dark matter profiles of eqn. (22). It is clear that the extrapolation of such an approximation out of the regime $\sigma_{\text{p,0}} \ll 1$ is unreliable. For larger velocity dispersions, the true $R_h(\sigma_{\text{p,0}})$ has, in all cases, a much steeper dependence on $\sigma_{\text{p,0}}$ than the power-law correlation obtained in the asymptotic limit. Finally, even if the qualitative shape of the three different $R_h(\sigma_{\text{p,0}})$ relations obtained for the three dark matter profiles of eqn. (13) remains similar, when looked in detail, they show several quantitative differences.

Figure 4. The $\rho_0(r_0)$ functional dependence generated by different halo profiles using Fornax as an illustrative example. Here, $r_0$ is measured in pc, and $\rho_0$ in $M_{\odot}\,\text{pc}^{-3}$. For each $(r_0, \rho_0(r_0))$, the Fornax half-light radius and central velocity dispersion is mapped onto the underlying theoretical $R_h(\sigma_{\text{p,0}})$ curve. The endpoints of the relations are marked with filled circles. (Full line, dashed line and dash-dotted line represent respectively the dark matter density profiles as ordered in eqn. (13).) For comparison, the red shaded area indicates the expectations of a $\Lambda$CDM cosmological model (see text for further description).

5 THE UNIVERSAL HALO HYPOTHESIS

For the moment, we continue to assume that the luminous components of all the dSphs are embedded in exactly the same universal dark halo, which in turn has one of the three density profiles of eqns. (14) and (15). By fitting the data-points in Table 1 to the respective $R_h(\sigma_{\text{p,0}})$ relation, we can both measure the best characteristic central density $\rho_0$ and radius $r_0$ associated with each profile and compare the quality of the three fits to gain insight into the closest match to actual dark halos, at least under the hypothesis of universality. We use the technique of Structural Analysis introduced in Section 2.3. The likelihood we maximize is

$$L(\rho_0, r_0) = -\frac{1}{2} \sum_{i=1}^{N} \left( \frac{y_i - Y_i \sqrt{\Phi_0}}{\sigma_{y_i}} \right)^2 + \left( \frac{z_i - r_0 R_h(Y_i)}{\sigma_{z_i}} \right)^2,$$

in which $(y_i, z_i)$ are the $N$ observed values of the data pairs with standard deviations $(\sigma_{y_i}, \sigma_{z_i})$, and $R_h(Y_i)$ is the theoretical $Z_i = R_h(\sigma_{\text{p,0}})$ relation generated by the halo model, whilst $\Phi_0$ has been defined in Section 3. The $N$ equations which determine the values of $Y_i$ are

$$\frac{\partial \ln L}{\partial Y_i} = \sqrt{\Phi_0} \frac{y_i - Y_i \sqrt{\Phi_0}}{\sigma_{y_i}^2} + r_0 \frac{\partial R_h}{\partial Y_i} \left( \frac{z_i - r_0 R_h(Y_i)}{\sigma_{z_i}^2} \right) = 0,$$

which can be easily solved numerically.

Note that each of the theoretical $R_h(\sigma_{\text{p,0}})$ relations comes with its own bounded domain of validity $(\sigma_{\text{p,0}} \in [0, \sigma_{\text{p, max}}], R_h \in [0, R_h(\sigma_{\text{p, max}})])$. This implies the existence of constraints for the possible physical parameters, $\rho_0$ and $r_0$. First of all, considering the Sgr dSph which has the largest half-light radius in Table 1 then we must have

$$r_0 R_h, \text{max} \gtrsim 1550 \, \text{pc}.$$ (33)

This must be so, otherwise the objects with the largest
three different dark matter profiles. Furthermore, the existence of an endpoint model \((\sigma_{p,\text{max}}, R_h(\sigma_{p,\text{max}}))\) implies constraints for \(r_0\) and \(\rho_0\), which respectively have a minimum and maximum available value for each dSph. For future use, we denote the minimum characteristic radius by \(r_{0,\text{min}}\) and the associated maximum value of the characteristic density by \(\rho_{0,\text{max}} = \rho_0(r_{0,\text{min}})\). Notice also that, for the specific case of the NFW profile, the existence of such a functional dependence naturally corresponds to the existence of an analogous link between the characteristic radius \(r_0\) and the concentration parameter \(c\), which, for clarity, we recall to be defined by

\[
\rho_0(c) = \frac{600 H_0^2}{8 \pi G} \ln(1 + c) - c/(1 + c) ,
\]

in which \(H_0\) is the Hubble constant [Navarro et al. 1996]. As a consequence, for each dSph, the maximum characteristic density \(\rho_{0,\text{max}}\) also corresponds to a maximum compatible value of the concentration, \(c_{\text{max}}\).

It is useful to compare to the predictions of numerical simulations of halo formation in ΛCDM. So, Fig. 4 also displays the translation in the \((\rho_0, r_0)\) plane of the \(c_{\text{max}}(M_{\text{vir}})\) functional relationships calculated in Bullock et al. [2001] and in Kuhlen et al. [2003] (see especially their Fig. 9). For the specific case of Fornax, this allows us to compare the

| Object | \(c_{\text{max}}\) | \(r_0(c = c_1)\) | \(r_0(c = c_u)\) |
|--------|----------------|-----------------|----------------|
| Carina | 44 | 640 \pm 150 | 285 \pm 70 |
| Draco  | 66 | 1030 \pm 200 | 460 \pm 80 |
| Fornax | 32 | 1090 \pm 100 | 560 \pm 60 |
| Leo I  | 56 | 970 \pm 200 | 450 \pm 90 |
| Leo II | 63 | 730 \pm 120 | 340 \pm 50 |
| Sculptor | 54 | 950 \pm 170 | 440 \pm 60 |
| Sextans | 23 | 1060 \pm 220 | 525 \pm 100 |
| UMi   | 52 | 970 \pm 170 | 470 \pm 80 |
| Bootes 1 | 44 | 620 \pm 200 | 265 \pm 80 |
| Bootes 2 | 195 | \(\leq 3600\) | \(\leq 1200\) |
| C Ven I | 26 | 805 \pm 20 | 490 \pm 40 |
| C Ven II | 81 | 560 \pm 200 | 160 \pm 70 |
| Coma | 79 | 570 \pm 120 | 210 \pm 70 |
| Hercules | 22 | 500 \pm 120 | 250 \pm 80 |
| Leo IV | 46 | 300 \pm 150 | 70 \pm 60 |
| Leo V | 76 | 150 \pm 280 | 60 \pm 100 |
| Leo T | 61 | 800 \pm 300 | 120 \pm 50 |
|Segue 1 | 154 | 920 \pm 400 | 130 \pm 80 |
|Segue 2 | 115 | 250 \pm 400 | 90 \pm 90 |
|UMa I | 56 | 1150 \pm 600 | 400 \pm 180 |
|UMa II | 67 | 730 \pm 200 | 180 \pm 70 |
|Williman 17 | 171 | 700 \pm 1000 | 200 \pm 200 |
|And II | 16 | 1650 \pm 450 | 850 \pm 250 |
|And IX | 25 | 770 \pm 220 | 370 \pm 140 |
|And XV | 59 | 1000 \pm 1900 | 400 \pm 400 |
|Cetus | 46 | 1680 \pm 250 | 820 \pm 160 |
|Sgr | 16 | 2250 \pm 200 | 1325 \pm 150 |
|Tucana | 95 | 2200 \pm 600 | 580 \pm 250 |

6 SINGLE OBJECT ANALYSES

6.1 The Method

As an illustrative example, let us consider the Fornax data-point, \((\sigma_{p,0}, R_h)_{\text{Fornax}}\), given in Table I. The problem is to choose the scaling \((r_0, \rho_0)\) that maps this dimensional data-point to models on the underlying dimensionless theoretical curve \(R_h(\sigma_{p,0})\) for a given dark halo profile.

As we have seen, any halo profile is naturally associated with a specific domain \([0, \sigma_{p,\text{max}}]\). For each \(\sigma_{p,0}\) in this interval, one and only one \((r_0, \rho_0)\) pair exists such that

\[
(\sigma_{p,0}, R_h)_{\text{Fornax}} = (\sigma_{p,0}\sqrt{\Phi_0}, r_0 R_h(\sigma_{p,0})) .
\]

We can thus define a one-to-one relation between the two characteristic scales \(\rho_0\) and \(r_0\), which we simply indicate by \(\rho_0(r_0)\), or equivalently by \(r_0(\rho_0)\). Each of the \((r_0, \rho_0)\) points corresponds to a different position of the Fornax data-point on the theoretical \(R_h(\sigma_{p,0})\) relation.

Fig. 4, for example, displays the three \(\rho_0(r_0)\) functional dependences generated by our three canonical dark matter profiles, together with the respective endpoints, for the case of Fornax. Clearly, the dependence changes for each of the half-light radii may not be compatible with the theoretical \(R_h(\sigma_{p,0})\) correlation in the universal approach. In the same way, by taking Tucana dSph which has the largest central projected velocity dispersion of \(\sigma_{p,0} = 15.8\) km s\(^{-1}\), we obtain

\[
\sigma_{p,\text{max}}\sqrt{\Phi_0} \geq 15.8\text{km s}\(^{-1}\) .
\]

Note that the detailed values provided by the constraints \((33)\) and \((34)\) depend on the dark matter profile under consideration, which fixes the endpoint \((\sigma_{p,\text{max}}, R_{h,\text{max}})\) of the curve \(R_h(\sigma_{p,0})\). This suggests that the best set of parameters in which to perform the numerical maximization of the likelihood \((31)\) is \([1/r_0, 1/\sqrt{\Phi_0}]\). We can easily transform our final results into the natural plane \((r_0, \rho_0)\), as shown in Fig. 4 for the three dark matter profiles. The 68% and 95% confidence regions according to the likelihood are shown. Note that the constraints \((33)\) and \((34)\) have been used in displaying the \((r_0, \rho_0)\) plane: in each panel, the red shaded area is forbidden. Table III lists the best-fit models, which are indicated with full dots in Fig. 4 together with the associated \(\chi^2\) value. Also listed are the results of similar fitting analysis performed on the classical dSphs only – that is, the first eight dSphs in Table I. In this case, the dSphs with the largest \(R_h\) and \(\sigma_{p,0}\) are respectively Sextans and Fornax.

The results of this analysis are not particularly encouraging for the universal halo hypothesis. Even if the NFW profile seems to provide the best fit to the available data, the resulting physical parameters \((r_0, \rho_0)\) do not make great sense, at least as judged from the expectations of numerical simulations of dwarf galaxy formation. The best fit characteristic radius is in fact extremely large, namely 4.9 kpc, while the characteristic density (to balance the extreme value of \(v^2_{\text{c}}\)) is very low, corresponding to a concentration parameter as small as \(c \approx 6.5\). If the analysis is restricted to the classical dSphs, the fits are better, suggesting that the universal approach is less inadequate.
Figure 5. Mass profiles $M(r)$ for Fornax assuming an NFW halo (upper left) and cored haloes (upper right and lower left). The curves correspond to different values of the halo scalelength $r_0$, in particular $1$ (in red), $2$, $4$, $8$ and $15$ (in black) times the minimum characteristic radius $r_{0,\min}$ associated with each dark matter profile. The lower right panel is a superposition of all the mass profiles in the preceding panels and shows the existence of a special radius at which – despite our ignorance of best choice of model or scalelength – the uncertainty in the enclosed mass is minimised.

expectations of a $\Lambda$CDM cosmological model with our results. It is interesting to note that the models selected by the $\Lambda$CDM scenario for the Fornax dSph are those at the highest end of the $\rho_0(r_0)$ relation, and that thus have a luminous and a dark component with similar scale lengths ($\approx 1$).

Now let us consider the halo enclosed mass:

$$M(\rho_0, r_0, r) \equiv \int_0^r 4\pi x^2 \rho(\rho_0, r_0, x) \, dx. \quad (37)$$

The existence of the functional relation $\rho_0(r_0)$ allows us to reduce the dimensionality of the free parameters of the halo mass functions by one, so that $M(\rho_0, r_0, r) = M(\rho_0(r_0), r_0, r) \equiv M(r_0, r)$. Thus, to produce a mass measure for each of the dSphs, we need only a dark matter profile and a characteristic radius.

However, even given our ignorance of the correct choice of dark matter halo density profile and characteristic radius $r_0$, we can still provide a reliable mass measure for each of the dwarfs by using only the single $(\sigma_p, 0, R_h)$ datapoint. This unexpected result is illustrated explicitly in the case of Fornax by Fig. 5. This shows the enclosed halo mass (37) for different plausible choices of $r_0$ and for our three different dark halo density laws (15).

It is clear that for each density profile there is a special radius $r_{\text{spec}}$, inside which the uncertainty of the mass measure caused by our ignorance of the characteristic radius $r_0$ is minimized. If we know the dark matter density law – for example from cosmological considerations – then the error that we make in measuring the enclosed mass in the absence of accurate information on the characteristic radius is barely $\lesssim 10\%$. Surprisingly, the locations of the special radii for the different halo profiles are close to each other, and amount to on average a value of $r_{\text{spec}} \approx 1.7 R_h$. This result derives directly from the shapes of the $\hat{R}_h(\sigma_p, 0)$ functional relations defined by our lowered isothermal models with flattish projected velocity dispersion.

Although we have shown the results of our calculation for Fornax only, we have carried out similar computations for all the dSphs, which leads to the conclusion that a uniformly good choice for the radius $r_{\text{spec}}$ is $r \approx 1.7 R_h$. (For the sceptical reader, this result can also be inferred from our later Fig. 7.)

This result has some superficially similar analogues in the recent literature. A number of authors have looked at classes of anisotropic models from a Jeans equation perspective and argued that the mass within the half-light radius is robust against changes in the halo model and the anisotropy (see e.g., Strigari et al. 2007a; Walker et al. 2009c; Wolf et al. 2010). The Jeans equations of course are a weak constraint, and there is an enormous freedom in solving for the enclosed mass in terms of multi-parameter models of the light and anisotropy profiles. Very often, the Jeans solutions are meaningless in that, although the stresses are positive, there is no physical distribution. A classic example is provided by cored light profiles in cusped dark halo profiles. The Jeans solution exists and is physical – but there
Figure 6. For each dSph, the upper panel displays the probability distributions for the halo mass inside $1.7R_h$, while the lower panel the distribution for the characteristic halo radius $r_0$. Black and red represent the lower and upper concentration parameter NFW models. Masses are in units of $10^7 M_\odot$, while radii are measured in kpc.
Figure 7. Comparison with previous mass measures, all radial coordinates are scaled to the half-light radius, while masses are measured in $M_\odot$. The shaded regions represent respectively: NFW halos with characteristic radii from $r_{0,\text{min}}$ to $5r_{0,\text{min}}$ in gray; cored halo with fall-off $\rho \sim r^{-3}$ with characteristic radii from $r_{0,\text{min}}$ to $5r_{0,\text{min}}$ in red; cored halo with fall-off $\rho \sim r^{-4}$ with characteristic radii from $r_{0,\text{min}}$ to $5r_{0,\text{min}}$ in green. The empty triangle indicates our mass measure inside $1.7R_h$ with error from the observational uncertainties. The blue filled circle is the measure of the mass enclosed in $R_h$ from Walker et al. (2010). Black dots are measures from Strigari et al. (2007) and (2008). The additional empty circle for Segue 1 is a measure from Geha et al. (2009).
is no distribution function \cite{An & Evans 2009}. Hence, this approach leads to weak constraints as a wide class of models, many of which are unphysical – being considered. Once this has been appreciated, it is less surprising that quasi-isothermal phase-space models allow us to fix the mass to much greater radii.

Hence, the Jeans based analyses are studying how the mass estimates are affected by anisotropy. They show that the mass within about a scale radius is largely unaffected by assumptions as to anisotropy. Our starting point here is different. We are arguing that the central parts of the dSphs are surely isotropic and are happy to build that hypothesis into our models. Our result is that by varying the halo profiles, the mass within 1.7 half-light radii is largely unaffected by changes in the mass profile.

6.2 The Masses of the dSphs

In fact, the effect of the observational errors on the mass measure inside $r_{\text{spec}}$ is larger than any differences caused by changing the halo models. To quantify this, we use a Monte Carlo numerical method. For each dSph, we study the distribution of the mass measures generated by assuming a normal distribution for the observational $(\sigma_{\mu,0}, R_h)$ points with means and standard deviations as in Table 1. Each random point extracted from these probability distributions generates a $r_0(r_0)$ relation, as in eqn (33), and hence a mass measure. We perform this analysis only for the NFW halo on the grounds of its cosmological importance. The value of $r_0$ used to produce each mass measure is calculated by fixing the concentration $c$ to a lower $c_1$ and an upper $c_0$ value.

Numerical simulations suggest that dSphs have dark haloes of NFW form with concentration values $20 \lesssim c \lesssim 30$ \cite{Navarro et al. 1996}, which we use as lower and upper bounds in the following analysis. However, we have already shown that there is a maximum value of the concentration $c_{\text{max}}$ for each dSph consistent with it falling on the theoretical curves $R_h(\sigma_{\mu,0})$. For some of the dSphs, $c_{\text{max}}$ is already less than 30 – namely, in the cases of Sextans ($c_{\text{max}} \approx 23$), C Ven I ($c_{\text{max}} \approx 26$), Hercules ($c_{\text{max}} \approx 22$), And II ($c_{\text{max}} \approx 16$), And IX ($c_{\text{max}} \approx 25$), and Sagittarius ($c_{\text{max}} \approx 16$). For these dSphs, we use $2c_{\text{max}}/3$ and $c_{\text{max}}$ as our lower and upper bounds.

The results are displayed in Fig. 6 which show the probability distributions for the enclosed masses $M(1.7R_h)$ and the characteristic scalelengths $R_0$ of the NFW halo. Note that, even though the two different assumed concentrations imply rather different distributions for the scalelengths, this has almost no effect on the mass measures because of our choice of $r_{\text{spec}} = 1.7R_h$. Only the objects with the smallest relative errors in the observational data, such as Fornax, C Ven I and Sgr, have recognizable differences in the modes of the two mass distributions, although these remain small compared to the intrinsic spread of the distributions.

The shape of the mass distributions in Fig. 6 depends critically on the quality of the observational data for the central velocity dispersion $\sigma_{\mu,0}$ or half-light radius $R_h$. If the uncertainties are small, then the probability distribution is a well-behaved, nearly normal function. We extract the mode or most likely value for the mass, together with the values at half maximum, which gives the the range (recall that for a Gaussian, the full-width at half maximum is $\approx 2.35\sigma$). However, if the observational datapoints are poor, the probability distribution acquires a highly skewed structure. The objects that suffer from this difficulty are Bootes 2, Leo IV, Leo V, Segue 2 and And XV. Here, the peak of the distribution occurs at mass scales too small to be clearly resolved, and we can give only the mass at half the maximum as an upper limit. A similar, but less dramatic problem, is encountered in Willman 1, and And IX. The mass estimates and range are listed in the first column in Table 5.

The shape of the distributions for the characteristic scalelength $R_0$ of the NFW haloes in Fig. 6 are well-behaved, although they do depend on the assumed concentration parameter. For the upper and lower values for the concentration, the mode or most likely value for $r_0$, together with the values at half maximum are listed in Table 4. Using the $(\sigma_{\mu,0}, R_h)$ datapoint only, we are unable to determine $r_0$ with any confidence. To achieve this, we would need further observational information, particularly the detailed photometric and kinematic profiles, which are only available for a handful of the brightest dSphs. Nevertheless, we can still deduce an interesting conclusion. The listed maximum concentration parameters $c_{\text{max}}$ range from values as small as 16 for Sgr and And II to very high values (over 150) for the faintest objects like Segue 1, Bootes 2 and Willman 1. The nearer $c_{\text{max}}$ is to the values of the concentration suggested by the cosmological simulations, the nearer the values inferred for the two characteristic radii $r_0(c = c_1)$ and $r_0(c = c_0)$ are to the value of the observational half-light radius $R_h$. However, a number of predominantly the fainter dSphs have a much higher $c_{\text{max}}$ (seven have $c_{\text{max}} > 70$, for example). Such dSphs, then, either they have a central density which is significantly higher than suggested by cosmological N-body simulations (see upper panel in Fig. 5), or they have a $r_0/R_h$ ratio which is much higher than unity. The only way to construct a model with a concentration in tune with cosmological predictions for is in fact to pick a larger scale radius $r_0$ for the halo. If we thus believe in the indications given by the simulations, we must also accept the existence of a systematic inverse trend between luminosity and the ratio $r_0/R_h$ (see lower panel in Fig. 5). In other words, the fainter luminosity dSphs must be embedded more deeply within their dark matter haloes.

Also given in Table 5 are the values $M^-$ and $M^+$, which respectively indicate the lowest and the highest mass measure inside $1.7R_h$ among all the models considered in Fig. 5. The spread $M^+ - M^-$ thus approximately represents the uncertainty due to the models only. Since the observational uncertainties are uniformly larger, it is evident that our method of measuring halo masses can provide useful information simply by improving the quality of the data in Table 4 before embarking on the laborious process of obtaining detailed photometric and kinematic profiles for each of the 28 dwarfs. Using the luminosities listed in Walker et al. (2009c), mass-to-light ratios are calculated accordingly.

For most of the dSphs, these measures extend out to much larger radii than previously.

Table 5 also lists a series of previous mass measurements, namely the measures of the mass enclosed in $R_h$ from Walker et al. (2010) and the measures of the mass inside 100pc, 300pc and 600pc from Strigari et al. (2007, 2008) A comparison between our results and the listed ones is displayed in Fig. 7. Note that the spread of the mass functions displayed in Fig. 7 only show the deviations of the
mass measures due to our uncertainty on the correct dark matter density profile and characteristic radius. Again, the importance of the choice $r_{\text{spec}} = 1.7R_h$ is self-evident.

6.2.1 The Common Mass Scale – Hercules and Leo IV

Strigari et al. (2008) claimed that all the dSphs shared a common mass scale of $\sim 10^7 M_\odot$ within 300 pc. It is worth noting at outset that, for the faintest objects like Willman 1 and Segue 1, 300 pc corresponds to nearly 10 half-light radii. The available data is limited to the very central parts of the putative dark halos, and the result is based on a stupendously bold extrapolation.

Nonetheless, it is interesting to see if there are any objects for which the assertion of Strigari et al. (2008) may be disproved. Fig. 7 shows that even within the uncertainties, Hercules and Leo IV are not compatible with a common mass scale of $\sim 10^7 M_\odot$. For Hercules, this is because the central velocity dispersion was revised downward to $3.7 \pm 0.9$ kms$^{-1}$ by Adén et al. (2009a), who used Strömgren photometry to discriminate between foreground Milky Way dwarf stars and Hercules giants. This comparison takes into account the $5.1 \pm 0.9$ kms$^{-1}$ originally given by Simon & Geha (2007). Even allowing for generous model uncertainties, Fig. 7 shows the mass of Hercules lies below $3 \times 10^6 M_\odot$ at 300 pc ($\approx 0.9R_h$). This agrees with the conclusion of Adén et al. (2009). Another counter-example is provided by Leo IV. Here, 300 pc corresponds to $\approx 2.6R_h$, and Fig. 7 shows that the interior mass is again at most $3 \times 10^6 M_\odot$.

We conclude that the common mass-scale of Strigari et al. (2008) is illusory. There are indeed dSphs for which the mass within 300 pc is $\sim 10^7 M_\odot$, such as Draco, U Mi and Fornax. There are also objects with such small half-light radii, like Segue 1 and Willman 1, that the uncertainties of the mass extrapolation to 300 pc are large. Consequently, they may be accommodated within a halo of mass $\sim 10^8 M_\odot$, although much smaller masses are probable. However, there are some objects with intermediate half-light radii, like Hercules and Leo IV for which the common mass-scale clearly fails.

It is worth noting that the agreement between our masses and earlier investigators is often good. Where results do differ – as for example in Hercules or Leo IV – it is often because velocity dispersion measurements have been revised. Such revisions are always downwards, as contaminants, binary stars and variables always introduce additional scatter if not properly accounted for.

6.2.2 Global Correlations

We are able now to refine the plot by Mateo (1998), also extended later by Gilmore et al. (2007). In both these works,
Figure 9. Correlations between global properties (luminosity, mass and mass-to-light ratios) of the dSphs as inferred in the present paper. Red indicates the dSphs associated with Andromeda.

Figure 8. Upper panel: The correlation between luminosity and concentration parameter necessary to accommodate a NFW model with $r_0/R_h \approx 1$ for all dSphs. Lower panel: the correlation between the depth of the embedding within the dark halo ($r_0/R_h$) and luminosity necessary to accommodate a NFW model with a similar value of the concentration ($c \approx 20$) for all dSphs.

Furthermore, it is interesting to look at the correlations between mass and mass-to-light ratio, as well as between luminosity and mass (see the left and right panels in Fig. 9). In both these planes, our 28 objects sample shows evidence for a correlation, although that considered first by Mateo (1998) seems to be the one with the smallest scatter.

It is curious that the correlation between mass and mass-to-light ratio is in fact the one with the largest scatter. This is slightly unexpected, since the mass and mass-to-light ratio plane is the one in which the use of the same mass diagnostic technique for all the objects is likely to have introduced the largest (regularizing) effect. By contrast, in the other two planes, the luminosity comes just as a label for the dSphs, which is completely external to the phase space modelling. This may suggest that the formations phases of the dSphs are characterized by a physical mechanism that links directly the mass size of their dark matter halo, and then potential well, to the extent and properties of the star formation history they have undergone.

6.2.3 Segue 1

Segue 1 has received considerable attention recently as one of the most promising targets for indirect detection of dark matter using the gamma ray signal generated by self-annihilation of neutralinos (Scott et al. 2010). This is because Segue 1 is relatively nearby ($\sim 23$ kpc) and has a high Galactic latitude. Geha et al. (2009) measured an internal velocity dispersion of $4.3_{-1.2}^{+1.8}$ km\,s$^{-1}$, which led them to claim an enormous mass-to-light ratio within 50 pc of $\sim 1320$. Our analysis is consistent with this result. Although the masses are comparable to the results of Strigari et al. (2007b, 2008), there are reasons to be cautious.

The case that Segue 1 is a dark-matter dominated dwarf galaxy has been seriously undermined by the recent work of Niederste-Ostholt et al. (2009), who showed that there is strong evidence from the Sloan Digital Sky Survey photometry for tidal effects in the outer parts of the object. This is hard to explain if indeed Segue 1 is embedded within a dark halo with mass $\sim 10^6$, as its tidal radius would then be much larger than its half-light radius. Still more worryingly, Niederste-Ostholt et al. (2009) raised the possibility of contamination in the Geha et al. (2009) sample, which may cause an artificial inflation of the central velocity dispersion. Segue 1 is in a confused area of the sky, close to both leading and trailing wraps of the Sgr stream and the Orphan stream. Niederste-Ostholt et al. (2009) showed that...
7 A SIMPLE MASS ESTIMATOR

It is helpful to summarise our results in a simple approximating expression which links the halo mass \( M(1.7R_h) \) to the available observational data \( (\sigma_{p,0}, R_h) \) (c.f. Illingworth 1976, Walker et al. 2010). The results listed in Table 5 lead to:

\[
M(1.7R_h) \equiv K \frac{\sigma_{p,0}^2 R_h}{G} \approx (5.8 \pm 1) \frac{\sigma_{p,0}^2 R_h}{G} .
\]

The scatter in the value of the coefficient \( K \) is caused by at least two different effects. First, the value of the concentration chosen to construct Table 5 correspond, for different dSphs, to different ratios \( \sigma_0/R_h \). However, as shown in Figs. 5 and 7, these differences generate rather small differences in the enclosed mass \( M(1.7R_h) \). Secondly, the shape of the probability distributions in Fig. 8 is influenced by the importance of the observational errors; in particular the peak of the complete probability distribution does not always coincide with the value of the mass we would obtain by simply ignoring the observational uncertainties.

A slightly different version of eqn. (38) can be obtained by analytic methods. Let us consider the halo mass function, which can be written as,

\[
M(r_0, r) = \rho_0(r_0)^3 r^3 g(r/r_0) ,
\]

and is valid for any choice of the halo density profile. We take the limit \( R_h/r_0 \to 0 \), appropriate for the case in which the stars are deeply embedded in the halo. Let us suppose that the central asymptotic behaviour of the halo density is \( \rho \propto (r_0/r)^{2-\delta} \), so that the gravitational potential in the central regions is simply \( \phi \propto (r/r_0)^{3/2} \). We obtain \( g \propto (r/r_0)^{1+\delta} \), while – using eqn. (1) – for the dimensionless half light radius we have \( R_h \propto \sigma_{p,0}^{2/3} \). Through eqn. (38), this determines the asymptotic functional relation \( \rho_0(r_0) \) between the inferred central density of the halo as a function of its characteristic scalelength : \( \rho_0 \propto r_0^{-7/2} \). By combining the information gathered so far in eqn. (39), we get for the NFW profile

\[
\lim_{R_h/r_0 \to 0} M_{\text{NFW}}(r_0, 1.7R_h) \equiv K_{\text{NFW}} \frac{\sigma_{p,0}^2 R_h}{G} 
\approx 5.78 \frac{\sigma_{p,0}^2 R_h}{G} ,
\]

while, for any cored model,

\[
\lim_{R_h/r_0 \to 0} M_{\text{core}}(r_0, 1.7R_h) \equiv K_{\text{core}} \frac{\sigma_{p,0}^2 R_h}{G} 
\approx 6.8 \frac{\sigma_{p,0}^2 R_h}{G} .
\]

In terms of our phase space models, the coefficient \( K \) is clearly a function of the radius within which the mass is calculated (typically \( 1.7R_h \) in this paper) as well as the ratio \( R_h/r_0 = \hat{R}_h \) for the specific halo model. In particular, we have

\[
K(\hat{R}_h, 1.7R_h) = \frac{M(\hat{R}_h, 1.7R_h)}{\rho_0(r_0)} \left( \hat{R}_h \sigma_{p,0}^2 \right)^{-1} .
\]

Now, requiring an approximate flatness for the projected velocity dispersion profile through eqn. (42) corresponds to selecting families of models whose coefficient \( K(\hat{R}_h, 1.7R_h) \) is almost constant in \( \hat{R}_h \), no matter what the precise dark matter density distribution is. This is easy to see in Figure 15.
which displays the coefficient \(K(R_h, 1.7R_h)\) as a function of \(R_h\) for different families of models. The black lines and the associated transparent area represent the families of models used in this paper. They satisfy eqn. (50) out to 6 half-light radii (\(\gamma = 6\)) and they have been plotted before in Fig. 2. The cored halos require a slightly higher value of \(K\), but the overall uncertainty on \(K\) is not larger than 20\%, even if the exact shape and scale length of the dark matter profile is unknown.

Red dashed lines in Fig. 10 display also analogous families of models which are characterized by tidally truncated dark matter density distributions (truncated in an exponential manner at 5 and 10 characteristic radii of the halo respectively, both for the NFW halo and for the cored halos). Note that, in this regard, tidal stripping does not represent a potentially dangerous unknown, as the uncertainty in \(K\) is not affected. By contrast, gray lines, and the associated gray shaded area represent families of models in which the requirement of eqn. (50) is much less restrictive and \(\gamma = 2\). In this case, the global uncertainty on \(K\) is larger than 50\%.

8 CONCLUSIONS

There are many examples of the modelling of the dwarf spheroidals (dSphs) using the Jeans equations in the recent literature. Typically, a photometric profile (King or exponential or Plummer) is combined with a dark halo density law and an assumption as to the behaviour of the anisotropy parameter. From this, simply requiring that the model can hold itself up against the dark halo gravity field via the Jeans parameter. From this, simply requiring that the model can

lead to results that are not generic, but merely a consequence of imposing the initial conditions. This is particularly the case regarding inferences concerning the behaviour of the dark matter density at the very centre (for example, whether it is cusped or cored).

The future should see more effort devoted to phase space modelling, in which a distribution of stellar orbits provides the observables, namely the density and the velocity dispersion. This rich field has so far been scarcely touched. A pioneering effort is that of Wilkinson et al. (2002), who modelled the Draco dSph with distribution functions. In these models, the scalelengths of the luminous and dark matter are the same, which restricts their widespread applicability. Wu’s (2007) impressive work constructs axisymmetric two and three integral models for three classical dwarf spheroidals, Draco, Ursa Minor and Fornax.

This paper has provided isotropic phase space models for all the Milky Way dSph galaxies. The observed flatness of the velocity dispersion profiles strongly suggests that the inner parts of the stellar populations are nearly isothermal, and so the families of lowered isothermal distribution functions made famous by Michie (1963) and King (1966) are natural starting points. For the dSphs, simple collisional relaxation has a physical timescale which is too slow to account for any central thermalization of velocities. However, a physical basis may be provided by the theory of tidal stirring (Mayer et al. 2001), in which dwarf irregular progenitors are transformed into dSphs by vigorous tidal shocking followed by bar and bending instabilities.

The distribution functions are all isotropic in velocity space. First, from the point of view of good scientific practice, the simplest assumption should be preferred until evidence to the contrary is found. In this respect, notice that the case of the dSphs is very different to that of elliptical galaxies, in which there is strong evidence for velocity anisotropy from the plot of observed flattening versus the ratio of ordered to random motions (Hillenbrand 1977; Binney 1978). Second, it is striking that most of the studies allowing anisotropy laws (Wu 2007; Lokas 2009) find fits suggesting that the dSph velocity distributions are nearly isotropic, especially in the central regions. Third, numerical simulations of tidal stirring produce dSphs that are typically isotropic (see Figure 23 of Mayer et al. 2001).

Clearly, though, the use of lowered isothermals is not the only possible way to construct models with a flattish velocity dispersion profile. For example, a tangentially biased in the structure of the orbits that increases with increasing radius could provide similar kinematic profiles. However, it seems unlikely that a formation scenario based on a collapse and subsequent tidal stirring can favour a tangentially biased velocity dispersion tensor in the outer parts.

When lowered isothermal distribution functions are embedded in dark haloes, then the stellar distribution relaxes in the gravity field of the dark matter. This leads to a prediction as to the half-light radius \(R_h\) as a function of the central velocity dispersion \(\sigma_{p,0}\). Modulo the overall scalings of the characteristic halo scalelength and the central dark matter potential, we have shown that this leads in practice to a one-parameter family of models. If the luminous length scale is much less than the halo scalelength, then the relationship between \(R_h\) and \(\sigma_{p,0}\) has a power-law form. In particular \(R_h \propto \sigma_{p,0}\) for a cored halo, and \(R_h \propto \sigma_{p,0}^2\) for an NFW halo. For systems whose luminous scalelength becomes comparable to the scalelength of the dark halo itself, significant deviations from a power-law correlation occur.

To match the observational data for any dSph, the existence of a theoretical curve between the dimensionless half-light radius and central velocity dispersion considerably restricts the set of suitable models. Note that solutions to the Jeans equations do not offer such a correlation – the set of physical solutions to the Jeans equations is overwhelmed by the infinitely more numerous unphysical solutions. In fact, in our approach, only one further ingredient is needed to fix the model, and this we choose as the halo scalelength (or equivalently the concentration of the dark halo). Once this is fixed, then there is only one, for example, NFW model that can place the observed datum for a dSphs onto the theoretical curve.

Even better, the mass within 1.7 times the half-light
radius $R_h$ is insensitive to the choice of halo scalelength. A larger halo scalelength enforces a compensating lower central dark matter density to ensure that the $(R_h, \sigma_{p,0})$ data-point lies on the theoretical curve, so that the mass interior to $1.7R_h$ remains largely unchanged. This result – perhaps more surprisingly – is also true when the halo model itself is altered, for example from cusped to cored. Accordingly, this enables us to provide reasonably reliable mass estimates for almost all the Milky Way dSphs out to $1.7R_h$. This result is valid for dSphs with flattish a projected velocity dispersion profiles.

The mass results do not support the conjecture of Strigari et al. (2008) that all dSphs have a common mass scale within 300 pc. Hercules and Leo IV have a mass interior to 300 pc of at most $3 \times 10^8 M_\odot$. It is probable that some of the more puny objects, like Segue 1 and Willman 1, also do not satisfy the putative common mass scale, but here the uncertainties of extrapolating the data on such physically small objects out to such a large distance as 300 pc does not allow us to be so definite.

The two most massive of the Milky Way dSphs are the most luminous, Sgr and Fornax. Within 1.7 half-light radii, we estimate that Sgr has a mass of $\sim 2.8 \times 10^8 M_\odot$ and Fornax $\sim 1.3 \times 10^8 M_\odot$. In particular, we do not reproduce the result of Peñarrubia et al. (2008) that physically smaller systems such as Draco and Sculptor are up to 5 times more massive than Fornax despite being roughly 70 times fainter. The least massive of the Milky Way satellites are Willman 1 ($\sim 4 \times 10^5 M_\odot$) and Segue 1 ($\sim 6 \times 10^5 M_\odot$). For 5 objects – namely Boo II, Leo IV, Leo V, Segue 2 and And XV – we are only able to provide an upper limit to the mass. It may well be that these objects are still more runty and unimpressive!

We also notice that the inferred values for the mass-to-light ratios for some of the most luminous dSphs – Fornax and Leo I particularly – could in fact mitigate against the validity of the standard approximation that the luminous component as a simple tracer. Actually, and especially in the case of a cored dark halo, it is quite likely that the stellar component itself has a non negligible effect on the kinematic structure. The point that for some of the bright dSphs, stars may contribute equally with the dark matter in the central regions has been noted before (Lokas et al. 2002; Strigari, Fr ank & White 2010).

The discovery of so many new dSphs over the last few years has thrown open a rich field for theorists. It is clear that our earlier ideas of common mass scales and universal haloes are giving way under the wealth of new data. The methods introduced in this paper constitute the first steps in providing the dSphs with phase-space models. Here, we have used the data on the half-light radius and central velocity dispersion to provide models for the whole sample. We plan to complement this with detailed models for some of the individual brighter dSphs in the near future, using the full photometric and kinematic profiles.

ACKNOWLEDGMENTS

We wish to thank Giuseppe Bertin for making interesting comments and suggestions, as well as the referee for a thorough reading of the paper. NA thanks STFC and the Isaac Newton Trust for financial support.

REFERENCES

Aaronson, M. 1983, ApJL, 266, L11
Adén, D., et al. 2009a, AA, 506, 1147
Adén, D., Wilkinson, M. I., Read, J. I., Feltzing, S., Koch, A., Gilmore, G. F., Grebel, E. K., & Lundstrø m, I. 2009b, ApJL, 706, L150
An, J. H., & Evans, N. W. 2009, ApJ, 701, 1500
Belokurov, V., et al. 2006, ApJL, 647, L11
Belokurov, V., et al. 2007, ApJ, 654, 897
Belokurov, V., et al. 2008, ApJL, 686, L83
Belokurov, V., et al. 2009, MN, 397, 1748
Belokurov, V., et al. 2010, ApJL, 712, L103
Binney, J. 1978, MN, 183, 501
Binney J., Tremaine S., 2008, Galactic Dynamics, Princeton University Press
Bullock, J. S., Kolatt, T. S., Sigad, Y., Somerville, R. S., Kravtsov, A. V., Klypin, A. A., Primack, J. R., & Dekel, A. 2001, MN, 321, 559
Chapman, S. C., Ibata, R., Lewis, G. F., Ferguson, A. M. N., Irwin, M., McCrackanie, A., & Tanvir, N. 2005, ApJL, 632, L87
Côté, P., Mateo, M., Olaszewski, E. W., & Cook, R. H. 1999, ApJ, 526, 147
Evans, N. W., An, J., & Walker, M. G. 2009, MN, 393, L50
Fraternali, F., Tolstoy, E., Irwin, M. J., & Cole, A. A. 2009, AA, 499, 121
Geha, M., Willman, B., Simon, J. D., Strigari, L. E., Kirby, E. N., Law, D. R., & Strader, J. 2009, ApJ, 692, 1464
Gilmore, G., Wilkinson, M. I., Wyse, R. F. G., Kleyna, J. T., Koch, A., Evans, N. W., & Grebel, E. K. 2007, ApJ, 663, 948
Hodgkin, S. T., Irwin, M. J., Hewett, P. C., & Warren, S. J. 2009, MN, 394, 675
Kendall, M., Stuart, A. 1979, London: Griffin, 4th ed., vol 2
King, I. R. 1962, AJ, 67, 471
King, I. R. 1966, AJ, 71, 64
Kleyna, J. T., Wilkinson, M. I., Evans, N. W., & Gilmore, G. 2001, ApJL, 563, L115
Kleyna, J., Wilkinson, M. I., Evans, N. W., Gilmore, G., & Frayn, C. 2002, MN, 330, 792
Kleyna, J. T., Wilkinson, M. I., Evans, N. W., & Gilmore, G. 2004, MN, 354, L66
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., et al. 2009, ApJL, 690, 453
Kuhlen, M., Strigari, L. E., Zentner, A. R., Bullock, J. S., & Primack, J. R. 2005, MN, 357, 387
Ibata, R. A., & Irwin, M. J. 1997, AJ, 113, 1865
Ibata, R., Martin, N. F., Irwin, M., Chapman, S., Ferguson, A. M. N., Lewis, G. F., & McConnachie, A. W. 2007, ApJ, 671, 1591
Illingworth, G. 1976, ApJ, 204, 73
Illingworth, G. 1977, ApJL, 218, L43
Irwin, M., & Hatzidimitriou, D. 1995, MN, 277, 1354
Irwin, M. J., et al. 2007, ApJL, 656, L13
Letarte, B., et al. 2009, MN, 400, 1472
Lewis, G. F., Ibata, R. A., Chapman, S. C., McCrackanie, A., Irwin, M. J., Tolstoy, E., & Tanvir, N. R. 2007, MN, 375, 1364
APPENDIX A: ASYMPTOTIC ANALYSIS OF THE PARAMETER SPACE

Here, we demonstrate some properties of the half-light radius $R_h$ in the two-dimensional parameter space $(\tilde{r}_t, \tilde{\sigma})$ by an analytic analysis in the asymptotic regimes. Note that we do not need to specify which dark matter density profiles [13]. As long as the associated gravitational potential is regular at large radii, the asymptotic behaviour of the $R_h(\tilde{r}_t, \tilde{\sigma})$ function is in fact the same.

Let start with the case $\tilde{\sigma}^2 \gg 1$, in which case the quantity $b$ defined in eqn (20) satisfies $b \ll 1$. The King density distribution (24) then behaves like

$$\rho_*(r) \sim \rho_{*,0} \left[ \frac{8b^5}{15\sqrt{\pi}} + O(b^7) \right]$$

$$\approx \rho_{*,0} \left[ \frac{8}{15\sqrt{\pi}} \left( \frac{\Phi(r) - \Phi(r_*)}{\sigma^2} \right)^2 \right]. \quad (A1)$$

Equation (A1) implies

$$\frac{\partial R_h}{\partial \tilde{\sigma}^2} = 0,$$  \quad (A2)

because the parameter $\tilde{\sigma}$ does not modify the profile of the stellar component $\rho_*(r)$, only its normalization. Thus, for any given value of the tidal radius $\tilde{r}_t$, if the velocity dispersion parameter is significantly higher than the central depth of the dark matter potential well, the half-light radius is constant.

Let us consider now the regime $\tilde{r}_t \gg 1$. For any value of $\tilde{\sigma}^2$, the tails of the luminous component are in the regime $b \ll 1$ described by the asymptotic eq (A1). The density profile has a divergent mass in the tails, because, as $\tilde{r}_t \rightarrow \infty$, the King distribution function reduces to the isothermal one. Let us indicate with

$$M_{\rho}(\tilde{r}_t) \sim \tilde{r}_t^3$$  \quad (A3)

the asymptotic behaviour of the projected total mass inside...
the tidal radius. A power-law behaviour for the enclosed projected mass as in eqn (A3) is associated to a power-law behaviour of the gravitational potential as the radius goes to infinity. In particular, \( \lambda = 1/2 \) for any regular potential for which \( \Phi \sim r^{-1} \) as \( r \to \infty \). We thus see that the half-light radius is linear in the tidal radius itself, because

\[
\hat{r}_t^\lambda \sim M_{\cdot \cdot \cdot, p}(\hat{r}_t) = 2M_{\cdot \cdot \cdot, p}(\hat{R}_h) \sim 2\hat{R}_h^\lambda.
\] (A4)

Furthermore, the ratio between the half-light radius and the tidal radius is fixed in the asymptotic regime (A3) as \( \hat{r}_t/\hat{R}_h = 2^{1/\lambda} \), implying that \( \hat{r}_t/\hat{R}_h \approx 4 \) for any regular dark matter potential.

Now, let us consider the limit of small tidal radii \( \hat{r}_t \ll 1 \), in other words the luminous component is in the very centre of the potential well. For any value of \( \hat{\sigma} \), the tidal radius can be made small enough so that the condition \( b \ll 1 \) is realized over the entire radial profile. Repeating the argument used in eq (A4), we see that the half-light radius is linear in the tidal radius and independent of the velocity dispersion parameter. The coefficient of this proportionality depends on the properties of the dark matter profile only, and amounts to \( \hat{r}_t/\hat{R}_h \approx 2.9 \) for an NFW profile and \( \hat{r}_t/\hat{R}_h \approx 2.5 \) for any cored profile.

Finally, the last is case \( \hat{\sigma} \ll 1 \). It is easy to see that, for any value of the tidal radius \( \hat{r}_t \), the velocity dispersion parameter \( \hat{\sigma} \) can be made small enough such that the ratio \( \hat{r}_t/\hat{R}_h \) is divergent.

Fig. A1 summarizes the properties of the models we recorded above by displaying the contours of the quantity \( \kappa = \hat{r}_t/\hat{R}_h \) in the \((\hat{r}_t, \hat{\sigma})\) plane. The parameter \( \kappa \) will provide a useful way to characterise the theoretical \( \hat{\sigma} - \hat{R}_h \) relations generated by these models. Note that the available parameter space is effectively divided into two regions: one for high values of \( \hat{\sigma} \) and/or of \( \hat{r}_t \), in which \( \kappa \leq 4 \) and one for small values of \( \hat{\sigma} \) in which \( \kappa > 4 \).