Orbit-based dynamical models of Local Group dwarf spheroidal galaxies

Breddels, Maarten

rijksuniversiteit groningen

Orbit-based dynamical models of Local Group dwarf spheroidal galaxies

Proefschrift

ter verkrijging van het doctoraat in de
Wiskunde en Natuurwetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. E. Sterken,
in het openbaar te verdedigen op
maandag 1 juli 2013
om 11.00 uur

door

Maarten Adriaan Breddels

geboren op 29 juli 1981
te Ooststellingwerf, Nederland

Promotor:	Prof. dr. A. Helmi
Beoordelingscommissie:	Prof. dr. P. T. de Zeeuw Prof. dr. O. Gerhard Prof. dr. E. Tolstoy

ISBN 978-90-367-6304-2
ISBN 978-90-367-6303-5 (electronic version)

Contents

Chapter 1

Introduction[1]

With absolute magnitudes ranging from $M_V \sim -9$ to ~ -13.5 and central surface brightness between $\mu_{0,V} \sim 22.5 - 27$ mag arcsec^{-2}, the "dwarf spheroidals" (dSph) are the faintest and lowest surface brightness galaxies known to date, beaten only by the relatively recently discovered ultra-faint dwarf galaxies (UFDs).

Although of dull appearance (see Fig. 1.1 for an example), dSph reveal an unexpectedly complex stellar populations mix (for a recent review see Tolstoy et al. 2009) what makes them very useful laboratories for understanding star formation and chemical enrichment processes at the faint end of the galaxy luminosity function. In terms of their internal dynamics, they might well be key in constraining the nature of dark matter. Even though the very first measurement of the line-of-sight velocity dispersion of a dSph was based on just 3 carbon stars in Draco (Aaronson 1983), it already hinted at a dynamical mass-to-light ratio about one order of magnitude larger than for globular clusters. Subsequent works have confirmed this result using larger samples that included red giant stars (e.g. Armandroff & Da Costa 1986; Aaronson & Olszewski 1987; Hargreaves et al. 1994a,b). If in dynamical equilibrium, dSph have the highest mass-to-light ratios known to date, with $M/L \sim 100\text{s}\ \mathrm{M}_\odot/\mathrm{L}_\odot$.

The primary focus of this Thesis is the development of orbit-based dynamical models for nearby dSph, with the goal of determining their mass content, dark matter density profile and internal orbital structure. In this Introduction, we put this work into context. In Sec. 1.1 we describe the latest observational surveys on the kinematics of dSph, place these systems in a cosmological setting and briefly discuss why most of these systems may be considered to be in dynamical equilibrium. In Sec. 1.2 we review the methods used to model the internal dynamics of spheroidal systems and their application to the dSph satellites of the Milky Way. Sec. 1.3 briefly summarizes our own findings as presented in the various Chapters of this Thesis. We end with a discussion on possible future developments in Sec. 1.4.

[1] Based on *"Internal kinematics and dynamical models of dwarf spheroidal galaxies around the Milky Way"* by Giuseppina Battaglia, Amina Helmi and Maarten Breddels, submitted to New Astronomy Reviews.

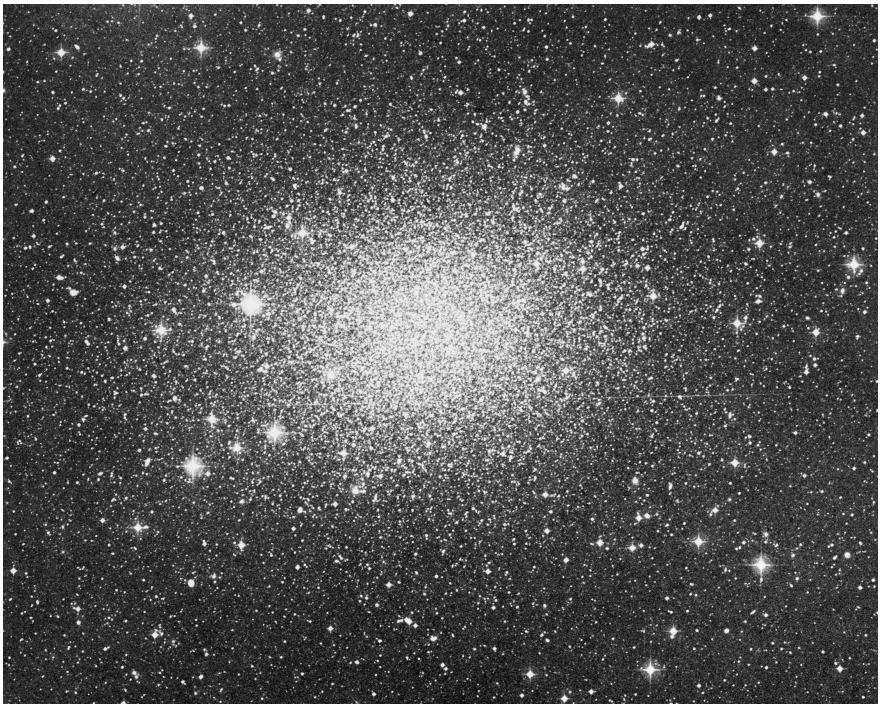


Figure 1.1: Optical image of the nearby ($\sim$ 80 kpc) dSph galaxy Sculptor, with its an almost featureless nearly spherical appearance. *Credit: David Malin, Anglo-Australian Observatory.*

1.1 Dwarf spheroidal galaxies

1.1.1 Surveys around the Milky Way

Determining the mass content of a system requires observations of the kinematics of suitable tracers. Since dSphs are devoid of a neutral interstellar medium, the only tracers available are stars. Because of their distance, to-date all measurements of their internal kinematics are based on line-of-sight velocities, as it is unfeasible to obtain accurate proper motions of individual stars in these galaxies with current facilities. The stars accessible for spectroscopic observations with current facilities are resolved for systems within the Local Group, since there is no crowding because of the low surface brightness of these galaxies. In this introduction, we concentrate on the dwarf galaxies satellites of the Milky Way (MW, hereafter. We refer the reader to Walker 2012, for a nice and comprehensive historical excursus on the growing kinematic samples for MW dSphs).

The first attempt to go beyond the determination of a global l.o.s. velocity dispersion of a dSph was made by Mateo et al. (1991) using a 2.5m telescope. These authors measured the kinematics of $\sim$ 30 stars in the Fornax dSph, in the center and in a field located at about two core radii. This first l.o.s. velocity dispersion "profile" turned out to be approximately flat, and this led the authors to suggest that it could be due to a dark halo spatially more extended than the visible matter. These results opened a whole line

of investigation to measure l.o.s. velocity dispersion profiles of dSph around the Milky Way, and to use these to determine their dark matter (hereafter, DM) distribution, orbital structure and dynamical state.

The samples of l.o.s. velocities collected in the 90s contained few dozens of individual members per dSph (e.g. Mateo et al. 1991; Hargreaves et al. 1994a,b; Olszewski et al. 1996; Queloz et al. 1995; Mateo et al. 1998). An increase in sample size became possible with multi-object spectrographs such as the KPNO/4m Hydra multi-fiber positioner (100 members in Draco and Ursa Minor, Armandroff et al. 1995), and the AF2/Wide Field Fibre Optical Spectrograph on the WHT (150 members in Draco, Kleyna et al. 2001).

In the second half of the 2000s several large spectroscopic surveys of the classical MW dSphs were carried out. In broad terms we can distinguish them in 3 main "streams":

1. Surveys that obtained l.o.s. velocities for typically ∼100-150 members per dSph, with a large success ratio of dSphs members/target stars thanks to an optimized target selection using Washington photometry (M, T2, and DDO51 filters, e.g. Majewski et al. 2005; Muñoz et al. 2005, 2006; Westfall et al. 2006; Sohn et al. 2007). These have made use of Keck/HIRES, Magellan/MIKE, CTIO/Hydra and Keck/DEIMOS.

2. Surveys to obtain several 100s of stars per dSph to determine both the internal kinematics and the metallicity distribution from Ca II triplet lines using intermediate resolution spectroscopy. This includes the Dwarf Abundances and Radial velocities Team (DART, PI: Tolstoy) (∼570, 800, 170 members for the Sculptor, Fornax, Sextans dSphs, respectively, at R∼6500 over the wavelength range 8200Å - 9400Å , Tolstoy et al. 2004, 2006; Battaglia et al. 2006; Helmi et al. 2006; Battaglia et al. 2008b; Starkenburg et al. 2010; Battaglia et al. 2011); and program 171.B-0520 (PI: Gilmore) "Towards the Temperature of Cold Dark Matter" (∼500, 170 members for the Carina and Leo II dSphs with the same set-up as for the DART data-set, Koch et al. 2006, 2007a). These have taken advantage of the VLT's large collecting area coupled to the wide-field, multi-object capability and stability of the FLAMES-GIRAFFE spectrograph (Pasquini et al. 2002) and, also of Keck/DEIMOS and GeminiN/GMOS (Koch et al. 2007c).

3. Surveys to obtain several 100s to 1000s l.o.s. velocities and spectral indexes (providing estimates of the *relative* metallicity of red giants) on a restricted wavelength range (5140 Å-5180Å) at resolution R∼20000 (PI: M.Mateo, e.g. ∼800, 2500, 1400, 400 members for Carina, Fornax, Sculptor and Sextans, respectively Walker et al. 2007a, 2009a). These have been mainly carried out with the Michigan/MIKE Fiber System (MMFS) at the Magellan/Clay (6.5m) telescope and with MMT/Hectochelle (see Mateo et al. 2008, for Leo I). With a comparable field-of-view to FLAMES (20 arcmin), MMFS has the advantage of almost double number of fibres (equally shared between the blue and red channel of the MIKE spectrograph).

Therefore, to-date the combined data-sets for the best studied dSphs have impressive sizes (∼2900 and 1700 probable members for Fornax and Sculptor, respectively), permitting studies of their internal properties to a level of detail that was unthinkable a little more than a decade ago. In particular, they have led to the characterization of

the line-of-sight velocity distributions (LOSVD) and its moments for the nearest dSph, crucial observables to constrain dynamical models.

While the global velocity dispersion tells us about the mass content of the system, the shape of the velocity dispersion profile provides information about the characteristics of the underlying potential and therefore the dark matter distribution of the system. However, it is also dependent on the internal orbital structure of the system (see also Sec. 1.2.1), and therefore just the second moment of a LOSVD is generally not sufficient to break this (mass- velocity anisotropy) degeneracy, and higher moments (especially the fourth) are necessary.

Instead of taking moments of the LOSVD, other possibilities are to express the LOSVD in, for instance, Gauss-Hermite moments (van der Marel et al. 1998) or analogous quantities (Amorisco & Evans 2012b). The (theoretically) best way would be to fit and model the full LOSVD (as in Chapter 4), since this incorporates all information and avoids spatial binning, which always leads to a loss of information.

1.1.2 DSph in a cosmological context

In our current understanding of the Universe, a mere 5% of the total mass/energy density budget consists of baryons, atoms essentially, with the remaining 95% comprising about 24% non-baryonic "dark matter" and 71% "dark energy" (see Hinshaw et al. 2012, for the 9-years WMAP results). This has become known as the Λ cold-dark matter (ΛCDM) model. As the evocative naming suggests, we ignore the nature of the great majority of constituents of the Universe.

There are several DM candidates such as weakly interacting massive particles, axions, sterile neutrinos, light gravitinos etc., whose existence is also motivated to solve problems in the Standard Model (for a review see Feng 2010). Some of these behave as cold and some as warm dark matter, where e.g. "cold" is defined as being non-relativistic at the time of structure formation. A wealth of experiments and strategies for direct and indirect detections of DM particles are underway (e.g. for reviews see Bertone et al. 2005; Hooper & Baltz 2008; Feng 2010), but at present the evidence for the existence of DM (based on the validity of Newton's law of gravity on all gravitational acceleration regimes) is provided by astrophysical observations on a variety of scales, from the smallest galaxies such as the dSph up to the largest structures in the Universe[2].

Potentially, astrophysical observations can provide important constraints on the dominant form of DM, as the characteristics of the DM particle are expected to influence the growth of structures, the substructure content and internal properties of DM halos. Rather than reviewing the extensive literature on the topic, we proceed to discuss results that are most directly related to this review, highlighting the crucial role of dwarf galaxies.

Cosmological pure DM N-body simulations, carried out in the ΛCDM framework, show that the halos formed follow very specific functional forms, such as the Navarro, Frenk & White profile (NFW, Navarro et al. 1996b, 1997)

$$\rho(r) = \frac{\rho_0}{r/r_s(1+r/r_s)^2} \tag{1.1}$$

[2] Alternative theories of gravity, or modifications of Newton's law have also been presented in the literature. We decided not to discuss these here because their application to model the dynamics of dSph has been very limited.

where ρ_0 and r_s are a characteristic density and radius. More recently the Einasto form has been found to provide better fits (e.g. Springel et al. 2008; Navarro et al. 2010)

$$\rho(r) = \rho_{-2} \exp\left\{-\frac{2}{\alpha_E}\left[\left(\frac{r}{r_{-2}}\right)^{\alpha_E} - 1\right]\right\}, \tag{1.2}$$

where ρ_{-2} and r_{-2} are the density and radius where the logarithmic slope $d\log\rho/d\log r = -\gamma_{DM} = -2$, and α_E is a shape parameter[3]. These density profiles are rather steep near the centre, with the NFW being cuspy with $\gamma_{DM} = 1$, while the Einasto profile has $\gamma_{DM} = 0$ at the centre.

Although not necessarily theoretically motivated, other density profiles are also often employed in the literature. Typically they have the form

$$\rho(r) = \frac{\rho_0}{(r/r_s)^\gamma (1 + (r/r_s)^\kappa)^{(\alpha-\gamma)/\kappa}}, \tag{1.3}$$

where $\alpha, \gamma, \kappa \geq 0$. Note that γ and α correspond to the inner and outer slopes respectively. The sharpness of the transition between these two regimes is thus given by κ. A cuspy profile has $\gamma > 0$, while for a cored one $\gamma = 0$ and $\kappa > 1$. This is because in the cored case, the profile must have a flat shape at the centre, i.e. $d\rho/dr = 0$. A profile that has $\gamma = 0$ and $\kappa \leq 1$ has at the centre $d\log\rho/d\log r = 0$ and a finite density, but in this case $d\rho/dr$ is non-zero, and hence this profile should not be confused with a core.

In the ΛCDM high-resolution cosmological N-body simulations described above the sub-halo mass function of MW-sized main halos is $dN/dM \propto M^{-\alpha}$, with $\alpha = 1.9$ down to the simulations resolution limit (Springel et al. 2008), which is smaller than the mass estimates for the faintest dSph (see Sec. 1.2). These simulations predict that MW halos contain 20% of the mass in subhalos, which results in a very large number of (mostly extremely low mass) satellites.

A comparison between the results of these pure DM N-body simulations with observations on galactic scales is not straightforward. Part of the issue lies in making the link between a luminous satellite to what should be its corresponding sub-halo in a DM simulation (e.g. of what mass? how dense?, see Strigari et al. 2010). This is particularly difficult because such simulations do not include baryons. This has motivated numerous theoretical efforts to provide a realistic treatment of baryonic effects using semi-analytical models and hydrodynamical simulations of dwarf galaxies (e.g. Revaz et al. 2009; Li et al. 2010; Font et al. 2011; Sawala et al. 2012; Starkenburg et al. 2013). Observationally, it is clearly important to obtain reliable estimates of the mass content and its distribution in dwarf galaxies.

For example, there is a debate about the inner shape of the density profiles of the DM halos hosting galaxies. For dSph, this issue is still very open (see Sec. 1.2 and Chapters 2 & 3). On the other hand, for isolated late-type dwarfs and low surface brightness galaxies, the rotation curves seem to favor cored rather than cusped DM distributions (e.g. de Blok 2010, and references therein). It has been suggested that feedback from supernovae explosions in these more massive systems could transform a cuspy halo into a cored one (e.g. Navarro et al. 1996a; Read & Gilmore 2005; Governato et al. 2010; Pontzen & Governato 2012; Teyssier et al. 2013). Note that in the case of an UFD, a single SN event releases an amount of energy comparable to the binding energy of the

[3] For $\alpha_E \sim 0.2$ the resulting profile resembles an NFW.

whole system. On the other hand, it is still to be assessed whether this mechanism is important or relevant on the scales of the MW dSphs, also given their low star formation rates (see Peñarrubia et al. 2012).

The "missing satellites" problem refers to the large mismatch between the observed number of dwarf galaxies satellites of the MW and M31 and the predicted number of DM subhalos (Klypin et al. 1999; Moore et al. 1999). The discovery of dozens low-luminosity dwarf galaxies in the Local Group, mainly by SDSS around the MW (e.g. Willman et al. 2005; Zucker et al. 2006; Belokurov et al. 2006, 2007, to mention a few) and the PandAS survey for M31 (e.g. McConnachie et al. 2009; Martin et al. 2009), has mitigated somewhat the "missing satellite" problem, after taking into account the surveys coverage and selection function (Koposov et al. 2009). The most appealing solution to reconcile predictions and observations is to suppress star formation, or gas accretion, in low-mass halos because of the joint effects of feedback and of a photo-ionizing background due to re-ionization (e.g. Bullock et al. 2000; Benson et al. 2002; Somerville 2002).

Another interesting issue was the recently reported "too big too fail problem" pointed out by Boylan-Kolchin et al. (2011), who used the Aquarius suite of DM simulations to argue that there exists a population of subhalos that are too massive and too dense to be consistent with the internal kinematics of the MW dSphs, and yet they do not have an observed stellar counter-part. However, as argued by Wang et al. (2012) and Vera-Ciro et al. (2012), the number of massive satellites is a stochastic quantity that also depends on the mass of the host. For example, if the mass of the MW is around $8{\times}10^{11}\,\mathrm{M}_{\odot}$, i.e. the least massive MW-like halos of the Aquarius suite (which reproduces well the observed MW satellite luminosity function, see Koposov et al. 2008; Starkenburg et al. 2013), the mismatch disappears. Furthermore, Vera-Ciro et al. (2012) show that M31, if assumed to be more massive than the Milky Way, does not miss such a population.

A plausible alternative to CDM is warm dark matter (WDM). The warm component has the effect of reducing the small-scale power in the primordial fluctuations spectrum, yielding fewer subhalos and of lower central densities (Colín et al. 2000, 2008; Lovell et al. 2012). Specifically, in the numerical simulations of Macciò et al. (2012, 2013), which explore a range of masses for the WDM particles, cored density profiles arise naturally. However, either the core sizes are too small to be consistent with those suggested in some studies of the internal kinematics of MW dSphs (see Sec. 1.2) or if large enough, they would be due to particles whose masses are inconsistent with the limits imposed by observations of the Lyman-α forest (e.g. Viel et al. 2005; Seljak et al. 2006; Viel et al. 2008). Note however that e.g. Busha et al. (2007) find in their WDM simulations that the halos are well described by an NFW form (i.e. cuspy) while Wang & White (2009) find this even holds for halos in hot dark matter simulations. Given that the state-of-the-art of WDM simulations is not as extensive and developed as for CDM, we await future developments.

From the above it is clear that there are numerous reasons to try and pin down the DM content and its distribution in the dSph. Given that the overall evolution of small systems like dwarf galaxies will most likely be sensitive to their relatively small potential well (e.g. Revaz & Jablonka 2012; Sawala et al. 2012), obtaining such measurements will also allow us to make sense of the variety of star formation and chemical enrichment histories of these galaxies, in particular in conjunction with the information on the dSphs orbital history that the Gaia satellite mission (Prusti 2011) will provide.

1.1.3 Are dSph in dynamical equilibrium?

A commonly made assumption in the dynamical modeling of dSph is that these objects are in dynamical equilibrium, while if they were significantly affected by tidal interactions with the MW this would need to be taken into account.

The possibility that dSphs are fully tidally disrupted dark-matter free galaxies has been excluded on the basis of their observed internal kinematic and structural properties (see for example Klessen et al. 2003; Muñoz et al. 2008), the large distances of some of these galaxies (up to 250 kpc from the MW) and a well-established luminosity-metallicity relation. It would also be difficult to explain the dSphs extended SFHs and broad metallicity distributions (see e.g. Tolstoy et al. 2004; Battaglia et al. 2006; Koch et al. 2006; Starkenburg et al. 2010; Battaglia et al. 2011) if the potential well would be due solely to the dSphs stars (amounting to typically 10^5-10^6 $M_\odot$, e.g. McConnachie 2012).

Partly because of the lack of knowledge of the orbits of dSph around the MW, the importance of tides on the stellar components of dSphs is largely unknown. This also depends on the degree of embedding of this component in its dark matter halo, as well as on the average density of the system. Mayer et al. (2001) propose that dSph galaxies are what results when a disky dwarf is tidally stirred by the MW. For this process to be effective, the stellar component of the dSph today has to be tidally limited, in which case tidal tails are expected. However, Peñarrubia et al. (2008b) find that the stars are very resilient to tides in their simulations where the stellar component follows a King-profile and is deeply embedded in an NFW halo. In any case, there is general consensus that the central velocity dispersion (or the dispersion at the half-light radius) continues being a good indicator of the present maximum circular velocity and bound mass, as long as the objects retain a bound core (e.g. Muñoz et al. 2008; Peñarrubia et al. 2008b; Klimentowski et al. 2009; Kazantzidis et al. 2011).

Besides the obvious case of Sagittarius, the only classical dSph presenting unambiguous signs of tidal disturbance such as tails and isophote twists is Carina (Battaglia et al. 2012). This object has been a candidate for tidal disturbance since a long time, with convincing arguments given by the presence of spectroscopically confirmed RGB stars probable members out to very large distances from its center (4.5 times the central King limiting radius), observed together with a break in the surface brightness profile, a velocity shear with turn-around, and a rising line-of-sight velocity dispersion profile (e.g. Muñoz et al. 2006). Among the classical dSphs, other candidates for tidal disruption are Leo I (e.g. Sohn et al. 2007; Mateo et al. 2008) and Ursa Minor (e.g. Martínez-Delgado et al. 2001; Palma et al. 2003; Muñoz et al. 2005), although the observational evidence is not as strong as for Carina. Note that even for Carina, the N-body simulations by Muñoz et al. (2008) show that large amounts of DM ($M/L \sim 40$ $M_\odot/L_\odot$) within the remaining bound core are still needed to explain its characteristics.

N-body simulations of tidally perturbed dSphs agree in predicting rising l.o.s. velocity dispersion profiles in the majority of cases, while only Carina and perhaps Draco (Walker 2012) are observed to show such feature. Together with the fact that most classical dSph show no tidal streams, this may taken as indicative that the outer parts of the stellar components of dSph have not been significantly affected by tides. All these arguments provide some justification for the assumptions made in this Thesis, namely that we may consider the dSph to be in dynamical equilibrium.

dwarf galaxy	Jeans		Schwarzschild		DF	M2M
	Spherical	Axis.	Spherical	Axis.		
Carina	G07,Wa07, L09,Wa09	HC12	Ch3			
Draco	L05,G07, Wa07,Wa09	HC12	J13		Kl01,Kl02	LM10
Fornax	L01,Wa07, L09,Wa09	HC12	Ch3	JG12†		
LeoI	Ko07b,G07, Wa07,M08, Wa09	HC12				
LeoII	Ko07a,G07, Wa07,Wa09					
Sculptor	Wa07,B08, L09,Wa09	HC12	Ch2,Ch3,Ch4		AmE12	
Sextans	G07,Wa07, Wa09,L09, B11	HC12	Ch3			
Ursa Minor	G07,Wa09					

Table 1.1: Overview of various modeling techniques applied to Local Group dSph galaxies. **AmE12:** Amorisco & Evans (2012a), **B08:** Battaglia et al. (2008a), **B11:** Battaglia et al. (2011), **Ch2:** Chapter 2 or Breddels et al. (2012), **Ch3:** Chapter 3 or Breddels & Helmi (2013), **Ch4**: Chapter 4, **G07:** Gilmore et al. (2007), **HC12:** Hayashi & Chiba (2012), **JG12:** Jardel & Gebhardt (2012), **J13:** Jardel et al. (2013), **Kl01:** Kleyna et al. (2001), **Kl02:** Kleyna et al. (2002), **Ko07a:** Koch et al. (2007b), **Ko07b:** Koch et al. (2007c), **L01:** Łokas (2001), **L01:** Łokas et al. (2005), **L09:** Łokas (2009), **LM10:** Long & Mao (2010), **M08:** Mateo et al. (2008), **Wa07:** Walker et al. (2007b), **Wa09:** Walker et al. (2009b), † Note that their dark matter halo is still spherical

1.2 Dynamical modeling

The techniques to model the internal dynamics of spheroidal systems have long been in place. However, their application to nearby dwarf spheroidals has only really taken off in the last decade, with the need for more sophisticated approaches thanks to the manifold increase in data samples. In this section we review the methods used, briefly discuss their limitations and the results obtained thus far for these systems. Table 1.1 gives an overview of the various modeling techniques applied to the MW dSph.

We divide this Section according to the groups of methods that have been used so far. In general, we can broadly classify methods on whether they are *parametric*, i.e. they assume a family of models, or *non-parametric*, in which the distribution function is expressed in more general terms, for example as an expansion of basis functions. Most works attempt to fit the moments of the velocity distributions, while the use of the velocities and positions of individual stars to determine the likelihood of a given model

(also known as discrete modeling) has been explored to a lesser extent in the literature.

1.2.1 Modeling with the Jeans Equations

To a very good approximation, a dwarf galaxy may be considered a collisionless system. The internal structure of such a dynamical system can be described through its distribution function $f(\mathbf{x}, \mathbf{v}, t)$, which in the collisionless case, obeys the Boltzmann equation

$$\frac{\partial f}{\partial t} + \mathbf{v}.\nabla_x f - \nabla_x \Phi.\nabla_v f = 0, \tag{1.4}$$

where $\Phi(\mathbf{x})$ is the total gravitational potential of the system (including stars and dark matter contributions Binney & Tremaine 2008). For our purposes $f(\mathbf{x}, \mathbf{v}, t)$ describes the probability of finding a star with a given position $\mathbf{x}$, and velocity $\mathbf{v}$ at time t.

In general we assume that the distribution function is time-independent (see Sec. 1.1.3), so that the first term in this equation may be dropped. Deriving the distribution function from Eq. (1.4) by comparison to observations is not straightforward (see below), so a commonly used approach is to take moments of this equation, and compare these moments to observables, since also low-order moments are easily measured from observations.

The zero-th moment corresponds to the continuity equation in hydrodynamics, and it is generally not used in data-model comparison. The first moment is obtained by multiplying Eq.(1.4) by v_j and integrating over all velocities. The resulting equation is

$$\frac{\partial \nu \langle v_i v_j \rangle}{\partial x_i} + \nu \frac{\partial \Phi}{\partial x_j} = 0, \tag{1.5}$$

where $\nu(\mathbf{x})$ is the stellar density, i.e. $\nu(\mathbf{x}) = \int d^3v f$, and the brackets $\langle\rangle$ denote moments, e.g. here $\langle v_i v_j \rangle = \int d^3v v_i v_j f$. Eq.(1.5) represents a set of 3 equations known as the Jeans equations. These are useful because they relate to observables, however, it should be born in mind that this is not a closed set of equations, in the sense that even if we knew the potential and the density, to derive the streaming (mean) velocities (3 components) and the full velocity ellipsoid (6 independent quantities), we only have 4 equations, i.e. the continuity and the Jeans equations. Although it is possible to use higher moments of the Boltzmann equation, this tends to be more cumbersome. Higher moments are also difficult to measure observationally reliably, and nonetheless the use of closure relations would still be necessary. Therefore, typically, as we shall see below, certain assumptions are made, regarding for example the form of the velocity ellipsoid, to find a solution to the system.

The distribution function of a steady state system depends on the integrals of motion. If the potential is time-independent, then the energy E is an integral of motion. For a spherical system, all components of the angular momentum $\mathbf{L}$ are conserved, while if the system is axisymmetric, then only L_z will be, but a third integral I_3 might exist. Therefore, in non-rotating spherical systems, the distribution function can be a function $f(E)$ or $f(E, L)$. Although it is possible for a spherical system to rotate (Lynden-Bell 1960), in which case the distribution function will be of the form $f(E, \mathbf{L})$ this is not the most general configuration. Rotation would be more natural in the axisymmetric case, when $f(E, L_z)$, i.e. there is a preferred axis (that about which the system rotates). As we will discus in Chapter 2, there is evidence of small velocity gradients in the dSph, however, their origin is unclear, and in many cases these can be explained by projection

effects. Therefore, in the rest of this Thesis we assume that our systems to do not rotate. In that case, the second moment and the variance of the velocity distribution are equal (after subtraction of the gradient), and we refer to these interchangeably.

Jeans equations for spherical systems

In the case of a spherical system, only one of the Jeans equations is non-trivially zero, and it relates the 2nd moment of the radial velocity $\langle v_r^2 \rangle$, the stellar density $\nu(r)$, the velocity anisotropy $\beta(r) = 1 - (\langle v_\theta^2 \rangle + \langle v_\phi^2 \rangle)/(2\langle v_r^2 \rangle)$, and the total gravitational potential $\Phi(r)$ as follows:

$$\frac{\mathrm{d}(\nu \langle v_r^2 \rangle)}{\mathrm{d}r} + 2\frac{\beta}{r}\nu \langle v_r^2 \rangle = -\nu \frac{\mathrm{d}\Phi}{\mathrm{d}r}. \tag{1.6}$$

An equivalent, often useful form of this equation is

$$\frac{GM(r)}{r} = \langle v_r^2 \rangle (\gamma_* - 2\beta - \alpha), \tag{1.7}$$

where r is the spherical radius, $\gamma_* = -d\log\nu/d\log r$ and $\alpha = d\log\langle v_r^2 \rangle / d\log r$. For example, if the radial velocity and stellar density have been measured, and we make an assumption on the velocity anisotropy β, we may be able to derive the mass distribution (gravitational potential) of the system. This is the most frequently used approach. The velocity ellipsoid can be isotropic, in which case $\beta = 0$, tangentially or radially anisotropic, when $\beta < 0$ or $\beta > 0$ respectively, and will in the most general case, vary with radius. In the case of $\beta = 0$, this implies that the velocity distribution is ergodic, i.e. it is only a function of energy $f = f(E)$, while for anisotropic systems, $f = f(E, L)$.

The above equations highlight a degeneracy between mass and anisotropy (if the stellar density is perfectly known from observations; otherwise this also enters the degeneracy). This is most easily seen if we assume that β is constant with radius. In that case, Equation (1.6) reduces to (Binney & Tremaine 2008)

$$\langle v_r^2(r) \rangle = \frac{1}{r^{2\beta}\nu(r)} \int_r^\infty \mathrm{d}r' r'^{2\beta} \nu(r') \frac{\mathrm{d}\Phi}{\mathrm{d}r'}. \tag{1.8}$$

We thus see directly that different combinations of the mass distribution, density and anisotropy might conspire to produce the same velocity dispersion profile in the radial direction. The situation is worsened by the fact that generally one deals with projected quantities, as discussed below.

A way to reduce the degeneracy is to use higher moments, in particular, the 4th moment equations are obtained by multiplying Eq. (1.4) by v_r^3 and $v_r v_t^2$ and integrating over velocity space (see Merrifield & Kent 1990):

$$\frac{\mathrm{d}(\nu \langle v_r^4 \rangle)}{\mathrm{d}r} - 3\frac{\nu}{r}\langle v_r^2 v_t^2 \rangle + \frac{2}{r}\nu \langle v_r^4 \rangle + 3\nu \langle v_r^2 \rangle \frac{\mathrm{d}\Phi}{\mathrm{d}r} = 0, \tag{1.9}$$

and

$$\frac{\mathrm{d}(\nu \langle v_r^2 v_t^2 \rangle)}{\mathrm{d}r} - \frac{\nu}{r}\langle v_t^4 \rangle + \frac{4}{r}\nu \langle v_r^2 v_t^2 \rangle + \nu \langle v_t^2 \rangle \frac{\mathrm{d}\Phi}{\mathrm{d}r} = 0. \tag{1.10}$$

If one assumes that the distribution function is of the form $f(E, L) = f_0(E) L^{-2\beta}$, it can be shown that the anisotropy is constant, and these equations simplify significantly to

$$\frac{\mathrm{d}(\nu \langle v_r^4 \rangle)}{\mathrm{d}r} + \frac{2\beta}{r}\nu \langle v_r^4 \rangle + 3\nu \langle v_r^2 \rangle \frac{\mathrm{d}\Phi}{\mathrm{d}r} = 0, \tag{1.11}$$

(Łokas 2002), whose solution may be expressed as

$$\langle v_r^4(r)\rangle = \frac{3}{r^{2\beta}\nu(r)}\int_r^\infty \mathrm{d}r' r'^{2\beta}\nu(r')\langle v_r^2(r')\rangle\frac{\mathrm{d}\Phi}{\mathrm{d}r'}. \tag{1.12}$$

As discussed above, the intrinsic moments are not directly accessible to the observer, and only projected moments of the line-of-sight velocity distribution and stellar density profile are measurable. Following Merrifield & Kent (1990) these projected moments take the form

$$\mu(R) = 2\int_R^\infty \nu(r)\frac{r\mathrm{d}r}{(r^2-R^2)^{1/2}}, \tag{1.13}$$

$$\langle v_{\rm los}^2(R)\rangle = \frac{2}{\mu}\int_R^\infty \nu(r)\left[\left(1-\frac{R^2}{r^2}\right)\langle v_r^2\rangle + \frac{1}{2}\frac{R^2}{r^2}\langle v_t^2\rangle\right]\frac{r\mathrm{d}r}{(r^2-R^2)^{1/2}}, \tag{1.14}$$

$$\langle v_{\rm los}^4(R)\rangle = \frac{2}{\mu}\int_R^\infty \nu(r)\left[\left(1-\frac{R^2}{r^2}\right)^2\langle v_r^4\rangle + 3\frac{R^2}{r^2}(r^2-R^2)\langle v_r^2 v_t^2\rangle + \frac{3}{8}\frac{R^4}{r^4}\langle v_t^4\rangle\right] \times\frac{r\mathrm{d}r}{(r^2-R^2)^{1/2}}. \tag{1.15}$$

Here R denotes the projected radial distance. Expressed in terms of the anisotropy β these equations take the form

$$\langle v_{\rm los}^2(R)\rangle = \frac{2}{\mu}\int_R^\infty \nu(r)\left(1-\beta\frac{R^2}{r^2}\right)\langle v_r^2\rangle\frac{r\mathrm{d}r}{(r^2-R^2)^{1/2}}, \tag{1.16}$$

$$\langle v_{\rm los}^4(R)\rangle = \frac{2}{\mu}\int_R^\infty \nu(r)\langle v_r^4\rangle g(r,R,\beta)\frac{r\mathrm{d}r}{(r^2-R^2)^{1/2}}, \tag{1.17}$$

where

$$g(r,R,\beta) = 1-2\beta\frac{R^2}{r^2}+\beta(1+\beta)/2\frac{R^4}{r^4}, \tag{1.18}$$

(Łokas & Mamon 2003). Eq. (1.17) is valid for the specific form of the distribution function that leads to a constant anisotropy, while Eq. (1.16) is more general.

In the recent past, Jeans modeling has been the most frequently used method to estimate the mass content of dSph (Łokas 2001; Kleyna et al. 2001; Koch et al. 2007b; Gilmore et al. 2007; Walker et al. 2007b; Battaglia et al. 2008a). For simplicity, many of the works assumed a constant anisotropy, and typically only the second moment is fit using the Jeans equation (although see below). The first modeling attempts already showed that mass following light models could not fit the relatively flat velocity dispersion profiles observed, and that extended dark matter halos were needed, for example in the case of Draco (Kleyna et al. 2001).

More recently, the focus has shifted to the type of dark matter halos that could host dSph. For example, Gilmore et al. (2007) assumed the velocity ellipsoid to be isotropic ($\beta = 0$), a cored light surface density distribution and a flat (inner) l.o.s. velocity dispersion profile, and found that dSph could be embedded in cored or cuspy dark matter halos (but shallower than the singular isothermal sphere). Walker et al. (2007b) assume NFW profiles and constant anisotropy together with an exponentially declining surface

brightness distribution. These authors fit the total mass $M_{\rm vir}$ and (constant) anisotropy β, and assume a particular value for the concentration from the virial mass-concentration relationship found in cosmological N-body simulations (e.g. Bullock et al. 2001; Macciò et al. 2007). It is important to stress that a quantity such as the total mass is not well constrained, but what is better constrained is the mass within a given radius (within the region spanned by the dataset). Such a quantity is less sensitive to the functional form of the density profile, and therefore preferable. Although the virial mass may be considered just another (free) parameter of the fit, its meaning as representing the total mass of the system is actually an extrapolation.

Walker et al. (2009b, 2010) have extended the modeling of their sample of dSph to allow for more general forms of the density profile of the dark matter ($\rho \propto 1/(x^{\gamma}(1+x^{\kappa})^{(3-\gamma)/\kappa})$, with $\gamma, \kappa \geq 0$, i.e. as in Eq. (1.3) with an outer slope $\alpha = 3$), while still assuming constant anisotropy. They use a Monte Carlo Markov Chain (MCMC) method to explore the space of parameters and find the best fit models. The results are shown in Fig. 1.2. An interesting finding is that they can strongly constrain the mass at the projected half-light radius $r_{\rm half}$ (the projected radius enclosing half of the total luminosity). Therefore, they also derive the circular velocity at $r_{\rm half}$, $V_{\rm half}$, and hence place lower limits on $V_{\rm max}$. This is only a lower limit because there is a degeneracy between $V_{\rm max}$ and the scale radius r_s of NFW profiles, as there are many such profiles consistent with a given measurement of $V_{\rm half}$ (Peñarrubia et al. 2008a)[4].

Łokas (2009) performed Jeans modeling of Car, Fnx, Sext and Scl using the 2nd and 4th moments of the l.o.s. velocity distribution. In this paper, they remove interloper stars iteratively, which leads to a velocity dispersion profiles that decrease with radius. They therefore find systematically lower masses than other authors, and that models in which mass follows light can in fact, fit the derived observables. However, the M/L derived are still much greater than expected from standard stellar populations (generally much greater than 10, see their Table 2). In the case of Draco, Łokas et al. (2005) found that the anisotropy was mildly tangential for a model with a r^{-1} density profile (with an exponential cut-off). This is consistent with the Jeans model by Walker et al. (2009b) and also with Jardel & Gebhardt (2012) orbital based Schwarzschild model of the system (within $\sim 1\sigma$, see below). Although the use of the 4th moment leads to a better constraint on the model parameters, its effect is relatively minor, and the solutions found are rather similar to those in which only the second moment is used. The reason for this may be attributed to the fact that the uncertainties on the measured kurtosis are large (samples are still small to measure moments very reliably), and that the differences with a Gaussian-like velocity distribution are not very big, in which case the 2nd moment suffices to characterize the LOSVD.

Other interesting results based on the Jeans equations

As discussed in previous sections, the use of the spherical Jeans equations requires assumptions on the functional form of the anisotropy and density profile of the system's dark matter halo. The solution to Eq. (1.14) gives us the parameters of the profile (a mass/density and scale radius, and an anisotropy). More generally, also the shape of the

[4] However, this degeneracy can be broken by measuring the velocity dispersion profile over a large extent in radius, as shown by Breddels & Helmi (2013) or Chapter 3. There is a second, more difficult to break degeneracy between the slope/functional form of the density profile and r_s.

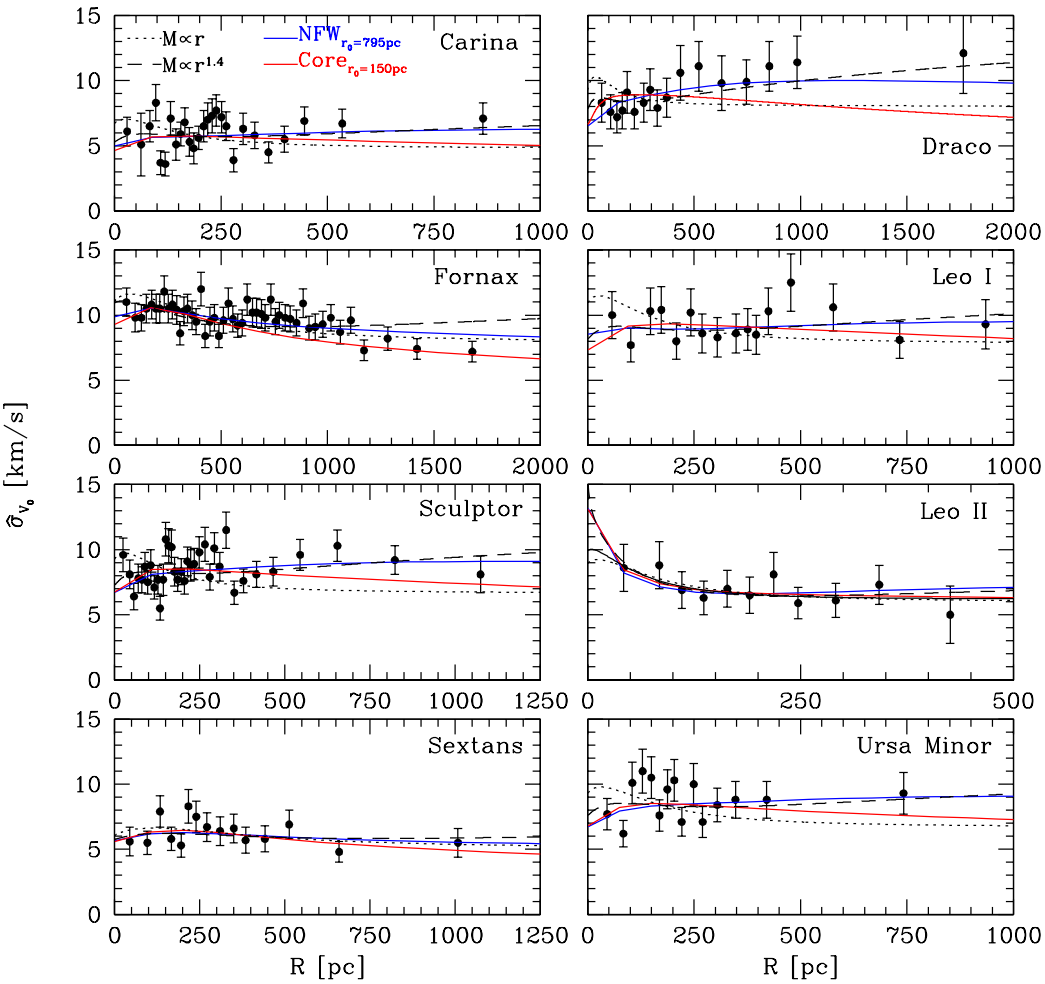


Figure 1.2: Projected velocity dispersion profiles for eight bright dSph obtained by Walker et al. (2009b). The profiles calculated from isothermal, power-law, NFW and cored halos (with $M(r) \propto r^{1.4}$ are also shown. These fits have been obtained using the spherical Jeans equations, and in particular Eqs. (1.8) and (1.16). For each type of halo these authors fit only for the anisotropy (assumed to be constant) and mass normalization (at a given scale/distance). The scale radius of the system is fixed for the various models as indicated in the panels.

density profile may be allowed to vary, as in Strigari et al. (2007b); Walker et al. (2009b).

A common mass scale? In two thorough studies, Strigari et al. (2007b, 2008) used MCMC numerical methods to explore a large range of models for the dark halos hosting dSph. These authors found that of all the parameters describing the model, the mass within 300 pc was a robust and well determined quantity, that was roughly independent of anisotropy or shape parameters. In an immediate application of this result, Strigari et al. (2008) found that most dSph must be embedded in dark matter halos of similar mass within this radius $M_{300} \sim 1 - 2 \times 10^7 M_\odot$, despite the fact that they span several orders of magnitude in luminosity. Although these results have been refined, especially for the ultrafaint dSph, which do not even extend up to 300 pc (and hence this M_{300} is an extrapolation), in general the M_{300} is confirmed to be a very weak function of luminosity $M_{300} \propto L^{0.03\pm0.03}$ (Rashkov et al. 2012).

Robust measurement of $M(r_{1/2})$ The virial theorem as well as the Jeans equations actually offer a plausible explanation for why the mass at a particular radius may be estimated reliably from the LOSVD only. The virial theorem tells us that $M_{\rm tot} = \sigma^2_{\rm tot} r_g / G$ where r_g is the gravitational radius of the system (see Eq. (2.42) of Binney & Tremaine 2008). On the other hand, Wolf et al. (2010) have shown analytically using the Jeans equation, that at the radius at which the slope of the stellar density profile $d\log\nu/d\log r = -\gamma_* = -3$, the mass is very well constrained independently of the anisotropy of the system. Thus

$$M(r_{-3}) = 3\frac{\langle\sigma^2_{\rm los}\rangle r_{-3}}{G} \tag{1.19}$$

for a system with a flat velocity dispersion profile. Since most of the dSph have such relatively flat profiles, Wolf et al. (2010) and also Walker et al. (2009b, 2010) in their MCMC analysis of the Jeans equation have been able to confirm this analytic result. In general, however, instead of estimating the radius r_{-3}, Wolf et al. (2010) use the half-light radius $r_{1/2}$ since the two are very similar for most profiles used to model the light distribution in dSph. Note that this is the 3D radius containing half of the total luminosity of the system, and not the effective radius obtained from the surface brightness profiles nor the 2D projected radius containing half of the luminosity, $r_{\rm half}$ in Walker et al. (2010).

These relations are also useful for ultrafaint dSph, provided these systems are in dynamical equilibrium. The sample sizes for most of these systems are too sparse to warrant a full dynamical model so general scalings as those just described may be more useful. See also An & Evans (2011) for more information on the theory of virial mass estimators.

General constraints on the df Not every solution to the Jeans equation has an associated distribution function that is physical, i.e. positive everywhere. This is why it is important to find additional conditions that can help identify when the assumptions made to solve the Jeans Equations will lead to plausible (physical) solutions.

An & Evans (2006, 2009) and Evans et al. (2009) use the Jeans equations to explore the asymptotic relations between the anisotropy β, the logarithmic slope of the light distribution γ_* and that of the underlying dark matter density profile near the center

of a spherical system γ_{DM}. They show that, if the tracer population is embedded in a spherical dark halo that is shallower than the singular isothermal sphere ($\gamma_{DM} < 2$) in the center, a finite central velocity dispersion $\sigma_{r,0}$ implies a relation between the central value of the logarithmic slope of the tracers $\gamma_{*,0}$ and the velocity anisotropy at the center β_0, namely $\gamma_{*,0} = 2\beta_0$. However, it is also possible that the system is dynamically cold at the center (i.e. $\sigma_{r,0} = 0$), in which case the condition is $\gamma_{*,0} > 2\beta_0$. This theorem highlights that care is required in the interpretation of results based on assumptions such as isotropy and spherical symmetry.

Ciotti & Morganti (2010) showed that there may be a more general relation that should hold at all radii, which is that $\gamma_* \geq 2\beta$. This may be seen to be related to the positivity of the mass (Eq. 1.7), as $M \propto \gamma_* - 2\beta - \alpha \geq 0$, where $\alpha = d\log\sigma_r^2/d\log r$. Ciotti & Morganti (2010) have demonstrated the above relation holds for particular forms of the distribution function (namely those in which the augmented density is a separable function of radius and potential, see their Eqs.(1 - 6) for more details), but the more general inequality (including α) should be always true.

However, it should also be born in mind that this analysis applies to intrinsic quantities implying for example, that even if $\sigma_{r,0} = 0$, $\sigma_{\rm los}$ can still be finite at the center, and hence the theorem is, although correct, less powerful in predicting the orbital behavior at the center, as we demonstrate in Chapter 5. Furthermore, a surface brightness profile might have a very shallow cusp (as considered in Strigari et al. 2010), in which case the velocity ellipsoid need not be isotropic.

1.2.2 Modeling through distribution functions

As stated previously, a solution to the Jeans equation is not necessarily physical since there is no guarantee that a distribution function will exist that is positive definite everywhere. This is one of the reasons why several authors have attempted to model directly the distribution function itself.

Dejonghe & Merrit (1992) have studied the issue of how the projected velocity distribution as a function of position $f_{\rm los}(v_{\rm los}, r_{\rm los})$ for a spherical system constrains the distribution function and gravitational potential. They show that if the form of the spherical potential is specified, then $f(E, L)$ is uniquely determined by $f_{\rm los}(v_{\rm los}, r_{\rm los})$. However, if the spherical potential is not known, they argue that there will be a family of possible potentials, but only those that lead to a df that is positive everywhere would be allowed, and not every potential will permit that.

Merritt & Saha (1993) explore the problem of inferring the gravitational potential of a spherical system from measurements of the l.o.s. velocity and positions for individual stars (or galaxies, in their case). They assume that the distribution function may be expressed as a polynomial expansion: $f(E, L) = \sum_{m,n} c_{m,n} f_{m,n}$ where $f_{m,n} = (-E)^{n-1/2} L^{2m}$, hence this distribution function is non-parametric. To determine the properties of the potential, however, a few parametric forms are considered. Thus from a discrete set of velocities of galaxies in the Coma cluster, they find best solutions in a maximum likelihood sense. These authors estimate that meaningful constraints are possible with datasets containing ~ 1000 objects.

Wu & Tremaine (2006) (see also Merritt 1993) take an even more general form for the distribution function, namely they divide the (E, L) space into $N_E \times N_L$ bins, and construct a set of top-hat basis functions, $h_{mn} = 1/V_{mn}$ where $V_{mn} = \int_{mn} d^3x d^3v$ is the

phase-space volumen associated to bin mn. Thus, $f(E,L) = \sum_{mn} w_{mn} h_{mn}$, and the task consists in finding the weights w_{mn} that fit the observables after assuming a specific gravitational potential. Wu & Tremaine (2006) use this technique to infer the mass of M87 from the motions of its globular clusters. These approaches are very powerful as they use maximally the datasets, without turning to moments to characterize the LOSVDs, and are also free of assumptions regarding the distribution function. In Chapter 4, we fit similar models where instead of the top-hat functions we use orbits as building blocks.

Wilkinson et al. (2002) introduce a family of anisotropic distribution functions for spherical systems, in which the dominant gravitational potential is cored and parametrized as $v_c^2 = v_0^2 r^2/(1+r^2)^{1+\delta/2}$. For different values of the characteristic parameters ($-2 \geq \delta \geq 1$), this leads to flat or declining rotation curves. The velocity ellipsoid is isotropic in the center, and may become radially or tangentially anisotropic at intermediate radii, while at large distances it is constant. The advantage of this family of distribution functions is that the expressions for the various moments (including the 2nd and 4th) are analytic, and depend only on the parameters of the distribution function. This means, that in principle, these characteristic parameters could be retrieved directly through comparison to observations. They also compute the projected (observable) quantities for different values of the parameters. The resulting l.o.s. velocity dispersion profiles (see their Fig. 3) can be flat, rising or declining depending on the distribution function. In Kleyna et al. (2002) they have applied this modeling to a dataset for Draco with $\sim$ 160 member stars, and found that the system is best fit by a slightly tangentially anisotropic ellipsoid and with a halo that falls off more slowly than a flat rotation curve model ($v_c \propto r^{0.17}$), while they are also able to rule out a mass-follows-light model and an extended harmonic core with 3σ confidence.

1.2.3 Made to measure

The Made-to-Measure (M2M) is a numerical method that integrates the orbits of test particles in a gravitational potential in order to reproduce a given set of observables (Syer & Tremaine 1996). Particles have associated weights, which themselves follow equations of motion. The system is evolved in time until a satisfactory solution has been found. The gravitational potential may be specified or determined self-consistently, and the resulting distribution function is completely non-parametric, and determined by the final particle's configuration that satisfies the observational constraints. The method can be used to model individual measurements or moments of a LOSVD (as with N-MAGIC in de Lorenzi et al. 2007).

Long & Mao (2010) have modeled Draco using the data from Kleyna et al. (2002) and assumed an isotropic velocity ellipsoid, and the same type of cored potentials as Wilkinson et al. (2002). The best fit model has asymptotic slope for the squared circular velocity v_c^2 of $\delta = -0.90^{+0.36}_{-0.35}$, while for the mass within three core radii they find $9.7 \pm 2.3 \times 10^7 M_\odot$, in comparison to Kleyna et al. (2002) who obtain $\delta \sim -0.34$ and a somewhat smaller mass. Long & Mao (2010) attribute this difference to their assumption of isotropy.

1.2.4 Schwarzschild modeling

This is the method of choice in this Thesis, and it is discussed and used extensively in Chapters 2, 3 and 4. Schwarzschild modeling is by now a traditional technique to derive

the mass distribution, especially in elliptical galaxies, from integrated light spectroscopy. It was initially developed in the 1980s (Schwarzschild 1979; Richstone & Tremaine 1984), and used extensively to derive M/L and black hole masses in the 1990s and the 2000s (e.g. Rix et al. 1997; van der Marel et al. 1998) where it was extended to allow for axisymmetric models, and even triaxiality (van den Bosch et al. 2008). The basic idea of the method is that the building blocks of galaxies are orbits, and through the right orbital superposition it is possible to match the light and kinematic distributions observed.

Therefore the method consists in assuming a specific gravitational potential, calculating the observables predicted for each orbit, and then weighting the orbits (with non-negative weights) to obtain a model that fits the observed data in a χ^2 sense. This approach guarantees that the distribution function obtained (which is reflected in the orbit weights) is non- negative. The fitting procedure thus allows the determination of the characteristic parameters of the best fit model for a specific gravitational potential. If one wishes to test different functional forms for the gravitational potential, then new orbit libraries need to be built, and the fitting procedure is repeated. The advantage of this method is that it does not make assumptions about the form of the anisotropy or the distribution function (rather these are an outcome of the model), and therefore it is less biased than some of the modeling techniques described above. Naturally, it is less flexible in the sense that it is more computationally intensive/expensive, and hence it is possible to explore a smaller variety of gravitational potentials, than for example, through Jeans models.

Despite the vast history, this method has not been applied systematically to the dynamical modeling of dwarf galaxies until very recently. Besides the work described in this Thesis, Jardel & Gebhardt (2012) have presented three-integral, Schwarzschild models of Fnx that take into account the non-spherical light distribution of this galaxy, although embedded in a spherical dark matter halo. These authors have tested a cored profile $\rho \propto (3r_c^2 + r^2)/(r_c^2 + r^2)^2$ and the NFW model. They find that the cored model is strongly favored, and that the velocity ellipsoid is mildly radially anisotropic. Their mass for Fnx $M(R_e) = 3.9^{+0.46}_{-0.11} \times 10^7 M_\odot$ is somewhat smaller than what the estimators by Wolf et al. (2010) or Walker et al. (2009b, 2010) would predict. Jardel & Gebhardt (2012) argue that this might be related to the fact that those estimators have been established (and shown to be independent of anisotropy) for spherical models. Another difference might lie in that the amplitude of the line-of-sight velocity dispersion profile they derive for Fnx is somewhat lower than that derived, for example, in Chapter 3 of this Thesis.

Jardel et al. (2013), return to spherical models, but assume that the density profile for the dark matter is non-parametric. They model Draco in this way, and find that the preferred model is a power-law, with a slope quite similar to the NFW, that is $\gamma_{DM} = 1$ for $20 \leq r \leq 700$ pc, and that the velocity ellipsoid is radial. Note that, in comparison to Wilkinson et al. (2002), Jardel et al. (2013) have allowed greater freedom in the form of the density profile (and have not forced cored models), and hence their results are potentially more robust.

1.2.5 Modeling dSph with composite stellar components

Several dSph host multiple stellar chemo-dynamical components. Since these components are embedded in the same gravitational potential, they allow to place more stringent constraints on the properties of this potential, since e.g. each of the component has to

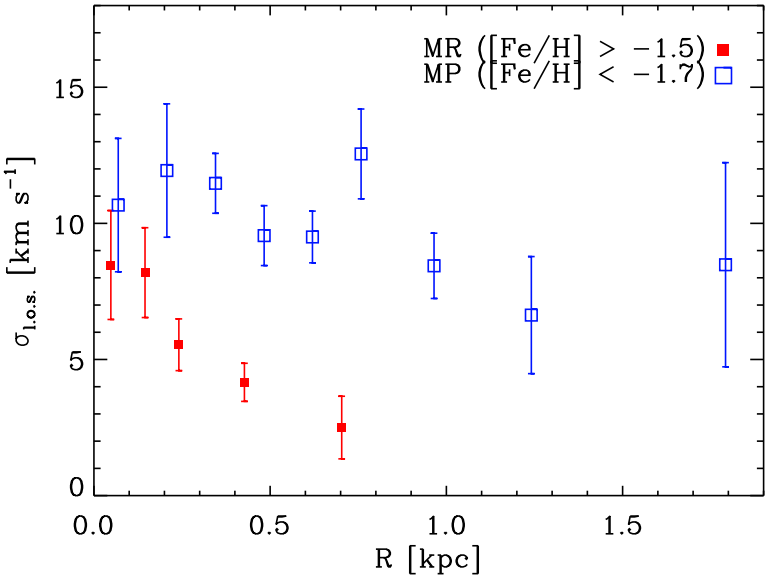


Figure 1.3: L.o.s. velocity dispersion profiles for the Sculptor members more metal-rich than [Fe/H]= −1.5 (filled squares) and more metal-poor than [Fe/H]= −1.7 (open squares), from rotation-subtracted velocities in the Galactocentric Standard of Rest system (see Battaglia et al. 2008a).

satisfy the Jeans equations independently. In practice this means that there are fewer free parameters since each component will follow its own distribution function entering the left-hand-side of Eq. (1.6), but the right-hand-side will be the same, thereby effectively leading to a reduction in the number of degrees of freedom.

This idea was first exploited by Battaglia et al. (2008a), who modeled Scl using two

components, a metal-rich centrally concentrated, and a metal-poor hot and extended, both embedded in a dark matter halo. These authors found, using Jeans models, that the metal-poor component was better fit with a nearly flat anisotropy profile, while the metal-rich one, because of its rapidly falling velocity dispersion profile (see Fig. 1.3), required a radially anisotropic ellipsoid. They found that cored models provided better fits but that NFW models could not be ruled out. In Chapter 4 we see that these two (likely the same) components will appear in the distribution function without using the chemical information.

Walker & Peñarrubia (2011) put forward another use of the composite stellar components, namely to infer the slope of the dark matter profile. These authors argue that one might consider Eq. (1.19) for each component separately, so that the mass of the host halo is constrained at the half-mass radius of each component independently. This then leads to two measurements of the mass at two different radii, and hence to a slope. They have performed many tests of their method, whose basic assumption is that the l.o.s. velocity profiles are flat, and found that their results are relatively robust to such (and other) assumptions, although systematic uncertainties affect the masses at $r_{\rm half}$ which depend on the density profile of the halo and the degree of embedding of the stars). These authors define

$$\Gamma = \Delta \log M / \Delta \log r = \frac{\log(M_{h,2}/M_{h,1})}{\log(r_{h,2}/r_{h,1})} \sim 1 + \frac{\log(\sigma_2^2/\sigma_1^2)}{\log(r_{h,2}/r_{h,1})}, \tag{1.20}$$

where $r_{h,\rm pop}$ and $M_{h,\rm pop}$ refer to the projected half-light radius and the mass at this point, while $\sigma^2_{\rm pop}$ is the global velocity dispersion that characterizes the population, and where pop $= 1, 2$, i.e. metal-rich or metal-poor components. In the limit of $r \to 0$, then $d \log M / d \log r = 3 - \gamma_{DM}$ where γ_{DM} is the central value of the slope of the dark matter density profile. Since $d \log M / d \log r$ decreases as r increases for any reasonable density profile, this implies that $3 - \gamma_{DM} > \Gamma$, or alternatively that $\gamma_{DM} < 3 - \Gamma$, as the slope Γ is measured at a finite distance from the center. Walker & Peñarrubia (2011) find $\Gamma = 2.61^{+0.43}_{-0.37}$ for Fnx, while for Scl $\Gamma = 2.95^{+0.51}_{-0.39}$. This thus implies that NFW-like profiles ($\gamma_{DM} = 1$) would be ruled out at significance levels $\gtrsim 96\%$ and $\gtrsim 99\%$ respectively for these systems. These results are much more stringent than any of the previously reported findings by other authors, where typically both profiles are consistent with the data.

More recently, Amorisco & Evans (2012a) have modeled the two populations in Scl using Michie-King models. These are isotropic in the center and become radially anisotropic in the outskirts. The validity of these assumptions for the velocity ellipsoid is taken from their analysis of the shape of the l.o.s. velocity distributions of Scl in Amorisco & Evans (2012b), whose estimates of the 4th moment would suggest a radially anisotropic ellipsoid (see however Chapter 2, where we find a kurtosis profile that is consistent with tangential anisotropy). Under these assumptions for the velocity ellipsoid (or the distribution function) these authors find that cored mass distributions are preferred over cusped ones. Given the uncertainties in the measurements of the 4th moments, this result could be related to the assumed shape of $\beta(r)$ rather than necessarily reflect the underlying mass distribution.

Agnello & Evans (2012) use the projected virial theorem and argue that the two

populations in Scl should satisfy independently the virial theorem,

$$\frac{K_{los,1}}{K_{los,2}} = \frac{W_{los,1}}{W_{los,2}} \tag{1.21}$$

from which the obtain the relation

$$\left(\frac{\sigma_{0,1}}{\sigma_{0,2}}\right)^2 > 2\left(\frac{R_{h,1}}{R_{h,2}}\right), \tag{1.22}$$

if the stars follow Plummer profiles and are embedded in NFW halos. Given their estimates of these various observables, Agnello & Evans (2012) conclude that no NFW halo can be compatible with the energetics of the two populations. Because the two populations should co-exist in virial equilibrium, the authors argue that this implies that the dark halo must be cored, and they estimate its size to be ~ 120 pc.

The results presented in this section all argue that the modeling using two (or multiple) components disfavor NFW/cuspy profiles for dSph, at least for Fnx and Scl. It is striking that the consideration of two components in dynamical equilibrium point all in the same direction. It would be important to confirm these results using fully-fleshed non-parametric methods, such as Schwarzschild or Made-to-Measure, that explore the presence of multiple populations and remove some of the (systematic) uncertainties in the use of global scaling relations. It would also be desirable to understand the extent down to which these systems' properties are better described using a few independent components, rather than to assume that the properties of the stars change gradually throughout the system, and specifically how these assumptions affect the dynamical models and their conclusions.

Figure 1.4 compares the results of various modeling approaches on Scl. In this figure we have plotted the mass distribution derived using Schwarzschild models from Chapter 3 (or Breddels & Helmi 2013). We have included here estimates of M_{300} by Strigari et al. (2008) and Walker et al. (2009b) and at the 3D half-light radius by Wolf et al. (2010). These estimates are all consistent with those obtained of Chapter 3 which is reassuring. The mass estimated at 1.8 kpc obtained by Battaglia et al. (2008a) assuming a cored density profile but modeling simultaneously metal-rich and metal-poor populations is also shown (open black circle). It is on the upper side of the curves, but is consistent within error bars, and is beyond the region where the mass profiles are indistinguishable, so this mass estimate is likely to be more model dependent. Finally, the two estimates of the mass derived by Walker & Peñarrubia (2011) are shown as diamonds in this Figure. These two estimates of the mass at the projected half-light radius of the metal-rich and metal-poor components of Scl, appear to be somewhat larger than what is found in Chapter 3. This is consistent with the systematic uncertainties that Walker & Peñarrubia (2011) reported from their Monte Carlo simulations. However, we notice that the mass at the half-light radius of the metal-poor component is more overestimated than that of the metal-rich one (and even higher than M_{300} or $M_{1/2}$ for example). In view of this, it seems plausible that the slope Γ that Walker & Peñarrubia (2011) derived could be overestimated, in which case, cuspy profiles with $\gamma_{DM} > 0$ could still be allowed.

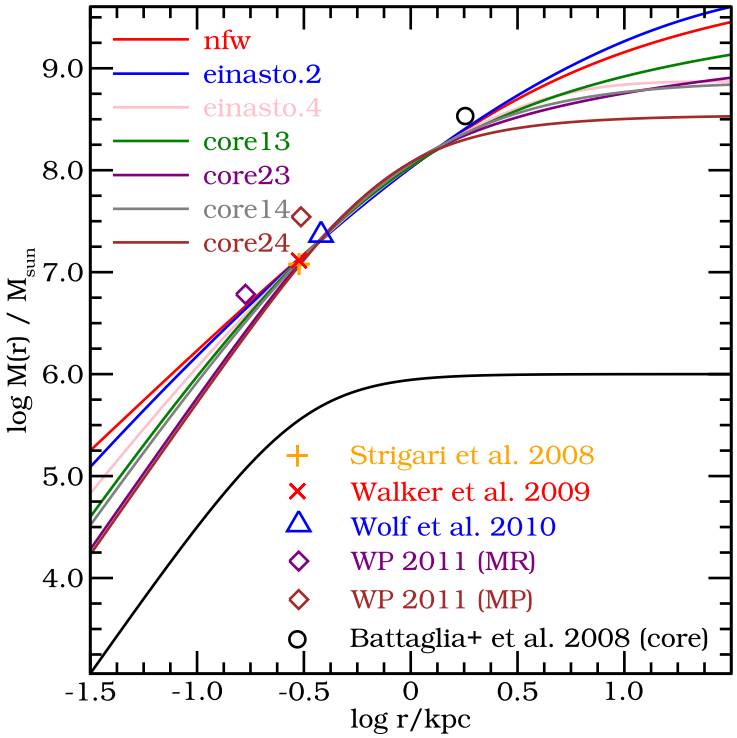


Figure 1.4: Mass distribution $M(r)$ for Scl derived in Chapter 3. The various curves represent the best fit models obtained using Schwarzschild's orbit-based method. As in the case of Fnx, the various density profiles appear indistinguishable from one another. We have overlaid various measurements for the mass at different radii, including those obtained considering separately the metal-rich and metal-poor populations by Battaglia et al. (2008a), and by estimating separately $M_{1/2}$ for the metal-rich (MR) and metal-poor (MP) component in Walker & Peñarrubia (2011).

1.3 This Thesis

In this thesis we focus on applying Schwarzschild modelling to fit the internal kinematics of Local Group dSph galaxies.

In Chapter 2 we have developed spherically symmetric dynamical models of dwarf spheroidal galaxies using Schwarzschild's orbit superposition method. We present our results for the Sculptor dwarf spheroidal galaxy, after testing our methods on mock data sets. We fit both the second and fourth velocity moment profile to break the mass-anisotropy degeneracy. For an NFW dark matter halo profile, we find that the mass of Sculptor within 1 kpc is $M_{1kpc} = (1.03 \pm 0.07) \times 10^8$ $M_\odot$, and that its velocity anisotropy profile is tangentially biased and nearly constant with radius. The preferred concentration ($c \sim 15$) is low for its dark matter mass but consistent within the scatter found in N-body cosmological simulations. When we let the value of the central logarithmic slope α vary, we find that the best-fit model has $\alpha = 0$, although an NFW cusp or shallower

is consistent at 1σ confidence level. On the other hand, very cuspy density profiles with logarithmic central slopes $\alpha < -1.5$ are strongly disfavoured for Sculptor.

In Chapter 3, we use the same models to fit both the 2nd and 4th moments of the line of sight velocity distributions of the Fornax, Sculptor, Carina and Sextans dwarf spheroidal galaxies. Our goal in this Chapter is to compare dark matter profile models of these four systems using Bayesian evidence. We consider NFW, Einasto and several cored profiles for their dark halos and present the probability distribution functions of the model parameters. When considering each system separately, we find there is no preference for one of these specific parametric density profiles. However, the combined evidence shows that is unlikely that all galaxies are embedded in the same type of cored profiles of the form $\rho_{DM} \propto 1/(1+r^2)^{\beta/2}$, where $\beta = 3, 4$. For each galaxy, we also obtain an almost model independent, and therefore accurate, constraint on the logarithmic slope of the dark matter density distribution at a radius $\sim r_{-3}$, i.e. where the logarithmic slope of the stellar density profile is -3. This is because all our best fit models essentially have the same mass distribution over a large range in radius (from just below r_{-3} to the last measured data point). This remarkable finding likely implies much stronger constraints on the characteristics that subhalos extracted from cosmological simulations should have in order to host the dSph galaxies around the Milky Way.

In Chapter 4 we fit the full line of sight velocity distribution of Sculptor, with again orbit-based dynamical models. Unlike previous work based the moments of the line of sight velocity distribution, we use the individual measured velocities. When we test our method on a Mock dataset, we find that we can recover well the distribution function and the parameters of the underlying potential. To estimate the effects of sampling, we test the method both using the expectation value of the log likelihood of the data and on a discrete dataset of 2,000 stars, i.e. the size available for Sculptor, as well as one that is $5\times$ larger. This enables us to show we accurately estimate the parameters of the potential and without bias with 6% (3%) uncertainty in the mass for a kinematic dataset of 2,000 (10,000) stars. When we apply our method to Sculptor, assuming a Navarro, Frenk & White profile, we find a mass of $M_{1kpc} = 10^{8.05\pm0.024} M_\odot$ and $r_s = 5.1^{+2.6}_{-1.7}$ kpc which are consistent, though slightly larger, than the estimates presented in Chapters 2 and 3. Interestingly, the distribution function of our best fit model appears to be bimodal, a result that is also present when the dataset is modeled using its moments and with different dark matter halo profiles. We show that an association to the metal-poor and metal-rich components known to exist in Sculptor is plausible. Hence this constitutes a demonstration of the full power of Schwarzschild's method, because the multi-component nature of Sculptor is not explicitly assumed in our modeling.

In the last chapter (Chapter 5), we demonstrate the existence of distribution functions that can be used to represent spherical massless cored stellar systems embedded in cuspy dark matter halos with constant mildly tangential velocity anisotropy. This provides full support to the fact that the models presented in Chapters 2, 3 and 4 are truly physical. In particular, we derive analytically the functional form of the distribution function for a Plummer stellar sphere in a Hernquist dark halo, for $\beta_0 = -0.5$ and for different degrees of embedding. This particular example satisfies the condition that the central logarithmic slope of the light profile $\gamma_0 > 2\beta_0$. Our models have velocity dispersion profiles similar to those observed in nearby dwarf spheroidal galaxies. Hence they can be used to generate initial conditions for a variety of problems, including N-body simulations that may represent dwarf galaxies in the Local Group.

1.4 Future directions

In the previous sections we have discussed the status of the field, and have begun to identify directions where more research would be desirable to understand the properties and dynamics of dSph. In the case of the dynamical modeling, much of this work, including this Thesis, has been done assuming that the dSph are embedded in spherical dark matter halos (and often, even assuming their light distribution is approximately spherical). First attempts to veer from this assumption have been made using the Jeans equations (Hayashi & Chiba 2012), but these have the limitation of exploring parametric models. It is therefore desirable to apply non-parametric modeling, for example along the lines of Jardel & Gebhardt (2012, although these authors still assume the dark halos are spherical), who have used Schwarzschild models assuming non-spherical light distributions. In fact, van den Bosch et al. (2008) have performed triaxial modeling of elliptical galaxies, implying that the tools needed for the dSph may already have been largely developed (see also van de Ven et al. 2006, for the modeling of OmegaCen in the limit of axisymmetry). Another example, is the M2M modeling of the Galactic bar by Long et al. (2013). Unfortunately the sample sizes for the most classical dwarfs may be still too small to warrant such sophisticated approaches, with the possible exceptions of Sculptor and Fornax. Nonetheless, it is worthwhile establishing what are the degeneracies/limitations in the modeling, and to what extent they can be broken by different datasets. This can be addressed by applying dynamical models on Mock datasets, for example extracted from N-body simulations or generated from known distribution functions.

The use of proper motion measurements of stars in dSph is another unexplored aspect of the dynamical modeling. The reason is, of course, that this has been beyond the capabilities of current instrumentation. However, the situation is likely to change in the coming years. For example, it is now possible to constrain the mean tangential motions of dSph using the Hubble Space Telescope (see Piatek & Pryor 2008, and references therein), and these measurements are likely to be significantly more accurate with the advent of Gaia[5]. The internal motions may still be just about beyond reach for individual stars in dSph. For example for a star at 70 kpc, an internal tangential velocity of $v_t \sim 10\ \mathrm{km\,s^{-1}}$ translates into proper motion of $\mu \sim 30\mu$as/yr. For a star of magnitude $G \sim 17$, the accuracy expected for the Gaia mission is $\sigma_\mu \sim 36\mu$as/yr, and hence the internal velocity and its error will be of comparable magnitude. This implies, however, that one ought to be able to bin the data to obtain a tangential velocity curves with reasonable accuracy, as the error on the dispersion is inversely proportional to the square root of the sample size. For the UFDs the situation is less clear, as these objects have faint and sparsely populated red giant branches. At these characteristic faint magnitudes Gaia's proper motion accuracies degrade quickly, from $\sim 80\mu$as/yr at $G = 18$, to 140μas/yr at $G = 20$. For an object at a distance of ~ 40 kpc, this implies an error in the tangential velocity of $\sim 15-25\ \mathrm{km\,s^{-1}}$. Although in principle one can reduce this error by binning, the sparsely populated RGBs prevent from obtaining the significant gains needed to characterize the internal kinematics of the UFDs. Nonetheless, these measurements clearly will allow the determination of their orbit, as well as aid in establishing membership and potentially finding extra-tidal stars and streamers. On the other hand, for the brightest stars in the LMC (those with $G \lesssim 15$), the expected accuracies are in the range $\sim 4-14\mu$as/yr, which

[5] See http:www.rssd.esa.intindex.php?project=GAIA&page=Science_Performance for the latest estimates of its performance.

at a distance of ~ 50 kpc, corresponds to a tangential velocity error of $\sim 0.1 - 3.3$ km/s, comparable to what can be obtained nowadays for the l.o.s. velocities routinely from the ground.

Wilkinson et al. (2002) have studied the impact of proper motion information following the specifications planned for the former SIM mission, namely 3 – 6 μas/yr, which translates into 1-2 km/s for stars in Draco of magnitude $V \sim 19-20$. These authors show that by adding proper motion information for samples as small as 160 stars, it is possible to obtain accurate estimations of both the velocity anisotropy and mass slope, and thereby break modeling degeneracies unambiguously. This is also confirmed by Strigari et al. (2007a), who show that, for general dark matter density and anisotropy profiles, the log slope of the dark matter profile at about $\sim 2r_c$ can be measured to within ± 0.2 if the proper motions of 200 stars (with tangential velocity errors of ~ 5 km/s) are added to the l.o.s. velocity measurements. This would allow to place tighter constraints on the type of dark matter halos hosting dSph, and hence possibly also on the nature of dark matter.

Bibliography

Aaronson, M. 1983, ApJ, 266, L11

Aaronson, M., & Olszewski, E. 1987, in IAU Symposium, Vol. 117, Dark matter in the universe, ed. J. Kormendy & G. R. Knapp, 153–158

Agnello, A., & Evans, N. W. 2012, ApJ, 754, L39

Amorisco, N. C., & Evans, N. W. 2012a, MNRAS, 419, 184

—. 2012b, MNRAS, 424, 1899

An, J. H., & Evans, N. W. 2006, ApJ, 642, 752

—. 2009, ApJ, 701, 1500

—. 2011, MNRAS, 413, 1744

Armandroff, T. E., & Da Costa, G. S. 1986, AJ, 92, 777

Armandroff, T. E., Olszewski, E. W., & Pryor, C. 1995, AJ, 110, 2131

Battaglia, G., Helmi, A., Tolstoy, E., Irwin, M., Hill, V., & Jablonka, P. 2008a, ApJ, 681, L13

Battaglia, G., Irwin, M., Tolstoy, E., de Boer, T., & Mateo, M. 2012, ApJ, 761, L31

Battaglia, G., Irwin, M., Tolstoy, E., Hill, V., Helmi, A., Letarte, B., & Jablonka, P. 2008b, MNRAS, 383, 183

Battaglia, G., Tolstoy, E., Helmi, A., Irwin, M., Parisi, P., Hill, V., & Jablonka, P. 2011, MNRAS, 411, 1013

Battaglia, G., et al. 2006, A&A, 459, 423

Belokurov, V., et al. 2007, ApJ, 654, 897

—. 2006, ApJ, 647, L111

Benson, A. J., Frenk, C. S., Lacey, C. G., Baugh, C. M., & Cole, S. 2002, MNRAS, 333, 177

Bertone, G., Hooper, D., & Silk, J. 2005, Phys. Rep., 405, 279

Binney, J., & Tremaine, S. 2008, Galactic Dynamics: Second Edition (Princeton University Press)

Boylan-Kolchin, M., Bullock, J. S., & Kaplinghat, M. 2011, MNRAS, 415, L40

Breddels, M. A., & Helmi, A. 2013, arXiv:13XX

Breddels, M. A., Helmi, A., van den Bosch, R. C. E., van de Ven, G., & Battaglia, G. 2012, arXiv:1205.4721

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

Bullock, J. S., Kravtsov, A. V., & Weinberg, D. H. 2000, ApJ, 539, 517

Busha, M. T., Evrard, A. E., & Adams, F. C. 2007, ApJ, 665, 1

Chanamé, J., Kleyna, J., & van der Marel, R. 2008, ApJ, 682, 841

Ciotti, L., & Morganti, L. 2010, MNRAS, 408, 1070

Colín, P., Avila-Reese, V., & Valenzuela, O. 2000, ApJ, 542, 622

Colín, P., Valenzuela, O., & Avila-Reese, V. 2008, ApJ, 673, 203

de Blok, W. J. G. 2010, Advances in Astronomy, 2010

de Lorenzi, F., Debattista, V. P., Gerhard, O., & Sambhus, N. 2007, MNRAS, 376, 71

Evans, N. W., An, J., & Walker, M. G. 2009, MNRAS, 393, L50

Feng, J. L. 2010, ARA&A, 48, 495

Font, A. S., et al. 2011, MNRAS, 417, 1260

Gerhard, O. E. 1993, MNRAS, 265, 213

Gilmore, G., Wilkinson, M. I., Wyse, R. F. G., Kleyna, J. T., Koch, A., Evans, N. W., & Grebel, E. K. 2007, ApJ, 663, 948

Governato, F., et al. 2010, Nature, 463, 203

Hargreaves, J. C., Gilmore, G., Irwin, M. J., & Carter, D. 1994a, MNRAS, 269, 957

—. 1994b, MNRAS, 271, 693

Hayashi, K., & Chiba, M. 2012, ApJ, 755, 145

Helmi, A., et al. 2006, ApJ, 651, L121

Hinshaw, G., et al. 2012, ArXiv e-prints

Hooper, D., & Baltz, E. A. 2008, Annual Review of Nuclear and Particle Science, 58, 293

Jardel, J. R., & Gebhardt, K. 2012, ApJ, 746, 89

Jardel, J. R., Gebhardt, K., Fabricius, M. H., Drory, N., & Williams, M. J. 2013, ApJ, 763, 91

Kazantzidis, S., Łokas, E. L., Callegari, S., Mayer, L., & Moustakas, L. A. 2011, ApJ, 726, 98

King, I. 1962, AJ, 67, 471

Klessen, R. S., Grebel, E. K., & Harbeck, D. 2003, ApJ, 589, 798

Kleyna, J., Wilkinson, M. I., Evans, N. W., Gilmore, G., & Frayn, C. 2002, MNRAS, 330, 792

Kleyna, J. T., Wilkinson, M. I., Evans, N. W., & Gilmore, G. 2001, ApJ, 563, L115

Klimentowski, J., Łokas, E. L., Kazantzidis, S., Mayer, L., Mamon, G. A., & Prada, F. 2009, MNRAS, 400, 2162

Klypin, A., Kravtsov, A. V., Valenzuela, O., & Prada, F. 1999, ApJ, 522, 82

Koch, A., Grebel, E. K., Kleyna, J. T., Wilkinson, M. I., Harbeck, D. R., Gilmore, G. F., Wyse, R. F. G., & Evans, N. W. 2007a, AJ, 133, 270

Koch, A., Grebel, E. K., Wyse, R. F. G., Kleyna, J. T., Wilkinson, M. I., Harbeck, D. R., Gilmore, G. F., & Evans, N. W. 2006, AJ, 131, 895

Koch, A., Kleyna, J. T., Wilkinson, M. I., Grebel, E. K., Gilmore, G. F., Evans, N. W., Wyse, R. F. G., & Harbeck, D. R. 2007b, AJ, 134, 566

Koch, A., Wilkinson, M. I., Kleyna, J. T., Gilmore, G. F., Grebel, E. K., Mackey, A. D., Evans, N. W., & Wyse, R. F. G. 2007c, ApJ, 657, 241

Koposov, S., et al. 2008, ApJ, 686, 279

Koposov, S. E., Yoo, J., Rix, H.-W., Weinberg, D. H., Macciò, A. V., & Escudé, J. M. 2009, ApJ, 696, 2179

Li, Y.-S., De Lucia, G., & Helmi, A. 2010, MNRAS, 401, 2036

Łokas, E. L. 2001, MNRAS, 327, L21

—. 2002, MNRAS, 333, 697

—. 2009, MNRAS, 394, L102

Łokas, E. L., & Mamon, G. A. 2003, MNRAS, 343, 401

Łokas, E. L., Mamon, G. A., & Prada, F. 2005, MNRAS, 363, 918

Long, R. J., & Mao, S. 2010, MNRAS, 405, 301

Long, R. J., Mao, S., Shen, J., & Wang, Y. 2013, MNRAS, 428, 3478

Lovell, M. R., et al. 2012, MNRAS, 420, 2318

Lynden-Bell, D. 1960, MNRAS, 120, 204

Macciò, A. V., Dutton, A. A., van den Bosch, F. C., Moore, B., Potter, D., & Stadel, J. 2007, MNRAS, 378, 55

Macciò, A. V., Paduroiu, S., Anderhalden, D., Schneider, A., & Moore, B. 2012, MNRAS, 424, 1105

—. 2013, MNRAS, 428, 3715

Majewski, S. R., et al. 2005, AJ, 130, 2677

Mamon, G. A., Biviano, A., & Boué, G. 2013, MNRAS, 429, 3079

Martin, N. F., et al. 2009, ApJ, 705, 758

Martínez-Delgado, D., Alonso-García, J., Aparicio, A., & Gómez-Flechoso, M. A. 2001, ApJ, 549, L63

Mateo, M., Olszewski, E., Welch, D. L., Fischer, P., & Kunkel, W. 1991, AJ, 102, 914

Mateo, M., Olszewski, E. W., Vogt, S. S., & Keane, M. J. 1998, AJ, 116, 2315

Mateo, M., Olszewski, E. W., & Walker, M. G. 2008, ApJ, 675, 201

Mayer, L., Governato, F., Colpi, M., Moore, B., Quinn, T., Wadsley, J., Stadel, J., & Lake, G. 2001, ApJ, 547, L123

McConnachie, A. W., et al. 2009, Nature, 461, 66

Merrifield, M. R., & Kent, S. M. 1990, AJ, 99, 1548

Merritt, D. 1993, ApJ, 413, 79

Merritt, D., & Saha, P. 1993, ApJ, 409, 75

Moore, B., Ghigna, S., Governato, F., Lake, G., Quinn, T., Stadel, J., & Tozzi, P. 1999, ApJ, 524, L19

Muñoz, R. R., et al. 2005, ApJ, 631, L137

Muñoz, R. R., Majewski, S. R., & Johnston, K. V. 2008, ApJ, 679, 346

Muñoz, R. R., et al. 2006, ApJ, 649, 201

Navarro, J. F., Eke, V. R., & Frenk, C. S. 1996a, MNRAS, 283, L72

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

—. 1997, ApJ, 490, 493

Navarro, J. F., et al. 2010, MNRAS, 402, 21

Olszewski, E. W., Pryor, C., & Armandroff, T. E. 1996, AJ, 111, 750

Palma, C., Majewski, S. R., Siegel, M. H., Patterson, R. J., Ostheimer, J. C., & Link, R. 2003, AJ, 125, 1352

Pasquini, L., et al. 2002, The Messenger, 110, 1

Peñarrubia, J., McConnachie, A. W., & Navarro, J. F. 2008a, ApJ, 672, 904

Peñarrubia, J., Navarro, J. F., & McConnachie, A. W. 2008b, ApJ, 673, 226

Peñarrubia, J., Pontzen, A., Walker, M. G., & Koposov, S. E. 2012, ApJ, 759, L42

Piatek, S., & Pryor, C. 2008, in IAU Symposium, Vol. 248, IAU Symposium, ed. W. J. Jin, I. Platais, & M. A. C. Perryman, 244–247

Plummer, H. C. 1911, MNRAS, 71, 460

Pontzen, A., & Governato, F. 2012, MNRAS, 421, 3464

Prusti, T. 2011, in EAS Publications Series, Vol. 45, EAS Publications Series, 9–14

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

Rashkov, V., Madau, P., Kuhlen, M., & Diemand, J. 2012, ApJ, 745, 142

Read, J. I., & Gilmore, G. 2005, MNRAS, 356, 107

Revaz, Y., & Jablonka, P. 2012, A&A, 538, A82

Revaz, Y., et al. 2009, A&A, 501, 189

Richstone, D. O., & Tremaine, S. 1984, ApJ, 286, 27

Rix, H., de Zeeuw, P. T., Cretton, N., van der Marel, R. P., & Carollo, C. M. 1997, ApJ, 488, 702

Sawala, T., Scannapieco, C., & White, S. 2012, MNRAS, 420, 1714

Schwarzschild, M. 1979, ApJ, 232, 236

Seljak, U., Makarov, A., McDonald, P., & Trac, H. 2006, Physical Review Letters, 97, 191303

Sersic, J. L. 1968, Atlas de galaxias australes

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

Somerville, R. S. 2002, ApJ, 572, L23

Springel, V., et al. 2008, MNRAS, 391, 1685

Starkenburg, E., et al. 2013, MNRAS, 429, 725

—. 2010, A&A, 513, A34

Strigari, L. E., Bullock, J. S., & Kaplinghat, M. 2007a, ApJ, 657, L1

Strigari, L. E., Bullock, J. S., Kaplinghat, M., Diemand, J., Kuhlen, M., & Madau, P. 2007b, ApJ, 669, 676

Strigari, L. E., Bullock, J. S., Kaplinghat, M., Simon, J. D., Geha, M., Willman, B., & Walker, M. G. 2008, Nature, 454, 1096

Strigari, L. E., Frenk, C. S., & White, S. D. M. 2010, MNRAS, 408, 2364

Syer, D., & Tremaine, S. 1996, MNRAS, 282, 223

Teyssier, R., Pontzen, A., Dubois, Y., & Read, J. I. 2013, MNRAS, 429, 3068

Tolstoy, E., et al. 2006, The Messenger, 123, 33

Tolstoy, E., Hill, V., & Tosi, M. 2009, ARA&A, 47, 371

Tolstoy, E., et al. 2004, ApJ, 617, L119

van de Ven, G., van den Bosch, R. C. E., Verolme, E. K., & de Zeeuw, P. T. 2006, A&A, 445, 513

van den Bosch, R. C. E., van de Ven, G., Verolme, E. K., Cappellari, M., & de Zeeuw, P. T. 2008, MNRAS, 385, 647

van der Marel, R. P., Cretton, N., de Zeeuw, P. T., & Rix, H.-W. 1998, ApJ, 493, 613

Vera-Ciro, C. A., Helmi, A., Starkenburg, E., & Breddels, M. A. 2012, arXiv:1202.6061

Viel, M., Becker, G. D., Bolton, J. S., Haehnelt, M. G., Rauch, M., & Sargent, W. L. W. 2008, Physical Review Letters, 100, 041304

Viel, M., Lesgourgues, J., Haehnelt, M. G., Matarrese, S., & Riotto, A. 2005, Phys. Rev. D, 71, 063534

Walker, M. G. 2012, arXiv:1205.0311

Walker, M. G., Mateo, M., & Olszewski, E. W. 2009a, AJ, 137, 3100

Walker, M. G., Mateo, M., Olszewski, E. W., Bernstein, R., Sen, B., & Woodroofe, M. 2007a, ApJS, 171, 389

Walker, M. G., Mateo, M., Olszewski, E. W., Gnedin, O. Y., Wang, X., Sen, B., & Woodroofe, M. 2007b, ApJ, 667, L53

Walker, M. G., Mateo, M., Olszewski, E. W., Peñarrubia, J., Wyn Evans, N., & Gilmore, G. 2009b, ApJ, 704, 1274

—. 2010, ApJ, 710, 886

Walker, M. G., & Peñarrubia, J. 2011, ApJ, 742, 20

Wang, J., Frenk, C. S., Navarro, J. F., Gao, L., & Sawala, T. 2012, MNRAS, 424, 2715

Wang, J., & White, S. D. M. 2009, MNRAS, 396, 709

Westfall, K. B., Majewski, S. R., Ostheimer, J. C., Frinchaboy, P. M., Kunkel, W. E., Patterson, R. J., & Link, R. 2006, AJ, 131, 375

Wilkinson, M. I., Kleyna, J., Evans, N. W., & Gilmore, G. 2002, MNRAS, 330, 778

Willman, B., et al. 2005, ApJ, 626, L85

Wolf, J., Martinez, G. D., Bullock, J. S., Kaplinghat, M., Geha, M., Muñoz, R. R., Simon, J. D., & Avedo, F. F. 2010, MNRAS, 406, 1220

Wu, X., & Tremaine, S. 2006, ApJ, 643, 210

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

Chapter 2

Orbit-based dynamical models of the Sculptor dSph galaxy

M. A. Breddels, A. Helmi,
R.C.E. van den Bosch, G. van de Ven, and G. Battaglia

Accepted for publication in MNRAS

Abstract

We have developed spherically symmetric dynamical models of dwarf spheroidal galaxies using Schwarzschild's orbit superposition method. This type of modelling yields constraints both on the total mass distribution (e.g. enclosed mass and scale radius) as well as on the orbital structure of the system (e.g. velocity anisotropy). This method is thus less prone to biases introduced by assumptions in comparison to the more commonly used Jeans modelling, and it allows us to derive the dark matter content in a robust way. Here we present our results for the Sculptor dwarf spheroidal galaxy, after testing our methods on mock data sets. We fit both the second and fourth velocity moment profile to break the mass-anisotropy degeneracy. For an NFW dark matter halo profile, we find that the mass of Sculptor within 1 kpc is $\mathrm{M}_{1\mathrm{kpc}} = (1.03 \pm 0.07) \times 10^8\ \mathrm{M}_\odot$, and that its velocity anisotropy profile is tangentially biased and nearly constant with radius. The preferred concentration ($c \sim 15$) is low for its dark matter mass but consistent within the scatter found in N-body cosmological simulations. When we let the value of the central logarithmic slope α vary, we find that the best-fit model has $\alpha = 0$, although an NFW cusp or shallower is consistent at 1σ confidence level. On the other hand, very cuspy density profiles with logarithmic central slopes $\alpha < -1.5$ are strongly disfavoured for Sculptor.

2.1 Introduction

The existence of dark matter has been invoked to explain discrepancies in the observed kinematics of (systems of) galaxies. Especially in the last 30 years it has become a key ingredient of our current cosmological model, the Λ cold dark matter paradigm (hereafter ΛCDM). N-body simulations have made clear predictions on how dark matter should be distributed in the Universe. Navarro, Frenk & White (1996) showed that simulated dark halos have a universal internal density distribution, now known as the NFW profile. Although there have been some revisions, the general form has remained, and the inner regions of simulated dark halos are found to be cusped with logarithmic slopes in the range -1.2 to -0.75 (Navarro et al. 2010). CDM simulations have also revealed the existence of a universal spin distribution and of relations between the characteristic parameters of a dark halo such as concentration and mass (e.g Bullock et al. 2001a).

The predictions of the ΛCDM model may be tested using kinematic data. Cleaner tests are generally obtained using tracers located at large distances, i.e. in the regions that are dominated by the dark matter (e.g Romanowsky et al. 2003; Battaglia et al. 2005, 2006; Xue et al. 2008). In these examples, a relatively accurate measurement of the mass contained within a given radius can be obtained, but constraints on the density profile depend on good knowledge of the spatial distribution of the tracers, which may be somewhat uncertain. Another possibility is to use galaxies that are dark matter dominated at all radii, such as low surface brightness systems (de Blok 2010).

An example of the latter class are the dwarf spheroidal (dSph) galaxies satellites of the Milky Way (Mateo 1998). These appear to be the most dark matter dominated galaxies with total dynamical mass to stellar light ratios in the order of 100-1000 $M_\odot/L_\odot$ derived under the assumption of dynamical equilibrium (e.g. Wolf et al. 2010). The nearby dSph galaxies have the additional advantage that individual stars can be resolved, and their red giant branch (RGB) stars are bright enough to measure line-of-sight velocities with errors of a few km s^{-1} (Mateo et al. 1991). The dynamical modelling of these objects is relatively simple since they are rather round, pressure supported and show little or no rotation. Their high dynamical mass-to-light ratios makes these systems ideal to study dark matter halos, especially their internal structure and to constrain their inner density profiles.

Most of the Milky Way dSph satellites have been modelled using the spherical Jeans equations (e.g. Kleyna et al. 2001; Battaglia et al. 2008; Strigari et al. 2008; Łokas 2009; Walker et al. 2009b), while for more distant objects, such as the dSph satellites of M31, masses have been derived from the average velocity dispersion and projected mass estimators (Kalirai et al. 2010; Collins et al. 2010). In Jeans models one has to specify (i) the form of the light distribution, (ii) the density profile (or equivalently the gravitational potential) of the dark matter component, and (iii) velocity anisotropy of the stars. These characterise a given Jeans model, from which the second velocity moment projected along the line-of-sight can be computed. This is then compared to the measured line-of-sight velocity dispersion of the stars at different locations across the galaxy to establish the performance and characteristic parameters of the specific model.

Jeans modelling suffers from a number of limitations. Firstly the functional form of the velocity anisotropy has to be specified a priori while it is generally unknown. This is because precise measurements of the proper motions of stars in dSph are well beyond reach with current instrumentation. Also inherent to the method is the comparison be-

tween the moments of the model to those of the data which requires binning of the data and generally implies loss of information. It is also important to note that there is no guarantee that the resulting distribution function is non-negative everywhere, a requirement for it to be physical. Nonetheless, there have been interesting discoveries based the use of the Jeans equations and which are robust to assumptions of the underlying anisotropy profile. These include for example, the existence of a possible common mass scale of dwarf spheroidals (e.g. Strigari et al. 2008), and the tight constraints on the total mass within the half-light radius of these systems (Walker et al. 2010; Wolf et al. 2010).

Recently, An & Evans (2009) demonstrated that if the tracer population is supported by a spherical dark halo with a core or a cusp (less steep than a singular isothermal sphere), then the central value of the logarithmic slope γ_0 of the light profile and the central velocity anisotropy β_0 are related as $\gamma_0 = 2\beta_0$. This is valid if $\sigma_r(0) > 0$, i.e. only if the stars are not dynamically cold in this region. This would imply that the derived existence of a cusp or core at the centre could merely be a consequence of the assumptions alone, if just the second velocity moment is modelled using Jeans equations (see also Ciotti & Morganti 2010, who show that a density slope-anisotropy inequality $\gamma > 2\beta$ holds at all radii, at least for a specific class of distribution functions for spherical systems). Thus care is required in interpreting the outcome of this type of models.

The above discussions shows clearly that there is a need to go beyond the modelling of the second moment using Jeans equations. For example, Łokas (2002) proposed to use higher moments to constrain the internal dynamics of dSphs since the kurtosis profile depends mostly on anisotropy while the velocity dispersion depends both on mass and anisotropy Łokas et al. (2005), hence this lifts some of the degeneracies. Other possibilities would be to use parametrised phase-space distribution functions as pioneered by Kleyna et al. (2001, 2002); Wilkinson et al. (2002a) (see also Amorisco & Evans 2012a), or the made-to-measure technique (Syer & Tremaine 1996; Long & Mao 2010).

In this chapter we take a different approach and use a technique named Schwarzschild modelling (Schwarzschild 1979) to probe the internal dynamics and characterise the dark matter content of the Sculptor dwarf spheroidal galaxy. The basic steps of the Schwarzschild method are to integrate a set of orbits in a given potential, calculate the predicted observables for each orbit, and then to weigh the orbits (with non-negative weights) to obtain a model that fits the observed data well in a χ^2 sense. This approach guarantees that the distribution function (which is reflected in the orbit weights) is non-negative. This method was originally used by Schwarzschild (1979) to prove that a self consistent solution in dynamic equilibrium exists for a triaxial system, but was only implemented to reproduce the density distribution. The method was later extended to include kinematic constrains (Richstone & Tremaine 1984; Pfenniger 1984). Since then many codes have been developed (e.g. Richstone & Tremaine 1984; Rix et al. 1997; van der Marel et al. 1998; Cretton et al. 1999; Valluri et al. 2004; van den Bosch et al. 2008). While first only the lowest moments of the line of sight velocity distribution (mean velocity and velocity dispersion) were fitted, better data have led to the inclusion of higher moments in the fits. While the use of moments allows one to use linear or quadratic programming to find the orbit weights, also likelihood methods using discrete data have been developed (e.g. Merritt & Tremblay 1993; Wu & Tremaine 2006; Chanamé et al. 2008). A great advantage of Schwarzschild modelling is that it does not require the specification of the anisotropy profile, this is in fact an outcome of the model (see also Jardel & Gebhardt 2012; Jardel et al. 2012, for applications on the Fornax and Draco dSph).

Sculptor (Scl) is a dwarf spheroidal galaxy satellite of the Milky Way. It lies at high galactic latitude and is located at a heliocentric distance of 79 kpc. With an ellipticity of 0.32 (axis ratio is 0.68) it is not extremely flattened (Irwin & Hatzidimitriou 1995), allowing us to approximate and model Sculptor as a spherical object. Its luminosity is $L_V = 2.15 \times 10^6 L_\odot$ and one recent estimate of its dynamical mass is $2-3 \times 10^8 \mathrm{M}_\odot$ within 1.8 kpc (Battaglia et al. 2008). Its (stellar) mass distribution can be well fitted with a Plummer profile with scale radius $b = 13.0$ arcmin ($\simeq 0.3$ kpc, Battaglia 2007). Two large kinematic data sets have been compiled by Battaglia et al. (2008) and by Walker et al. (2009a), leading to a total ~ 2000 member stars with radial velocity measurements with errors of ~ 2 km/s. As we show below, the combination of these two data sets together with the Schwarzschild method allows us to constrain the dark matter distribution of Sculptor and its internal orbital structure.

This chapter is organised as follows. In §2.2 we will describe the basic ingredients of Schwarzschild modelling, especially focusing on how it can be applied to dSph data. In §2.3 we validate our model on a mock data set motivated by the current Sculptor data. In §2.4 we apply the technique to Scl data, we present a brief discussion in §2.5, and leave our conclusions to §2.6.

2.2 Dynamical model

In this section we review some of the theory that provides the basis for our Schwarzschild method. We then describe how to generate models and focus later on how these can be fit to the observables.

2.2.1 Generalities

The phase-space structure of a galaxy can be specified by its distribution function (hereafter df) $f(\boldsymbol{x}, \boldsymbol{v})$, where $\boldsymbol{x}$ and $\boldsymbol{v}$ are the position and velocity coordinates respectively. The probability of finding a star in the volume $\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}$ is given by $f(\boldsymbol{x}, \boldsymbol{v})\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}$. All observables may be derived from knowledge of the df. For example the normalised surface density:

$$\mu(x, y) = \int \mathrm{d}z \mathrm{d}\boldsymbol{v} f(\boldsymbol{x}, \boldsymbol{v}), \tag{2.1}$$

where z is the direction along the line-of-sight.

According to the (strong) Jeans (1915) theorem, the df of a steady-state stellar system in which almost all orbits are regular, is a function of the isolating integrals of motion (see also Binney & Tremaine 2008). Spherically symmetric systems (both in the tracer's density and the underlying potential) have only regular orbits and generally respect 4 integrals of motion, the energy and the 3 components of the angular momentum vector. However, if the galaxy shows no rotation, due to symmetry, the df will depend only on the energy and the length of the angular momentum vector, i.e. $f(\boldsymbol{x}, \boldsymbol{v}) = f(E, L)$. Furthermore if the velocity distribution is isotropic, the df can only depend on energy and $f(\boldsymbol{x}, \boldsymbol{v}) = f(E)$.

Most dSph galaxies are so distant that the only phase-space coordinates that may be measured currently are the projected stellar positions on the sky, and the line-of-sight velocities of (a subset of) its stars. These can be used to derive the surface density $\mu_0(R)$ and the moments of the line-of-sight velocity distribution:

$$\mu_0(R) = \int \mathrm{d}z\mathrm{d}\boldsymbol{v} f(E, L), \tag{2.2}$$

$$\mu_2(R) = \frac{1}{\mu_0(R)} \int \mathrm{d}z\mathrm{d}\boldsymbol{v} v_\parallel^2 f(E, L), \tag{2.3}$$

$$\mu_4(R) = \frac{1}{\mu_0(R)} \int \mathrm{d}z\mathrm{d}\boldsymbol{v} v_\parallel^4 f(E, L). \tag{2.4}$$

Here R is the projected distance on the sky from the centre of the galaxy and $v_\parallel$ the velocity along the line-of-sight, after subtraction of the centre of mass mean motion.

The above equations suggest that through comparison to the observables it should be possible to derive the form of the df. In some cases, it may be better to parametrise the df and try to estimate its characteristic parameters by comparison to the data (Wilkinson et al. 2002b; Amorisco & Evans 2012a). However, in this work we prefer to use a non-parametric approach such as the Schwarzschild method. This method uses orbits integrated in a specific gravitational potential as building blocks. From these, light and kinematical profiles may be derived and compared to observations through appropriate weighing of the orbits.

In the case of a dwarf galaxy embedded in a spherical dark matter halo, the gravitational potential can be characterised by a few parameters such as: i) the (enclosed) mass of the dark matter halo M_{DM}, and ii) its scale parameter r_{DM}. Due to the high dynamical mass-to-light ratios of dSphs, we do not expect the stellar mass to have a significant influence on the dynamics of the galaxy. We assume a fixed stellar mass-to-light ratio of $\mathrm{M}_\odot/\mathrm{L}_\odot = 1$ as in Walker et al. (2007), and hence from the light distribution we may directly derive the gravitational potential associated to the stars. In the remainder of the chapter we shall refer to properties related to the stellar mass and luminosity interchangeably.

Thus in practise, for a given set of parameters of the potential, we integrate orbits and match these to the observations by adjusting the orbital weights. We then repeat this exercise for other values of these parameters. This can be used to establish the values of the set of parameters which result in a better fit to the observables.

2.2.2 From the model to the observables

Our Schwarzschild method is based on many of the ideas of Rix et al. (1997) and van den Bosch et al. (2008). It is however, a new implementation that is optimised for spherical symmetry. Among other small improvements, our code can be run in parallel and is therefore significantly faster; furthermore for each orbit, we do not store the full line-of-sight velocity distribution but only its moments, which also reduces the computational load.

We now focus on how to generate the observables, namely the surface density and moments of the line-of-sight velocity distribution of the models and how to compare these to data.

For convenience we define $l = L/L_{\mathrm{max}}$ the relative angular momentum (where L_{max} is the angular momentum of a circular orbit of energy E), such that $l \in [0, 1]$. This enables us to define a rectangular grid in energy and relative angular momentum. Since

the Schwarzschild method is based on orbit integrations, the df may be seen as a sum of Dirac delta functions:

$$f(E,L) = \sum_{i,j} \hat{f}_{i,j} \delta(E - E_i) \delta(L - l_j L_{\max,i}), \tag{2.5}$$

where $\sum_{i,j} \hat{f}_{i,j} = 1$ and $\hat{f}_{i,j} \geq 0$.

To define the grid in energy and (relative) angular momentum we proceed as follows. For the energy we choose N'_E radii between a minimum and maximum radius spaced logarithmically, and take the corresponding energy of a purely radial orbit. The minimum and maximum radii we consider are 0.033 kpc and 24.492 kpc, respectively. For each energy we choose N'_l relative angular momenta spaced linearly between 0 and 1. All orbits are integrated starting from their apocentre.

We also define N_R radial bins on the sky, defined by radii at the edges R_k ($k = 0...N_R$). The borders are determined by the kinematic data, by requiring for instance that each bin contains a particular number of stars.

In general, it is convenient to work with the (normalised) mass in a given radial bin:

$$\frac{\mathrm{d}m_*(R)}{M_*} = 2\pi R \mu_0(R) \mathrm{d}R. \tag{2.6}$$

Thus the mass contributed by orbit of energy E_i and relative angular momentum l_j in the radial bin k is:

$$\frac{\Delta m_{*,i,j,k}}{M_*} = \int_{R_k}^{R_{k+1}} 2\pi R \mu_{0,i,j}(R) \mathrm{d}R. \tag{2.7}$$

In the Schwarzschild method this quantity is obtained by integrating the i,j orbit and calculating the fractional time this orbit spends in radial bin k. Since we integrate the orbit with a fixed time step, this is simply equivalent to counting the number of times the orbit crosses bin k, divided by the number of time steps. To reflect the spherical symmetry, at each time step the position and velocities are rotated randomly $N_{\mathrm{rot}} = 25$ times, as in Rix et al. (1997, Eq. 2). Each orbit is integrated for 100 orbital timescales t_{orb}, with $t_{\mathrm{orb}} = 2\pi r_{\mathrm{a}}/v_{\mathrm{circ}}$, and where r_{a} is the apocentre radius and v_{circ} the circular velocity at r_{a}. Each orbit is stored at 1000 points (separated by a constant time step). Therefore the total mass (contributed by all orbits) in bin k is:

$$\begin{aligned} \frac{\Delta m_{*k}}{M_*} &= \sum_{i=1}^{N'_E} \sum_{j=1}^{N'_l} g(E_i, L_j) \hat{f}_{i,j} L_{\max} \Delta E_i \Delta l_i \times \frac{\Delta m_{*,i,j,k}}{M_*} \\ &= \sum_{i=1}^{N'_E} \sum_{j=1}^{N'_l} c'_{i,j} \times \frac{\Delta m_{*,i,j,k}}{M_*}, \end{aligned} \tag{2.8}$$

where $g(E,L)$ is the density of states. The coefficients $c'_{i,j}$ are known as the orbital weights.

We may now proceed to calculate the light-weighted second and fourth moments of the line-of-sight velocity distribution in a given projected radial bin k as:

$$\mu_{2,k} = \frac{M_*}{\Delta m_{*k}} \sum_{i=1}^{N'_E} \sum_{j=1}^{N'_l} c'_{i,j} \int_{R_k}^{R_{k+1}} 2\pi R \mu_{0,i,j}(R) \mu_{2,i,j}(R) \mathrm{d}R, \tag{2.9}$$

$$\mu_{4,k} = \frac{M_*}{\Delta m_{*k}} \sum_{i=1}^{N'_E} \sum_{j=1}^{N'_l} c'_{i,j} \int_{R_k}^{R_{k+1}} 2\pi R \mu_{0,i,j}(R) \mu_{4,i,j}(R) \mathrm{d}R, \tag{2.10}$$

where $\mu_{2,i,j}(R)$ and $\mu_{4,i,j}(R)$ are the second and fourth moment respectively of orbit i, j. The integral is also derived from the orbit integrations. However, instead of counting each time the orbit is found in bin k, we add the corresponding second moment in quadrature (and to the fourth power for the fourth moment) and at the end divide by the number of time steps. Note that the moments are linear in the orbital weights, which allows us to find a solution using quadratic programming, while for instance the kurtosis ($\gamma_2 = \mu_4/\mu_2^2$) is not.

It is possible to consider the orbit weights ($c'_{i,j}$) as free parameters whose exact values will be determined through comparison to the observables. However this would imply that the number of orbits that are integrated to reproduce the observables is exactly equal to the number of free parameters that define the df. Decoupling these two sets of quantities is clearly desirable, see e.g. Cretton et al. (1999). This procedure is known as *dithering* and results in smoother density distributions while keeping the number of free parameters in the distribution function small.

While we may use $N'_E \times N'_l$ orbits to reproduce the observables, we choose only $N_E \times N_l = N'_E \times N'_l / (N_{d_E} \times N_{d_l})$ free parameters to characterise the distribution function, where we take $N_{d_E} \times N_{d_l} = 8 \times 8 = 64$. The coefficients of the distribution function $c_{i,j}$ are related to the orbit weights ($c'_{i,j}$) as follows:

$$c'_{i,j} = \frac{1}{N_{d_E} \times N_{d_l}} c_{i \setminus N_{d_E}, j \setminus N_{d_l}}, \tag{2.11}$$

where $\setminus$ indicates the integer part, e.g. $[i/N_{d_E}]$. Therefore $N_{d_E} \times N_{d_l}$ orbits share the same df coefficient. In practice, one can simply average the quantities obtained from the individual orbits. We choose $N_E = 20$ and $N_l = 8$, which results in $20 \times 8 = 160$ free parameters for the distribution function, but we integrate $20 \times 8 \times 8 \times 8 = 10250$ orbits.

To fit models to the data we generally use projected quantities (i.e. the observables). However, if one knows (or has derived) the df coefficients, it is also possible to make predictions for quantities that are not (yet) directly observable, such as the intrinsic (3d) density distribution or moments of the full velocity distribution. For example, the mass contained in the (spherical) radial bin m contributed by orbit i, j is

$$\frac{\Delta m_{*,3d,i,j,m}}{M_*} = \int_{r_m}^{r_{m+1}} 4\pi r^2 \nu_{*,i,j}(r) \mathrm{d}r, \tag{2.12}$$

where the integral is computed from the orbital integrations, and $\nu_{*,i,j}(r)$ is the radial density profile of orbit i, j. In practise we use $N_r = 50$ (3d) radial bins, spaced linearly between $r_{\min} = 0$ kpc and $r_{\max} = 1.5$ kpc. Similarly we also store the radial and tangential velocity dispersions in these bins. Although we do not store the intrinsic properties beyond 1.5 kpc, this has no effect on the way the projected properties are

determined. Note that the intrinsic properties are not used in any of the fitting routines but may be used for inferring for instance the intrinsic velocity anisotropy profile.

Orbits are integrated using the GNU Scientific Library (GSL) ordinary differential equation solver using an 8th order (Runge Kutta) Prince-Dormand method. We found that the energy is conserved to better than 0.1%.

2.2.3 Fitting procedure

Light distribution

Our first requirement is for the model to fit the observed light distribution. We assume that this is known accurately. We require that the projected mass (or light) in each bin is matched within 1 per cent. Given our assumption of a constant stellar mass-to-light ratio, we make no distinction between surface brightness and stellar mass surface density in what follows. From the assumed brightness profile $\mu_*(R)$, we calculate:

$$\frac{\Delta m_{*,\mathrm{true},k}}{M_*} = \int_{R_k}^{R_{k+1}} 2\pi R \mu_*(r) \mathrm{d}R, \tag{2.13}$$

and thus require for each projected radial bin k that:

$$\left| \frac{\Delta m_{*,\mathrm{true},k}}{M_*} - \frac{\Delta m_{*,k}}{M_*} \right| \leq 0.01. \tag{2.14}$$

Note that the number of bins for the light does not have to equal the number of bins for the kinematics, in this work we choose 250 bins for fitting the light distribution.

Kinematics

To derive the line-of-sight velocity dispersion profile we calculate the second and fourth moment estimators of the line-of-sight velocity distribution $\hat{\mu}_{2,k}$ and $\hat{\mu}_{4,k}$ in bins containing at least 250 stars. Assuming that the measurement errors are normally distributed, and all measurements and errors are independent and uncorrelated, we can obtain $\hat{\mu}_2$ of the population as follows. The expectation value of the second moment is

$$E[m_2] = E\left[\frac{1}{N}\sum_i^N (v_i + \epsilon_i)^2\right] = \mu_2 + s_2, \tag{2.15}$$

where ϵ_i is the unknown noise of measurement i, which we assume is drawn from a normal distribution with dispersion σ_i (i.e. this is the formal error of measurement i). Hence $s_2 = \langle \sigma_i^2 \rangle = E\left[\frac{1}{N}\sum_i^N \epsilon_i^2\right]$ is the average of the estimated squared errors. Here μ_2 the true value of the second moment. Therefore, our best estimate for the second moment of the underlying population is:

$$\hat{\mu}_2 = \frac{1}{N}\sum_i^N (v_i + \epsilon_i)^2 - s_2. \tag{2.16}$$

Similarly, the expectation value of the fourth moment

$$E[m_4] = E\left[\frac{1}{N}\sum_i^N (v_i + \epsilon_i)^4\right] = \mu_4 + 3s_2^2 + 6\mu_2 s_2, \tag{2.17}$$

where we have used that the fourth moment of a normal distribution is $3\sigma^4$. Therefore our estimate for the fourth moment is:

$$\hat{\mu}_4 = \sum_i^N (v_i + \epsilon_i)^4 - 3s_2^2 + 6\mu_2 s_2, \tag{2.18}$$

where we have assumed $\mu_2 \approx \hat{\mu}_2$.

The variance of the second moment $\mathrm{var}(m_2)$, can be determined using $\mathrm{var}(x) = E[x^2] - (E[x])^2$, which yields

$$\mathrm{var}(m_2) = \frac{1}{N}\left(\mu_4 - \mu_2^2 + 2s_2^2 + 4\mu_2 s_2\right). \tag{2.19}$$

Although we formally need $\mathrm{var}(\hat{\mu}_2)$, we have found by testing with a Gaussian distribution, that for our purposes $\mathrm{var}(\hat{\mu}_2) \approx \mathrm{var}(m_2)$. For the variance of the fourth moment we find:

$$\begin{aligned}\mathrm{var}(m_4) = \mu_8 + 105s_2^4 + 204\mu_4 s_2^2 + 420\mu_2 s_2^3 \\ + \quad 28\mu_6 s_2 - 9s_2^4\end{aligned} \tag{2.20}$$

which require the 6^{th} and 8^{th} moments:

$$E[m_6] = \mu_6 + 15\mu_4 s_2 + 45\mu_2 s_2^2 + 15s_2^3, \tag{2.21}$$

$$E[m_8] = \mu_8 + 210\mu_4 s_2^2 + 28\mu_6 s_2 + 420\mu 2 s_2^3, \tag{2.22}$$

and again we use $\mathrm{var}(\hat{\mu}_4) \approx \mathrm{var}(m_4)$.

The likelihood of the kinematic data given a model is:

$$p(\text{kinematic data}|\text{model}) \propto e^{-\frac{1}{2}\chi^2_{\mathrm{kin}}} \tag{2.23}$$

where

$$\chi^2_{\mathrm{kin}} = \sum_k^{N_{\mathrm{bins}}} \frac{\left(\hat{\mu}_{2,k} - \mu_{2,k}\right)^2}{\mathrm{var}(\hat{\mu}_{2,k})} + \sum_k^{N_{\mathrm{bins}}} \frac{\left(\hat{\mu}_{4,k} - \mu_{4,k}\right)^2}{\mathrm{var}(\hat{\mu}_{4,k})}. \tag{2.24}$$

Here $\mu_{2,k}$ is given by Eq. (2.9), $\hat{\mu}_{2,k}$ is the estimate from the data for bin k and similarly for the fourth moment[1].

Finding a solution

We need to find the $c_{i,j}$ that maximise the probability (Eq. 2.23) or minimise the χ^2_{kin}, under the condition that all $c_{i,j}$ are positive (and sum up to unity) and the light distribution is reproduced to within 1 per cent. This problem can easily be solved by quadratic programming (QP), since the minimisation is quadratic in the df coefficients, and the constraints are linear. Note however that for this non-parametric problem, the parameter space is very large, and a solution will often yield an unrealistically spiky df. To effectively reduce the parameter space and yield a smoother df, we add a regularisation constraint, in analogy to Cretton et al. (1999) and van den Bosch et al. (2008), by including a penalty term to the total χ^2. This term has the form:

$$\chi^2_{\mathrm{reg}} = \chi^2_{\mathrm{reg,E}} + \chi^2_{\mathrm{reg,L}}, \tag{2.25a}$$

[1] Here we have neglected correlations between the moments, although these may exist in practice.

$$\chi^2_{\rm reg,E} = \left(\lambda_E \sum_{j=0}^{N_L} \sum_{i=1}^{N_E-1} -\xi_{i-1} c_{i-1,j} + 2\xi_i c_{i,j} - \xi_{i+1} c_{i+1,j} \right)^2 , \tag{2.25b}$$

$$\chi^2_{\rm reg,L} = \left(\lambda_L \sum_{j=1}^{N_L-1} \sum_{i=0}^{N_E} -\xi_i c_{i,j-1} + 2\xi_i c_{i,j} - \xi_i c_{i,j+1} \right)^2 , \tag{2.25c}$$

where $\chi^2_{\rm reg,E}$ and $\chi^2_{\rm reg,L}$ are small for a smooth df. This smoothness requirement is implemented by demanding the second order derivatives of the df to be small, which we compute by taking second order finite differences (Eqs. 2.25b-2.25c).

In our case we found $\lambda_L = \lambda_E/8$ to work well, and we calibrate λ_E in the next section. The ξ_i terms are the inverse of the (normalised) masses inside the radii defined by our energy grid (§2.2.2) (see also van den Bosch et al. 2008, Eq. 29). Since the regularisation term $\chi^2_{\rm reg}$ is quadratic in the df coefficients, it can also be optimised using the QP.

The total χ^2 now becomes:

$$\chi^2 = \chi^2_{\rm reg} + \chi^2_{\rm kin}, \tag{2.26}$$

Minimising this equation, in combination with the linear constrains of the $c_{i,j}$ and the linear constraints on the light distribution (Eq. 2.14) defines the problem for the QP.

2.3 Testing the method

2.3.1 Plummer profile embedded in an NFW dark matter halo

Mock Sculptor

We now create a mock galaxy that may be representative of Sculptor according to previously published dynamical models of this system (Battaglia et al. 2008). The goal is to test our method in the region of parameter space where we expect Sculptor to be. For the stellar component we choose a Plummer profile with total mass $M_* = 10^6$ M$_\odot$ and a scale radius $b = 0.3$ kpc. The stellar component is embedded in a spherical NFW dark matter halo with scale $r_s = 0.5$ kpc, and enclosed mass at 1 kpc of $M_{\rm DM}(< 1\,{\rm kpc}) = 10^8 {\rm M}_\odot$. The radial density profile for the NFW halo is of the form $\rho_{DM}(r) = \rho_0 (r/r_s)^{-1}(1 + r/r_s)^{-2}$. We set the velocity anisotropy to be constant, $\beta = -1$. Recall that $\beta(r) = 1 - \sigma_t^2(r)/\sigma_r^2(r)$ and $\sigma_t^2(r)$ (where $\sigma_t^2 = \sigma_\phi^2 = \sigma_\theta^2$ for every r) and $\sigma_r^2(r)$ are the second moments of the intrinsic velocity distribution at radius r in the tangential and radial directions respectively. Note that in this model, although the central velocity dispersion is null[2], the line-of-sight velocity dispersion is finite, and has a value $\sigma_{los} = 7.71$ km s^{-1}. By assuming the df to be separable, i.e. $f(E, L) = f_E(E) f_L(L)$, we may compute it explicitly (numerically) as described in Appendix 2.A.

As an extra check that our model galaxy is physical and stable, we have generated phase-space coordinates for 100 000 stars from its df, and simulated it numerically using GADGET-2 (Springel 2005). In this simulation the stars are represented as N-bodies and they are embedded in the static potential given by the dark halo of our mock Sculptor model. We found that, even after 10 Gyr of evolution, the density distribution, velocity dispersion profiles and the anisotropy match the initial values well.

[2] which implies there is no conflict with the An & Evans (2009) theorem.

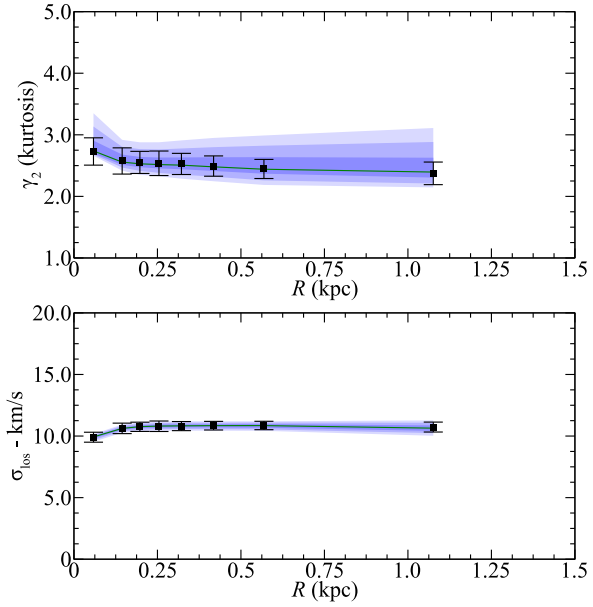


Figure 2.1: The line-of-sight velocity dispersion (bottom) and kurtosis (top) for mock Sculptor. **Black symbols:**: Values for the moments in radial bins from the mock Sculptor data, with 1σ error bars. **Blue contours**: Recovered profiles from the models, where the regions correspond to the 68.3, 95.4 and 99.7 per cent confidence intervals.

To generate observations of our mock Sculptor we could draw a random sample of ~ 2000 stars from its distribution function. However, this has the disadvantage that many realisations would be required to test if the mean of the recovered quantities matches the known input values. Therefore, for the purposes of testing our modelling technique we prefer to compute the moments of the line-of-sight velocity distribution at different radii directly from the known distribution function, as this is less susceptible to randomness. We add uncertainties in the moments and choose the location of the radial bins to match the Sculptor data set. Fig. 2.1 shows the line of sight velocity dispersion profile and the kurtosis derived in this way. Note however in the model fitting we use the second and fourth moments since these are linear in the df coefficients. We calculate the uncertainties in the moments using Eqs.(2.19) and (2.20), assuming no measurement errors since these contribute only $\sim 1\%$ of the error budget for the typical measurement errors of 2 $\mathrm{km\,s^{-1}}$ and line of sight velocity dispersions of 10 $\mathrm{km\,s^{-1}}$ found in dSph. Therefore the uncertainties in the moments are only due to the number of objects per bin. Here we we choose to have 250 stars per bin, which gives a total 8 bins for a sample of 2000 objects.

We proceed to test our code in two steps. In the first instance our aim is to establish how well the method recovers the intrinsic properties of our mock galaxy if the df is known. Thus in this first test we use the known df to compute the df coefficients. These define the orbital weights which our Schwarzschild code uses to calculate the observables. The df coefficients are shown in the upper left panel of Fig. 2.2. The

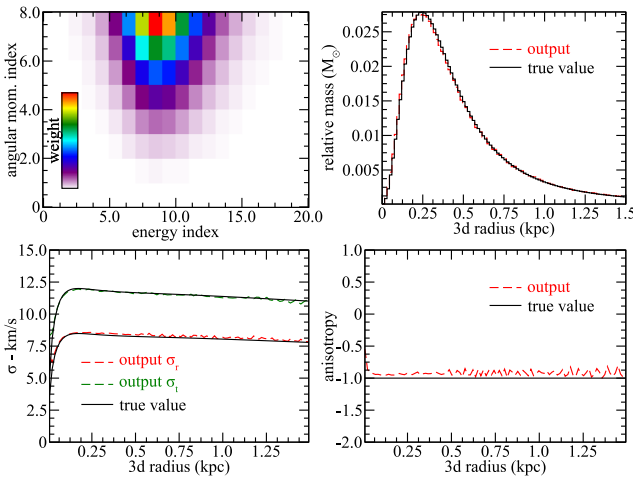


Figure 2.2: Result of a test of our Schwarzschild code on mock Sculptor. Here we have assumed knowledge of the df coefficients and recovered the intrinsic properties of the model.

recovered (normalised) mass per intrinsic (3d) bin (Eq. 2.12), is plotted in the top right panel of the same figure. The red dashed curve shows the output of the Schwarzschild code, while solid black corresponds to the true values. In the lower left panel we plot the velocity dispersions for the radial (red) and tangential (green) directions. The solid curves indicate the true values, whereas in dashed we showed the recovered dispersions. Here the "true" velocity dispersion has been calculated using the Jeans equations (Binney & Tremaine 2008, chapter 4). The lower right panel shows the true (solid black) and the recovered (dashed red) anisotropy as a function of radius. This exercise shows that given the correct weights we are indeed able to recover the known intrinsic properties of our mock galaxy.

The small deviations from the true values especially visible in the anisotropy profile are expected since the df coefficients only approximate the true df. These deviations can thus be removed by increasing the number of df coefficients. For example, if we double the number of coefficients in the energy and angular momentum directions, the small offset between the true and recovered anisotropy profiles disappears. The increase in the resolution in the energy direction also leads to the elimination of the wiggles in the anisotropy profile. On the other hand, the turnover of the anisotropy profile seen at small radii is related to the sampling of orbits with the highest binding energy. Recall that we sample orbits from a minimum radius $r_{\rm min} \sim 0.03$ kpc, so that the highest binding energy radial orbit has its apocentre at $r_{\rm min}$. The orbits that contribute to the region $r \lesssim r_{\rm min}$ are those which are very elongated with pericentres inside this radius and with large apocentres (beyond $r_{\rm min}$), and the set of orbits with the highest binding energy but which have more angular momentum. These more circular orbits only contribute within a small range in radii, and hence the resulting velocity ellipsoid is radially biased. Clearly if we were to reduce $r_{\rm min}$, i.e. increase the sampling of orbits in the central regions, this will lead to a decrease in the radius at which the velocity anisotropy turns over. However, we deem this unnecessary as the amount of mass associated to this region is negligible,

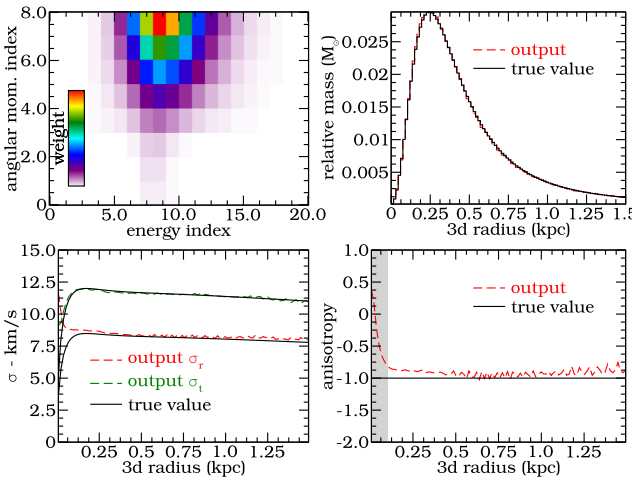


Figure 2.3: Result of the application of the Schwarzschild code on our mock Sculptor. The figure shows that the intrinsic structure is recovered through the QP when the underlying gravitational potential is known. The grey region in the lower right plot indicates where we cannot recover the anisotropy.

and this regime is in fact outside the reach of observations since we only have access to observables along the line-of-sight, and a star at small projected radius could be located at larger physical radii from the centre. Furthermore, the size of the currently available data sets is a strongly limiting factor (see next paragraph).

We now use the full Schwarzschild method, and solve for the df using QP. For the regularisation parameters we found $\lambda_E = 0.1$ to give good results. Fig. 2.3 summarises our findings. The overall properties of the df are well recovered as well as the remaining characteristics (see Fig. 2.2 for comparison). The anisotropy is recovered accurately except for $r \lesssim 0.1$ kpc. This is not due to sampling of highly-bound orbits discussed above, but is mostly driven by the small number of stars in this (3d) inner region. Running the same experiment with a larger data set (10 000 and 50 000 stars) we see the mismatch in the anisotropy to occur at smaller radii. In practise, this means that with the current data sets we are not sensitive to the anisotropy at $r \lesssim 0.1$ kpc.

Global halo parameter recovery

In the above tests we showed that the Schwarzschild code accurately recovers the df and therefore the kinematic properties of our mock dwarf galaxy. This test was done assuming that the (enclosed) mass within 1 kpc of the NFW halo (M_{1kpc}) and its scale (r_s) were known. We now focus on how to estimate these parameters directly.

We proceed to calculate the probability of a model for a set of parameters values. In our case these parameters are M_{1kpc} and r_s. However, instead of calculating this probability on a regular grid as done in e.g. Gebhardt et al. (2007) and van den Bosch et al. (2008), we use an adaptive method, similar to Gebhardt & Thomas (2009). This first finds the probability density function (pdf) on a coarse grid and then determines where the pdf needs to be refined, and does so hierarchically. This allows us to obtain

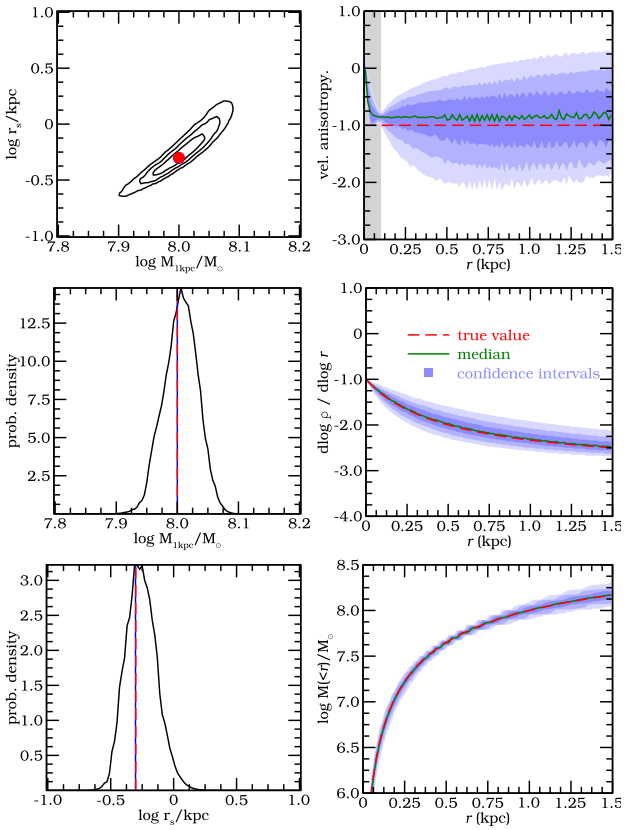


Figure 2.4: Left column: Probability density functions (joint and marginalised) for mass and scale parameters of the NFW dark matter halo potential recovered for mock Sculptor model. Blue dot and blue lines (left column) indicate the maximum likelihood value (of the unmarginalised pdf), while the red dot and vertical dashed lines indicate the input values for the mock Sculptor model. The green solid line indicates the median value and the blue regions (or black contour lines in the top left panel) the 68.3, 95.4 and 99.7 per cent confidence intervals. **Top right:** Recovered anisotropy profile. The grey region indicates where we cannot recover the anisotropy. **Middle right:** Recovered logarithmic density slope (see text) for the dark matter. **Bottom right:** Recovered enclosed mass profile.

a relatively smooth pdf via the evaluation of a small number of models. For each set of model parameters we calculate the (relative) probability as $p \propto e^{-\frac{1}{2}\chi^2_{\rm kin}}$ (Eq. 2.24). This results in estimates of the best fit parameters, as well as in confidence intervals. We assume the prior on $\mathrm{M}_{1\rm kpc}$ to be uniform in $\log \mathrm{M}_{1\rm kpc}$ in the range $\log \mathrm{M}_{1\rm kpc} \in [7.6, 8.2]$[3] and the prior on r_s uniform in $\log r_s$ in the range $\log r_s \in [-1, 1]$.

The pdf for the parameters $\mathrm{M}_{1\rm kpc}$ and r_s for our mock Sculptor model is shown in the top left panel of Fig. 2.4. The pdf is nicely centred on the input values $\mathrm{M}_{1\rm kpc} = 10^8 \mathrm{M}_\odot$ and $r_s = 0.5 \approx 10^{-0.3}$ kpc. The maximum likelihood value (blue dot or lines) almost equals the input value, where the small deviation is caused by the discretisation of the pdf. Although the enclosed mass at 1 kpc is recovered both accurately and precisely (mean $\mathrm{M}_{1\rm kpc} = 1.02{\times}10^{8.00\pm0.03}\mathrm{M}_\odot$, corresponding to a 7% uncertainty, or $\mathrm{M}_{1\rm kpc} = 1.02^{+0.075}_{-0.070}\times 10^8\mathrm{M}_\odot$), the scale radius is more poorly constrained (mean $r_s = 0.56 \times 10^{\pm0.14}$ kpc, corresponding to a 37% uncertainty, or $r_s = 0.56^{+0.21}_{-0.15}$ kpc). Note that the marginalised pdf for $\mathrm{M}_{1\rm kpc}$ and r_s are somewhat asymmetric (a reflection of what is seen in the upper left panel of Fig. 2.4), and this leads to slightly biased mean values for the parameters of the model.

Each Schwarzschild model (i.e. for a given $\mathrm{M}_{1\rm kpc}$ and r_s) results in a single anisotropy profile. To find the pdf of the velocity anisotropy profile one should integrate (marginalise) over all possible df coefficients (as in Magorrian 2006). However this is not always feasible due to the high dimensionality of the parameter space required to specify the df ($N_E \times N_l = 160$ for this model). Instead we take the single anisotropy profile of each model, and calculate the probability density function for the anisotropy as a function of radius as follows:

$$p(\beta|r) = \int \mathrm{dM}_{1\rm kpc} \int \mathrm{d}r_s p(\beta|r, \mathrm{M}_{1\rm kpc}, r_s) p(\mathrm{M}_{1\rm kpc}, r_s). \tag{2.27}$$

We plot the median anisotropy as a function of radius in green in the top right panel of Fig. 2.4, together with the 68.3, 95.4 and 99.7 percent confidence intervals in blue. Note however, that the anisotropy values at different radii are not independent. The input anisotropy is indicated by the red dashed line. The anisotropy seems to be reproduced quite accurately, except at small radii. Since our technique recovers nearly perfectly the input values of the model, the anisotropy profile found is essentially equivalent to that derived in Fig. 2.3. The mismatch at small radii is explained in the previous section, and the apparent small uncertainty in the anisotropy in this region may be understood from the following argument. Using the Jeans equation, we may express the mass within a given radius as

$$GM(r)/r = \sigma_r^2(\gamma - 2\beta - \alpha),$$

where $\gamma = d\log\nu_*/d\log r$, β is the anisotropy, and $\alpha = d\log\sigma_r^2/d\log r$. For any model without a black hole in the centre, the $lhs \to 0$ as $r \to 0$. For a cored profile (as we have assumed) $\gamma = 0$ in this limit. This implies that there is quite a strong restriction on the behaviour of β (and σ_r) at small radii. The above equation implies that as $r \to 0$,

$$2\sigma_r^2 - 2\sigma_t^2 - r d\sigma_r^2/dr = 0,$$

and since $\sigma_r \to 0$ as $r \to 0$ to have a physical solution in a cuspy dark matter halo according to An & Evans (2009), then this means that there is only one possible σ_t at $r = 0$, for any model, i.e. value of $M_{\rm DM}$ and r_s.

[3] Outside this interval the pdf is essentially zero.

In the central right panel of Fig. 2.4 we plot the logarithmic slope (η) of the dark matter density ($\rho_{\rm DM}$) as a function of radius:

$$\eta(r) = \frac{\mathrm{d}\log\rho_{\rm DM}}{\mathrm{d}\log r}. \tag{2.28}$$

In the inner parts $\eta = -1$ and in the outer regions $\eta = -3$ due to our choice of the NFW profile. In the next section we also explore a different functional form for the halo density profile, which makes this plot more meaningful and useful for later comparison.

In the bottom right panel we plot the enclosed (dark matter) mass as a function of radius. The least uncertainty in the enclosed mass is at $r \approx 0.5 - 0.6$ kpc. This radius is close to the half light radius $r_{1/2} \approx 1.3b \approx 0.4$ kpc where Walker et al. (2009b, 2010) and Wolf et al. (2010) find the enclosed mass to be most robustly determined and to be independent of anisotropy.

The line of sight velocity dispersion and the kurtosis profiles obtained from the models are shown as the blue contours in Fig. 2.1. These have been computed in an analogous manner to the anisotropy profile, i.e. as in Eq. (2.27). This figure shows that the resulting curves are in excellent agreement with the input profiles.

To gain further confidence in our methodology, we have also performed a similar set of tests for different anisotropy profiles, while keeping the same stellar and dark matter density profiles. In one case the anisotropy varied from $\beta = -1$ in the centre to $\beta = +0.25$ at larger radii (i.e. from tangentially to radially biased). The other case we have tested has an anisotropy profile that changes from $\beta = 0$ at the centre to $\beta = -1$ at larger radii (i.e. from radial to tangential anisotropy). Also in these cases all the quantities recovered are in excellent agreement with the input values, indicating that our methodology works well and is robust.

2.3.2 Changing the dark matter halo density profile

In reality we will not know the actual density profile of the dark matter halo hosting a galaxy like Sculptor, and we would like to determine this from the data. A particularly interesting quantity is the inner slope of the density profile since this depends on the nature of the dark matter particles themselves, i.e. whether it is cold, warm or self-interacting (Avila-Reese et al. 2001; Spergel & Steinhardt 2000).

Therefore, in this section we use our mock Sculptor, which is embedded in an NFW profile, but we assume a more general functional form to test the performance of our Schwarzschild method, i.e. we take:

$$\rho_{\rm DM}(r) = \rho_0 \left(r/r_s\right)^{\alpha} \left(1 + r/r_s\right)^{-(3+\alpha)}, \tag{2.29}$$

such that for $\alpha = -1$ this reduces to the NFW case. For the orbit integration we need to know the potential (or rather the forces) generated by this density distribution. Since no general analytic expression exists for these general potentials, we have to solve Poisson's equation numerically. We do this using the FEM (Finite Element Method) method (e.g. Pepper & Heinrich 1992). Our basis functions are Lagrange polynomials of degree 0 to 3 (cubic), which leads to a force field of order 2 (quadratic). We use a grid of 200 points in log radius, from $r = 10^{-6} - 10^{4}$ kpc. Testing this in the case of the NFW profile we find that the relative errors in the force in this range are $\sim 10^{-6}$.

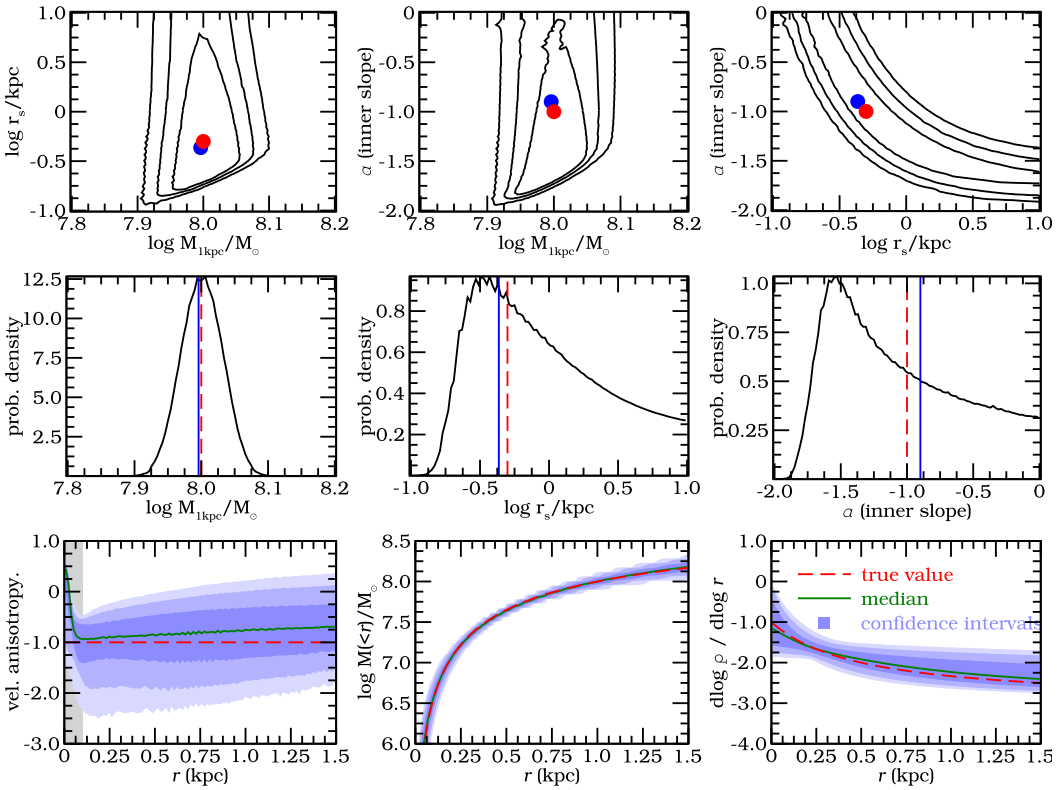


Figure 2.5: Top two rows: Probability density functions (joint and marginalised) for mass, scale and inner slope parameters of the dark matter halo potential recovered for our mock Sculptor model. Blue dots (top row) and blue lines (middle row) indicate the maximum likelihood value (of the unmarginalised pdf), while the red dot and vertical dashed lines indicate the input values for the mock Sculptor model. The green solid line the median value and the blue regions (or black contour lines in the top row) the 68.3, 95.4 and 99.7 per cent confidence intervals. **Bottom left:** Recovered anisotropy profile. The grey region indicates where we cannot recover the anisotropy. **Bottom centre:** Recovered enclosed mass profile. **Bottom right:** Recovered logarithmic density slope (see text) for the dark matter.

We use our Schwarzschild code to find the best model that fits our mock Sculptor data, now with an additional unknown parameter α, assuming a uniform prior in the range $\in [-2, 0]$ ($\alpha > 0$ corresponds to a central hole in the dark matter distribution, which we do not consider). The results are given in Fig. 2.5. The top row in shows the joint pdfs, marginalised over the remaining parameter. The middle row shows the pdfs of the single parameters, marginalised over the other two parameters. The blue dots and blue lines indicate the maximum likelihood value (of the unmarginalised pdf). In the bottom row the recovered anisotropy, mass and density profile are shown.

In general, all quantities are recovered quite well. However, the pdf of α versus $\log r_s$ shows an important degeneracy between these parameters, indicating that it is hard to determine either of these quantities reliably from our mock data set. The maximum likelihood (the blue dot) is slightly offset from the input value (red dot), which may indicate small systematic errors due to for instance the discretisation of the distribution function. However, note that since this systematic offset is in the direction of the degeneracy, the systematic error is small compared to the statistical uncertainty and therefore we do no consider this to be a problem for data sets of this size and quality. This analysis suggests that the current data is not sufficient to provide a good estimate of the inner slope for these models. The limitation lies in the number of stars with spectroscopic measurements (which in the case tested here is 2000) and/or their spatial distribution.

2.4 Application to the Sculptor dSph galaxy

2.4.1 Data and extracted velocity moments

We use the line-of-sight velocities of Battaglia et al. (2008, 1073 stars, hereafter B08)and Walker et al. (2009a, 1541 stars, hereafter W09)[4]. In the case of duplicates (stars in common in the datasets) we average the line-of-sight velocities and the errors (in quadrature). Two observations are considered to be from the same star when the astrometry agrees within 1 arcsec, and a velocity difference less than 3ϵ, where ϵ is the average velocity error. Inspection of the relative distances between stars in the datasets shows that this criterion is optimal to sieve duplicates. This procedure led to the identification of 308 duplicates, roughly 11% of the combined dataset.

To create a velocity dispersion profile, we first need to convert the measurements of the heliocentric line-of-sight velocities into line-of-sight velocities that take into account the space motion of Sculptor. We provide below a brief summary of this procedure and refer the reader to Appendix 2.B for more details.

The heliocentric line-of-sight velocities of Sculptor's stars are shown in Fig. 2.6. As can be seen from this figure, there appears to be a velocity gradient along the major axis (see also Fig. 1 of B08). The presence of such a gradient could be due to intrinsic rotation in Sculptor, as suggested by B08. On the other hand it is also possible that the gradient is a result of the projection of the proper motion of the centre of mass of Sculptor (or a mix of both), in which case it can be used to infer its space velocity (Walker et al. 2008). In absence of independent and direct measurements of the proper motion of Sculptor, it remains debatable what the source of the gradient is. For simplicity, here we assume that

[4] Although Amorisco & Evans (2012b) have reported a systematic velocity offset of 1.5 km/s in the dataset of W09 compared to B08's, we here perform no correction. The various tests we have done show that this offset has no visible effect on the results.

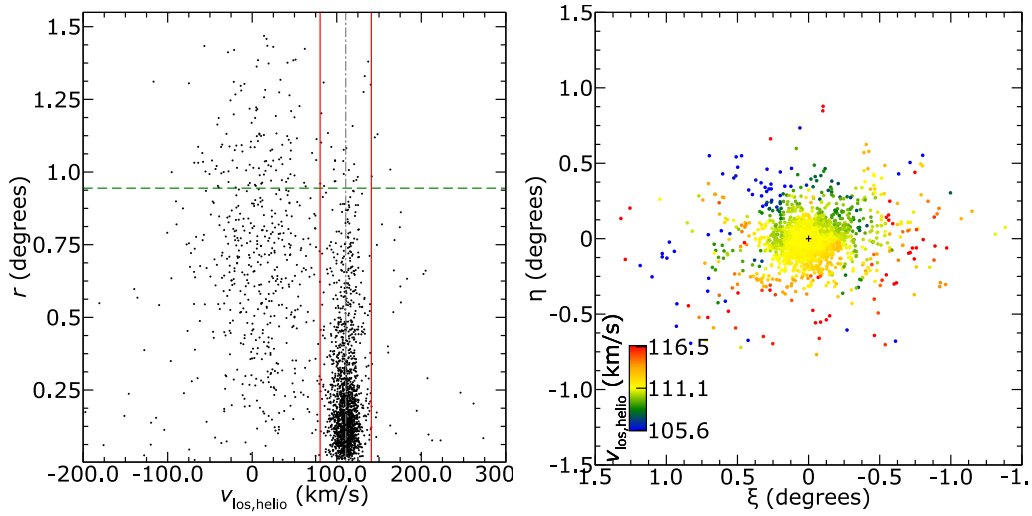


Figure 2.6: Top: Line of sight velocities for Sculptor versus radius. We only use the stars at radii $r < 3400$ arcsec. The grey line indicates the systemic radial velocity for Sculptor and red lines $\pm 3\sigma$ the mean velocity dispersion. **Bottom:** Heliocentric line-of-sight velocities from the combined data set of Battaglia et al. (2008) and Walker et al. (2009a). The velocities have been smoothed by taking the median in cells of 0.2 degrees on a side.

Sculptor does not rotate and we derive the velocity of the centre of mass of Sculptor from the line-of-sight measurements in Appendix 2.B. We note that in practice, our procedure simply removes the gradient, which one might say is equivalent to having removed (solid body) rotation.

Velocity dispersion profile

For our dynamic modelling, we need to calculate the velocity dispersion profile of Sculptor in radial bins. To this end, we initially make a rough selection of the likely members of Sculptor, and then perform a more thorough analysis including the effects of Milky Way contaminants. In the first step, we take the systemic heliocentric radial velocity ($v_{\rm Scl,sys,helio} = 110.6$ km s^{-1}) and the mean velocity dispersion ($\sigma_{\rm Scl} = 10.1$ km s^{-1}) from B08. We require that the member stars are within 3σ of the systemic velocity of Sculptor, as indicated by the red solid lines in the right panel of Fig. 2.6. Furthermore we also require that they are located within $r < 3400$ arcsec (~ 0.94 deg, 1.3 kpc), indicated by the green dashed line in the same panel. We add this requirement since we are not confident that outside this radius a reliable velocity dispersion can be measured due to the low number density of (probable) Sculptor members compared to Milky Way stars. An improved method for discriminating Milky Way contaminants based on surface gravity in the data set of B08 has been developed by (Battaglia & Starkenburg 2012), see also Walker et al. (2009c).

We then define radial bins such that each but the last bin contains at least 250 stars that match these criteria. From the total of unique (i.e. non-duplicates) 2306 stars, 1695

match the above two criteria, resulting in 7 radial bins, where the last one contains 195 stars.

After we defined our bins to include at least 250 probable members, we remove the requirement of being within 3σ of the systemic velocity of Sculptor. We now only require $r < 3400$ arcsec (= 1.3 kpc), so all stars below the green dashed line in top panel of Fig. 2.6 are considered for calculating the velocity dispersions (2153 stars). We now use a model for the velocity distribution of the foreground contamination and of Sculptor itself, which then allows us to calculate the most likely velocity dispersion in each radial bin.

Following B08 and W09 we model the velocity distribution in a radial bin as a sum of Gaussians. The velocity distribution of Sculptor itself is modelled as a single Gaussian, while that of the Milky Way is modelled as a sum of two Gaussians[5], following B08. Then the probability of the velocity dispersion of Sculptor in radial bin j with data D_j is:

$$\begin{aligned} p(\sigma_j|D_j) &= \frac{p(D_j|\sigma_j)p(\sigma_j)}{p(D_j)} = \prod_i^{N_j} \frac{p(D_{j,i}|\sigma_j)p(\sigma_j)}{p(D_{j,i})} \\ &= \prod_i^{N_j} \frac{p(R_{j,i}, v_{j,i}|\sigma_j)p(\sigma_j)}{p(R_{j,i}, v_{j,i})} \\ &\propto p(\sigma_j) \prod_i^{N_j} (p(R_{j,i}, v_{j,i}, m|\sigma_j) \\ &\qquad + p(R_{j,i}, v_{j,i}, \neg m|\sigma_j)) \,, \end{aligned} \tag{2.30}$$

where $R_{j,i}$ and $v_{j,i}$ are the radius and velocity of the i^{th} star in the j^{th} bin, N_j is the number of stars in bin j, $p(\sigma_j)$ is the prior, which we take flat between the range $0 \leq \sigma_j \leq 30$ km s^{-1} and m and $\neg m$ indicate the Boolean value of being a member star of Sculptor or not. The proportionality can be used since the denominator is a normalisation constant. The first terms in the last line of Eq. (2.30) can be expanded further (for each j):

$$\begin{aligned} p(R_i, v_i, m|\sigma) &= p(R_i, v_i|m, \sigma)p(m|\sigma) \\ &= p(R_i, v_i|m, \sigma)p(m) \\ &= p(R_i|m)p(v_i|m, \sigma)p(m) \end{aligned} \tag{2.31}$$

We take the prior on membership, to be equal $p(m) = p(\neg m) = \frac{1}{2}$. Using the model of Sculptor as described above, $p(R_i|m) = \mu_{\text{Scl}}(R_i)$, the normalised surface density and $p(v_i|m, \sigma)$ is a Gaussian convolved with the individual measurement errors on v_i.

The second term in Eq. (2.30) can similarly be derived by replacing m with $\neg m$ in Eq. (2.31), $p(R_i|\neg m) = \mu_{\text{MW}}$ is the density of the Milky Way foreground. Since the normalisation is not important, we only need to know the ratio $\mu_{\text{MW}}/\mu_{\text{Scl}}(R_i)$ in each bin[6] . This can be estimated by the ratio of stars outside the 3σ and inside the 3σ

[5] This gives a good fit to the Besançon model in this region of the sky and for stars with colours and magnitudes in the observed range (Robin et al. 2003).

[6] Although the sampling of B08 and W09 is different, we have found in tests that this has no influence on our results.

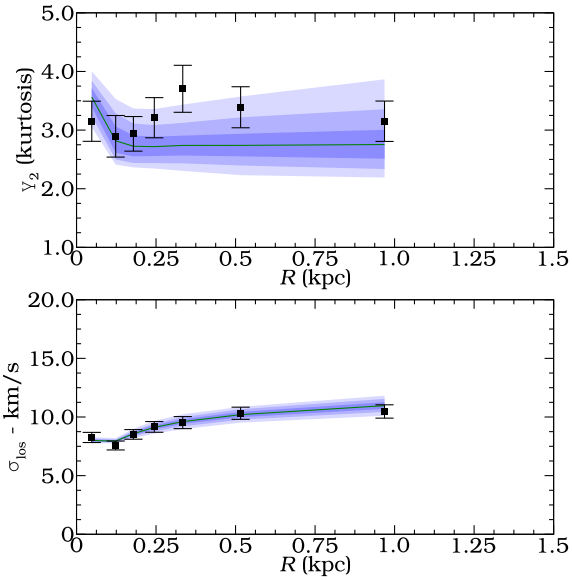


Figure 2.7: Line-of-sight velocity dispersion profile obtained in radial bins for the data by Battaglia et al. (2008) and Walker et al. (2009a) of Sculptor, taking into account the foreground contamination by the Milky Way. The dashed curve corresponds to the (pdf weighted) median line-of-sight velocity dispersion profile from the Schwarzschild models presented in Sec. 4.2, while the contours indicate the 1, 2, and 3σ uncertainties around this curve. The last bin extends to 1.3 kpc.

velocity dispersion. Furthermore, if there is any bias in the sampling of the kinematic data (which usually is the case), it will affect both the Sculptor data and the foreground data in equal ways, and will cancel out in the ratio. The term $p(v_i|\neg m, \sigma)$ is the weighted sum of two Gaussians as described in Battaglia et al. (2008).

For each radial bin we find the maximum likelihood value for the velocity dispersion. After this, we perform a 3σ clipping around the mean, and estimate the second and fourth moments for the remaining stars using Eqs. (2.16) and (2.18). The errors are computed from Eqs. (2.19) and (2.20). The final sample contains 1696 member stars. Fig. 2.7 shows the resulting velocity dispersion profile and the kurtosis ($\hat{\mu}_4/\hat{\mu}_2^2$). The line-of-sight velocity dispersion is well-constrained, it is relatively flat although it appears to be slightly rising with radius. The kurtosis has larger error bars, and this implies that additional modelling is required to establish in a robust statistical way what the shape of the velocity ellipsoid is (Gerhard 1993).

2.4.2 Schwarzschild method applied to Sculptor

We now apply the Schwarzschild method to the data from the Sculptor dSph and model this galaxy as a (non-rotating) spherically symmetric system. For the light distribution we assume a Plummer profile with scale radius $b = 0.3$ kpc (Battaglia 2007).

We first assume that Sculptor is embedded in an NFW dark matter halo, as we did for

mock Sculptor in §2.3.1. The results of this modelling are shown in Fig. 2.8. We obtain a tight constraint on the enclosed dark matter mass of $\mathrm{M_{1kpc}} = 1.03 \times 10^{8.00\pm0.03}\mathrm{M_\odot}$ (7% uncertainty, or $\mathrm{M_{1kpc}} = 1.03^{+0.075}_{-0.070} \times 10^8\mathrm{M_\odot}$). The scale radius at $r_s = 2.15 \times 10^{\pm0.25}$ kpc (76% uncertainty, or $r_s = 2.15^{+1.6}_{-0.93}$ kpc) is less well constrained, similar to what we find for mock Sculptor. In comparison to our mock model, the Sculptor dwarf galaxy would seem to have a larger scale radius (see Fig. 2.4).

Our estimates are consistent with those derived in previous work for the NFW family of mass models. For example, Walker et al. (2009b, 2010) derive a mass of $10^{+3.2}_{-5.0}\times10^7\mathrm{M_\odot}$ within 1.1 kpc, while we estimate $10^{+1.3}_{-1.2}\times10^7\mathrm{M_\odot}$ within the same distance with smaller error bars. On the other hand, Battaglia et al. (2008) obtained a mass of $2.2^{+1.0}_{-0.7}\times10^8\mathrm{M_\odot}$ within 1.8 kpc, while our measurement at this radius is $1.9^{+0.4}_{-0.3} \times 10^8\mathrm{M_\odot}$. The mass estimates by Strigari et al. (2008) Walker et al. (2010, MCMC value) and Wolf et al. (2010) are over plotted in the bottom right panel of Fig. 2.8, and all three agree very well with ours and are within the confidence regions.

The top right panel of Fig. 2.8 shows that Sculptor's anisotropy is mostly tangential and fairly constant with radius, except near the centre where it becomes slightly more isotropic (even after talking into account our limitations due to the projection effects shown and discussed in the context of Fig. 2.3). This anisotropy profile at $r > 0.1$ kpc is consistent with the constant anisotropy assumed in Jeans models of Sculptor, as by Walker et al. (2007), who find $\beta = -0.5$.

We plot the joint pdf of $\mathrm{M_{1kpc}}$ and r_s again in Fig. 2.9. In the left panel we plot lines of constant virial mass M_{200} in blue[7], with the blue dotted line indicating a value of $\log \mathrm{M_{200}} = 8.5$, increasing with steps of 0.5 dex until $\log \mathrm{M_{200}} = 10.5$. Orange lines indicate constant concentration values, with the orange dashed line corresponding to $c = 10$, increasing with steps of 5 until $c = 40$. This shows that the concentration of Sculptor is $\sim 15 \pm 6$ and that the virial mass is not well determined (not better than within factor of 100 at a 3σ level uncertainty).

Cosmological N-body simulations of dark matter have shown that there is a relation between the concentration of dark matter halos and their virial masses, the so called mass-concentration relation (e.g. Bullock et al. 2001b). In the right panel of the Fig. 2.9, we show as the dashed black line the mass-concentration relation of Macciò et al. (2007):

$$\log c_{200} = -0.109 \log(M_{200}/\mathrm{M_\odot}) + 2.34. \tag{2.32}$$

Judging solely from this relationship this would suggest that Sculptor is not compatible with the current ΛCDM cosmology. If we however plot the intrinsic scatter of $\sigma_{\ln c_{200}} = 0.33$ in the same panel (solid red lines, 1,2 and 3σ contours) we see that Sculptor lies well within the 1 and 2σ contours. We can also use the mass-concentration relation as a prior in our models. The results are shown as the green contours in this figure and they are slightly smaller than the original contours. The effect is small, but leads to a narrowing down of the possible values for r_s.

[7] M_{200} is the virial mass (mass enclosed within r_{200}), where r_{200} is the distance at which the average density of a dark matter halo is 200 times the cosmological density ρ_c (e.g. Binney & Tremaine 2008, §2.2).

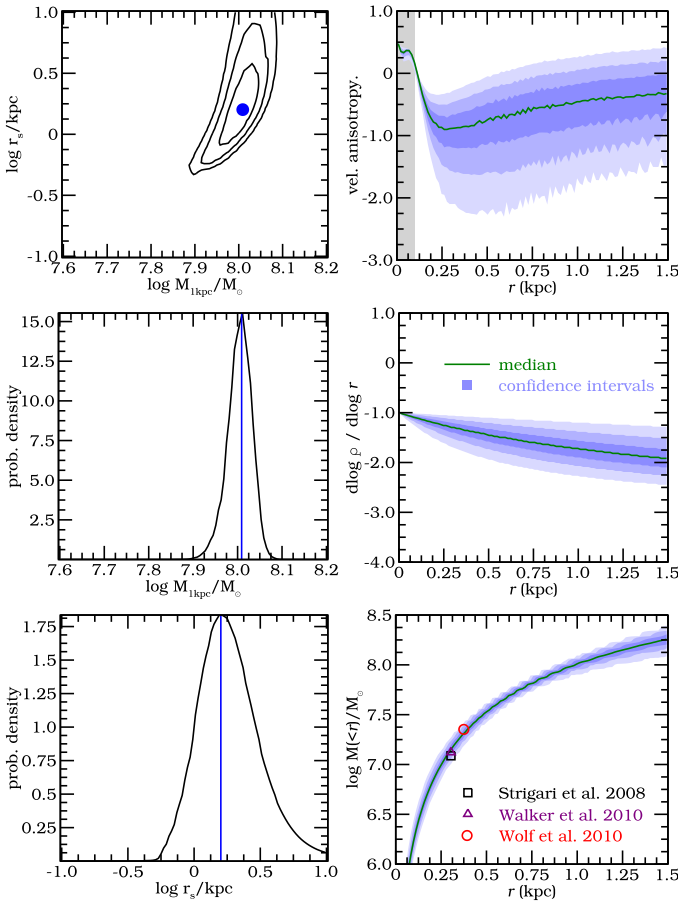


Figure 2.8: Left column: Probability density functions (joint and marginalised) for mass and scale parameters of the NFW dark matter halo potential recovered for Sculptor. Blue dot and blue lines (left column) indicate the maximum likelihood value (of the unmarginalised pdf). The green solid line indicates the median value and the blue regions (or black contour lines in the top left panel) the 68.3, 95.4 and 99.7 per cent confidence intervals. **Top right:** Recovered anisotropy profile. The grey region indicates where we cannot recover the anisotropy. **Middle right:** Recovered logarithmic density slope (see text) for the dark matter. **Bottom right:** Recovered enclosed mass profile.

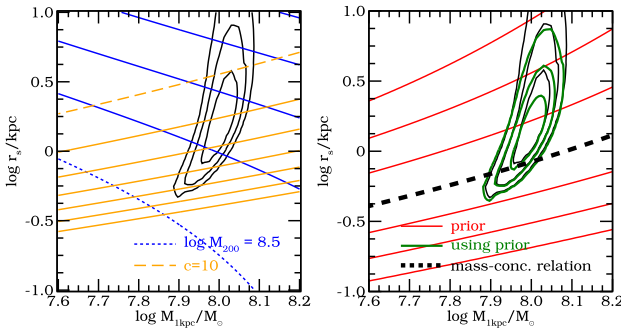


Figure 2.9: Left: The black contours correspond to the same pdf as that shown in the bottom right panel of Fig. 2.8. Blue lines indicate curves of constant M_{200}, with the blue dotted line corresponding to a value of $\log M_{200} = 8.5$, increasing with steps of 0.5 dex until $\log M_{200} = 10.5$. Orange lines indicate values of constant concentration, with the orange dashed line corresponding to $c = 10$, increasing with steps of 5 until $c = 40$. **Right:** Red contour lines indicate the cosmologically motivated prior, with the black dashed line the mean value. The green contours are the pdf obtained using this prior for Sculptor.

2.4.3 Dark matter inner density profile

We now consider a more general dark matter profile for the dark matter halo of Sculptor as we did for our mock models in §2.3.2 by allowing in the inner slope α to vary (see Eq. 2.29). The results are shown in Fig. 2.10.

This figure shows that the maximum likelihood value for $M_{1\mathrm{kpc}}$ and that the velocity anisotropy recovered by the Schwarzschild method are in very good agreement with the values obtained when α is fixed to -1 as in Fig. 2.8. However, as discussed in Sec. 2.3.2 the strong degeneracy between r_s and α implies that the scale radius is less well determined.

The middle right panel of Fig. 2.10 shows that the distribution of values for the inner slope α is very broad. Nonetheless it is clear that very steep cuspy profiles ($\alpha < -1.5$) are excluded. The maximum likelihood value is reached for a cored profile ($\alpha = 0$), although this is statistically indistinguishable from slightly cuspier slopes as evidenced by the pdfs in this Figure. The bottom right panel of Fig. 2.10 shows that at a distance of 250 pc (where the anisotropy profile begins to change its shape, and which according to our tests in Sec. 2.3.2 is the inner most point where it is reliably determined) the median logarithmic slope profile (green line) takes a value of ~ -1.25, which is larger that found in our mock Sculptor model (~ -1.75). Since the maximum likelihood value of r_s estimated by the Schwarzschild method is not very different from that assumed in mock Sculptor, this comparison would suggest that the density profile of Sculptor is shallower than NFW, although the uncertainties are still too large to make a very firm statement.

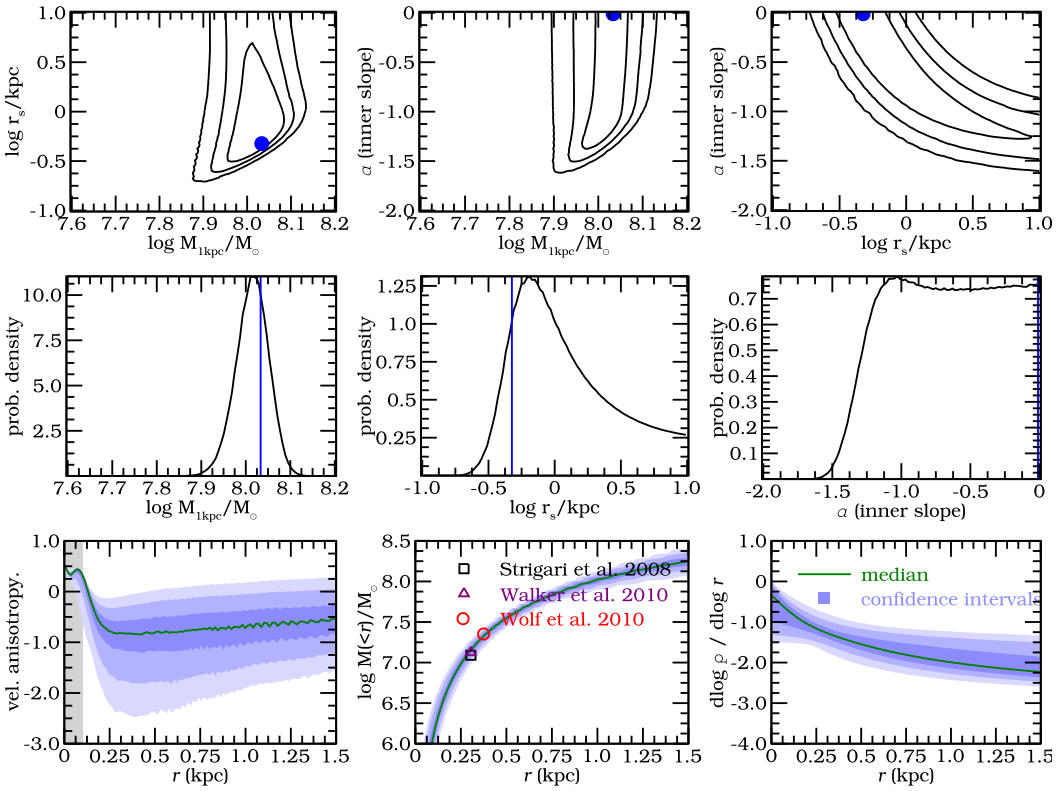


Figure 2.10: Top two rows: Probability density functions (joint and marginalised) for mass, scale and inner slope parameters of the dark matter halo potential recovered for Sculptor. Blue dots (top row) and blue lines (middle row) indicate the maximum likelihood value (of the unmarginalised pdf). The green solid line indicates the median value and the blue regions (or black contour lines in the top row) the 68.3, 95.4 and 99.7 per cent confidence intervals. **Bottom left:** Recovered anisotropy profile. The grey region indicates where we cannot recover the anisotropy. **Bottom centre:** Recovered enclosed mass profile. **Bottom right:** Recovered logarithmic density slope (see text) for the dark matter.

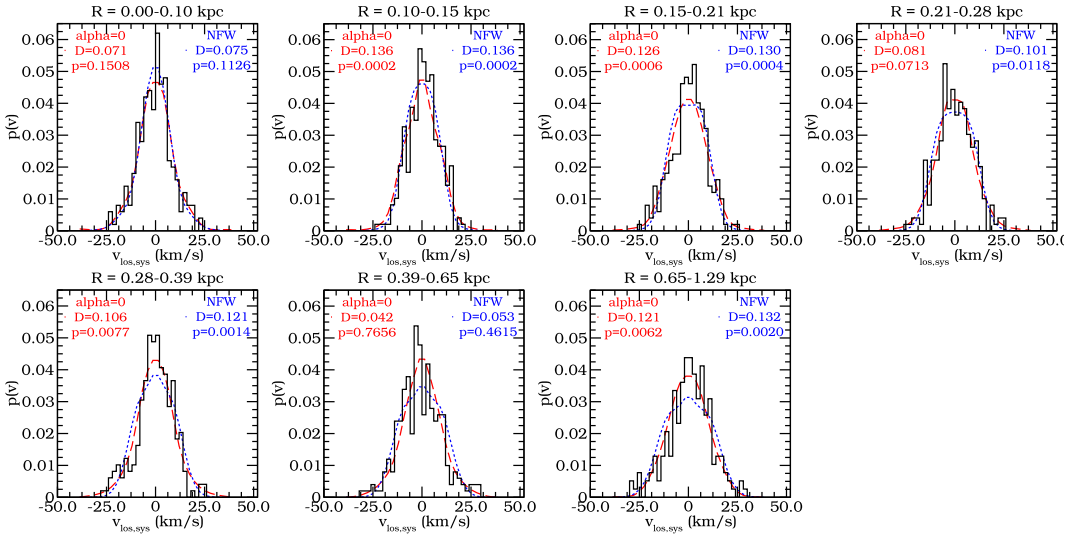


Figure 2.11: Line-of-sight velocity distributions for Sculptor stars for different radial bins (black histogram). The red curve corresponds to the best-fit $\alpha = 0$ model, while the best-fit NFW model ($\alpha = -1$) is shown in blue. The p-values correspond to the probability that the observed and best-fit model are drawn from the same parent distribution, as quantfied by a KS-test.

2.5 Discussion

Our results agree with previous studies of Sculptor that a central logarithmic slope $\alpha = 0$ is more likely than the NFW $\alpha = -1$ cusp (Battaglia et al. 2008; Walker & Peñarrubia 2011; Amorisco & Evans 2012a; Agnello & Evans 2012), although in our case the evidence is clearly not strong enough to rule out the latter. Note however, that many, though not all, of these works have tested the presence of a true core, namely $d\rho/dr = 0$ and not just $d \log \rho/dr = \alpha = 0$ at the centre. A comparison of the statistical significance of our results with Walker & Peñarrubia (2011) or Agnello & Evans (2012) is not straightforward because of the very different methods employed to estimate the inner slope. These authors use the existence of two distinct populations (metal-rich and metal-poor) to constrain the mass distribution (which is modelled non-parametrically), and this may or may not be the cause of the difference. Amorisco & Evans (2012a) favour a cored profile over an NFW with a high significance, but their conclusion is based on the assumption that two populations follow Michie-King phase-space distribution functions, which are radially anisotropic. In Amorisco & Evans (2012b) these authors present evidence that the velocity anisotropy of Sculptor might in fact radial. This is in conflict with our results, since we find, with high confidence levels, that the orbits of stars in Sculptor are tangentially biased (also when marginalised over all models), especially at radii beyond 250 pc, where the dominant population is the metal-poor one. Furthermore, also Walker et al. (2007); Battaglia et al. (2008) and Łokas (2009) favour a tangentially biased constant anisotropy profile in their Jeans models of this system.

Given these seemingly contradictory results, it is worthwhile taking a closer look

at the line-of-sight velocity distributions to understand where the discrepancies might arise. Figure 2.11 shows these distributions (black histograms) together with the results obtained for the best fit NFW (blue dotted) and $\alpha = 0$ (red dashed) models. This figure shows that the l.o.s. velocity distribution is more peaked at small radii than in the outskirts, where it is more flat-topped. This is consistent with our measurements of the l.o.s. kurtosis, and also with our derived anisotropy profile. As shown by Dejonghe (1987), systems with a tangentially biased velocity ellipsoid have a flat-topped l.o.s. velocity distribution only at large radii, while in the centre, this distribution is always more peaked. This is because the l.o.s. towards the centre has contributions from stars located at a larger range of radii, and hence also from radial plunging orbits, which drives the shape of the projected velocity distribution to be more peaked. This is known as the "complementarity property", and the results of our modeling would be consistent with such a scenario.

Figure 2.11 also shows the small differences between the $\alpha = 0$ and $\alpha = -1$ profiles, and lend support to our conclusion that the two profiles are both relatively good representations of the data. This is quantified by a KS-test, whose probabilities are indicated in the corners of each of the panels of this figure. In a few of the radial bins, none of the models fair particularly well. The $\alpha = 0$ model tends to fit better the peak of the histogram, and this could be partly to the lower but still tangential anisotropy since $\beta \sim -0.3$ for most radii.

The question thus arises as why do Amorisco & Evans (2012b) find a radial anisotropy. Just like us, these authors have utilized the W09 dataset. However, they use stars with a membership probability of 0.5 (as estimated by W09), and do not model the contamination by the Milky Way any further. In the presence of contaminants, l.o.s. velocity distributions have extended wings, and this produces a peaky distribution akin that of truly radially anisotropic systems. We have tested this idea by measuring the kurtosis for two different membership probability values $p = 0.5$, as in Amorisco & Evans (2012b), and $p = 0.9$ (which is more in line with our more sophisticated modeling of the foreground), and found a significant difference: the kurtosis is > 3 in the first case while in the second case it is consistent with that shown in Fig. 2.6 of this chapter.

In conclusion, care is required when contamination is present, and the differences between profiles that have $\alpha = -1$ such as the NFW or $\alpha = 0$, although present, are perhaps not as dramatic as maintained in other published work.

2.6 Conclusions

We have presented a spherically symmetric dynamical model for the Sculptor dwarf spheroidal galaxy using the Schwarzschild orbit superposition method. This method fits a set of observables, which in our case are the light, the second and fourth moments of the line of sight velocity distribution. We have tested this method on a mock model for the Sculptor dSph galaxy embedded in an NFW profile, and generated with similar sampling and velocity errors as the data currently available for this system.

In our tests we have found our method to give precise (7% uncertainty) and accurate estimates for the mass within 1 kpc, when assuming that the underlying gravitational potential is of NFW form. However the scale radius is recovered less precisely (37% uncertainty) for data sets containing $\sim$ 2000 member stars. We have also explored a more general model for the dark matter halo and found that we are able to measure the

logarithmic slope of its density profile, although the central value is weakly constrained. Nonetheless we find that the maximum likelihood value for the inner slope is very close to the input value.

We then used the Schwarzschild method on Sculptor after having estimated the second and fourth line of sight velocity moments for this galaxy. Assuming an NFW profile for the dark matter profile, we derive a mass within 1 kpc of $M_{1kpc} = (1.03 \pm 0.07) \times 10^8$ $M_\odot$, and find the concentration ($c \sim 15$) to be compatible with current ΛCDM predictions, given the expected scatter in the mass-concentration relation (Macciò et al. 2007). When we try to constrain the inner slope of the dark matter density profile of Sculptor, we can exclude very cuspy profiles ($\alpha < -1.5$). However, given the current data set, our method does not seem to be able to discriminate in a statistically significant way between a $\alpha = -1$ cusp and a central logarithmic slope $\alpha = 0$, although the latter is the most likely value. We are, however, able to determine that the logarithmic slope of the density profile falls off to the value of -2 at a distance of ~ 1 kpc.

The Schwarzschild method is also able to derive the velocity anisotropy profile, except near the centre where we are limited by the number of tracers. For Sculptor we find this to be tangentially biased with a hint that it may become more isotropic for $r \lesssim 250$ pc. This result is nearly independent of the assumed shape of the dark matter density profile, whether NFW or its generalised form. This nearly flat tangentially anisotropic ellipsoid should hold clues to the formation and dynamical evolution of Sculptor but it is as yet unclear whether a model exists that can reproduce this trend.

Models in which stars follow the dark matter are inconsistent with our results, as they predict a more radially anisotropic velocity ellipsoid (Diemand et al. 2004). On the other hand, the tidal stirring of a disky galaxy (see e.g. Mayer 2010), can lead to a tangentially biased ellipsoid. However, this model predicts that the ellipsoid becomes increasingly tangential with radius as a consequence also of tidal stripping, and this is not what we derive at face value.

Schwarzschild modelling does not have to assume a parametric form for the velocity anisotropy as for instance in the commonly used Jeans modelling. We therefore believe that we are less affected by biases due to assumptions compared to such class of models. Furthermore, by construction we are guaranteed that our models are physical in the sense of having non-negative distribution functions.

We plan to develop the Schwarzschild method further to work with the full line of sight velocity distribution, instead of binning the data and comparing it the the velocity moments profile. Avoiding the loss of information when binning, we expect that this may give us better estimates for the inner slope and the anisotropy profile. Also, since neither Sculptor nor any of the other dwarf spheroidal galaxies are spherical, we are developing a non spherical orbit-based dynamical model. We also plan to apply this modelling to other dwarf spheroidal galaxies such as Fornax, Carina and Sextans in future work.

Acknowledgements

AH and MB are grateful to NOVA for financial support. AH acknowledges financial support the European Research Council under ERC-Starting Grant GALACTICA-240271. We are also grateful to Carlos Vera-Ciro and Simon White for various interesting discussions.

Bibliography

Agnello, A., & Evans, N. W. 2012, ApJ, 754, L39

Amorisco, N. C., & Evans, N. W. 2012a, MNRAS, 419, 184

—. 2012b, MNRAS, 424, 1899

An, J. H., & Evans, N. W. 2009, ApJ, 701, 1500

Avila-Reese, V., Colín, P., Valenzuela, O., D'Onghia, E., & Firmani, C. 2001, ApJ, 559, 516

Battaglia, G. 2007, PhD thesis, Kapteyn Astronomical Institute, University of Groningen

Battaglia, G., et al. 2005, MNRAS, 364, 433

—. 2006, MNRAS, 370, 1055

Battaglia, G., Helmi, A., Tolstoy, E., Irwin, M., Hill, V., & Jablonka, P. 2008, ApJ, 681, L13

Battaglia, G., & Starkenburg, E. 2012, A&A, 539, A123

Binney, J., & Tremaine, S. 2008, Galactic Dynamics (Princeton University Press)

Bullock, J. S., Dekel, A., Kolatt, T. S., Kravtsov, A. V., Klypin, A. A., Porciani, C., & Primack, J. R. 2001a, ApJ, 555, 240

Bullock, J. S., Kolatt, T. S., Sigad, Y., Somerville, R. S., Kravtsov, A. V., Klypin, A. A., Primack, J. R., & Dekel, A. 2001b, MNRAS, 321, 559

Chanamé, J., Kleyna, J., & van der Marel, R. 2008, ApJ, 682, 841

Ciotti, L., & Morganti, L. 2010, MNRAS, 408, 1070

Collins, M. L. M., et al. 2010, MNRAS, 407, 2411

Cretton, N., de Zeeuw, P. T., van der Marel, R. P., & Rix, H. 1999, ApJS, 124, 383

de Blok, W. J. G. 2010, Advances in Astronomy, 2010

Dehnen, W., & Binney, J. J. 1998, MNRAS, 298, 387

Dejonghe, H. 1987, MNRAS, 224, 13

Diemand, J., Moore, B., & Stadel, J. 2004, MNRAS, 352, 535

Gebhardt, K., et al. 2007, ApJ, 671, 1321

Gebhardt, K., & Thomas, J. 2009, ApJ, 700, 1690

Gerhard, O. E. 1993, MNRAS, 265, 213

Irwin, M., & Hatzidimitriou, D. 1995, MNRAS, 277, 1354

Jardel, J. R., & Gebhardt, K. 2012, ApJ, 746, 89

Jardel, J. R., Gebhardt, K., Fabricius, M., Drory, N., & Williams, M. J. 2012, ArXiv e-prints

Jeans, J. H. 1915, MNRAS, 76, 70

Kalirai, J. S., et al. 2010, ApJ, 711, 671

Kleyna, J., Wilkinson, M. I., Evans, N. W., Gilmore, G., & Frayn, C. 2002, MNRAS, 330, 792

Kleyna, J. T., Wilkinson, M. I., Evans, N. W., & Gilmore, G. 2001, ApJ, 563, L115

Łokas, E. L. 2002, MNRAS, 333, 697

—. 2009, MNRAS, 394, L102

Łokas, E. L., Mamon, G. A., & Prada, F. 2005, MNRAS, 363, 918

Long, R. J., & Mao, S. 2010, MNRAS, 405, 301

Macciò, A. V., Dutton, A. A., van den Bosch, F. C., Moore, B., Potter, D., & Stadel, J. 2007, MNRAS, 378, 55

Magorrian, J. 2006, MNRAS, 373, 425

Mateo, M., Olszewski, E., Welch, D. L., Fischer, P., & Kunkel, W. 1991, AJ, 102, 914

Mateo, M. L. 1998, ARA&A, 36, 435

Mayer, L. 2010, Advances in Astronomy, 2010

Merritt, D., & Tremblay, B. 1993, AJ, 106, 2229

Navarro, J. F., et al. 2010, MNRAS, 402, 21

Pepper, D., & Heinrich, J. 1992, The finite element method: basic concepts and applications (Hemisphere Pub)

Pfenniger, D. 1984, A&A, 141, 171

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

Richstone, D. O., & Tremaine, S. 1984, ApJ, 286, 27

Rix, H., de Zeeuw, P. T., Cretton, N., van der Marel, R. P., & Carollo, C. M. 1997, ApJ, 488, 702

Robin, A. C., Reylé, C., Derrière, S., & Picaud, S. 2003, A&A, 409, 523

Romanowsky, A. J., Douglas, N. G., Arnaboldi, M., Kuijken, K., Merrifield, M. R., Napolitano, N. R., Capaccioli, M., & Freeman, K. C. 2003, Science, 301, 1696

Schwarzschild, M. 1979, ApJ, 232, 236

Schweitzer, A. E., Cudworth, K. M., Majewski, S. R., & Suntzeff, N. B. 1995, AJ, 110, 2747

Spergel, D. N., & Steinhardt, P. J. 2000, Physical Review Letters, 84, 3760

Springel, V. 2005, MNRAS, 364, 1105

Strigari, L. E., Bullock, J. S., Kaplinghat, M., Simon, J. D., Geha, M., Willman, B., & Walker, M. G. 2008, Nature, 454, 1096

Syer, D., & Tremaine, S. 1996, MNRAS, 282, 223

Valluri, M., Merritt, D., & Emsellem, E. 2004, ApJ, 602, 66

van den Bosch, R. C. E., van de Ven, G., Verolme, E. K., Cappellari, M., & de Zeeuw, P. T. 2008, MNRAS, 385, 647

van der Marel, R. P., Cretton, N., de Zeeuw, P. T., & Rix, H.-W. 1998, ApJ, 493, 613

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

—. 2009a, AJ, 137, 3100

Walker, M. G., Mateo, M., Olszewski, E. W., Gnedin, O. Y., Wang, X., Sen, B., & Woodroofe, M. 2007, ApJ, 667, L53

Walker, M. G., Mateo, M., Olszewski, E. W., Peñarrubia, J., Wyn Evans, N., & Gilmore, G. 2009b, ApJ, 704, 1274

—. 2010, ApJ, 710, 886

Walker, M. G., Mateo, M., Olszewski, E. W., Sen, B., & Woodroofe, M. 2009c, AJ, 137, 3109

Walker, M. G., & Peñarrubia, J. 2011, ApJ, 742, 20

Wilkinson, M. I., Kleyna, J., Evans, N. W., & Gilmore, G. 2002a, MNRAS, 330, 778

—. 2002b, MNRAS, 330, 778

Wolf, J., Martinez, G. D., Bullock, J. S., Kaplinghat, M., Geha, M., Muñoz, R. R., Simon, J. D., & Avedo, F. F. 2010, MNRAS, 406, 1220

Wu, X., & Tremaine, S. 2006, ApJ, 643, 210

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

Appendix 2.A Numerical approximation to the distribution function

We take the following separable form for the distribution function:

$$f(E,L) \quad = \quad f_E(E) f_L(L). \tag{2.33}$$

For a constant anisotropy for instance, $f_L(L) \propto L^{-2\beta}$. Now we assume that the distribution function $f(E,L)$ can be approximated by $\hat{f}(E,L)$, where $\hat{f}_E(E)$ is a sum of delta functions, such that:

$$\hat{f}(E,L) \quad = \quad \frac{1}{N}\sum_{i=1}^{N} w_i \delta(E-E_i) f_L(L). \tag{2.34}$$

The density distribution corresponding to this distribution function is:

$$\begin{aligned}
\hat{\nu}(r) &= 2\pi \int_{-v_{r,\max}}^{v_{r,\max}} dv_r \int_0^{v_{t,\max}} v_t dv_t \hat{f}(E,L) && (2.35)\\
&= \frac{4\pi}{r^2}\int_0^{-\Phi(r)} dE \int_0^{L_{\max}} dL L \frac{\hat{f}(E,L),}{\sqrt{-2(E-\Phi(r))-\frac{L^2}{r^2}}} && (2.36)\\
&= \frac{4\pi}{r^2}\frac{1}{N}\sum_{i=1}^{N} w_i \int_0^{L_{\max}} dL L \frac{f_L(L)}{\sqrt{-2(E_i-\Phi(r))-\frac{L^2}{r^2}}} && (2.37)\\
&\quad \times\Theta(-(E_i-\Phi(r))) && (2.38)\\
&= \frac{1}{N}\sum_{i=1}^{N} w_i \hat{\nu}_i(r), && (2.39)
\end{aligned}$$

where Θ is the Heaviside step function and the $\hat{\nu}_i(r)$ the densities that correspond to the each of the energy delta functions.

Given a stellar density distribution $\nu(r)$ and a gravitational potential $\Phi(r)$, it may be possible to find the weights w_i such that $\nu(r) \approx \hat{\nu}(r)$. In this case we may state that we have found a numerical approximation to the distribution function that generates the proper stellar density distribution and is embedded in the potential $\Phi(r)$. A solution can be found for instance using a non-negative least square method. An even simpler method is to start with the $\hat{\nu}_j$ corresponding to the lowest binding energy. All $\hat{\nu}_i$ associated with higher binding energies can only contribute to the density at smaller radii, therefore by weighing $\hat{\nu}_j$ this can account for the density out to the outermost radius. Now one can proceed with the next $\hat{\nu}_i$. Thus we start from the lowest binding energy components, use appropriate weights and build the density distribution from outside in. Care should be taken to make sure all weights are positive.

In the case of the mock Sculptor model discussed in the main text, $\nu(r)$ is the Plummer profile and $\Phi(r)$ is the sum of the potentials of the Plummer mass distribution describing the stellar component and that generated by the NFW profile associated to the dark halo. In this case we have chosen $f_L(L) \propto L^{-2\beta}$, where $\beta = -1$. For our purpose we choose a logarithmically spaced radial grid of 600 points between $r_{\min} = 10^{-3}$ kpc and

$r_{\max} = 10^3$ kpc. For each r_i on the grid, we calculate the potential energy, giving us a grid of energies, which we take the energies for our distribution function (E_i in Eq. 2.34). For each E_i we calculate the density on the same radial grid. The last step is to find the weights w_i using the above procedure. A small mismatch (few %) of the density at large radii (> 300 pc) occurs due to the distribution function missing lower binding energy components. The cumulative mass distribution of the stellar mass deviates $< 10^{-4}$ from the true mass distribution, and within 300 pc the relative density deviates $< 2 \times 10^{-4}$. Outside this radius the density does not match very well, but since this is at large radii and its mass contribution is very small (note also that the cumulative mass distribution shows only small deviations) this is of no importance.

Appendix 2.B Centre of mass velocity of Sculptor

In this Appendix we transform the observed line-of-sight velocities to velocities with respect to the centre of mass of Sculptor. This requires knowledge of the latter, which is what we derive here using a maximum likelihood method.

The observed (heliocentric) line-of-sight velocity of a star can be expressed as:

$$\begin{aligned} v_{*,hel}(l,b) &= \boldsymbol{e}_{los}(l,b) \cdot (\boldsymbol{v}_{*,Scl}(l,b) + \boldsymbol{v}_{Scl,GSR} - \boldsymbol{v}_{\odot,GSR}) \\ &= v_{*,Scl}(l,b) + v_{Scl,GSR}(l,b) - v_{\odot,GSR}(l,b), \end{aligned}$$

where $\boldsymbol{e}_{los}(l,b)$ is the line-of-sight unit vector in the direction of the star, $\boldsymbol{v}_{*,Scl}(l,b)$ the velocity of the star with respect to the centre of mass of Sculptor, $\boldsymbol{v}_{Scl,GSR}$ the systemic velocity of the centre of mass of Sculptor with respect to the Galactic Standard of Rest (hereafter GSR), $v_{\odot,GSR}$ the velocity of the Sun with respect to the GSR and $\cdot$ indicates the inner product. The component of the line-of-sight velocity we are interested in is $v_{*,Scl}(l,b)$. Since $v_{*,hel}(l,b)$ is measured, and assuming we know $\boldsymbol{v}_{\odot,GSR}$, we only need to find $v_{Scl,GSR}(l,b)$. For the velocity of the Sun we use $v_{\odot,GSR} = v_{\odot,LSR} + v_{LSR,GSR} = (10.0, 5.2, 7.2) + (0, 220, 0)$ km s^{-1}, where $_{LSR}$ denotes Local Standard of Rest (Dehnen & Binney 1998).

To determine which stars are likely members of Sculptor we make a rough first selection. We take the systemic heliocentric radial velocity ($v_{\rm Scl,sys,helio} = 110.6$ km s^{-1}) and the mean velocity dispersion ($\sigma_{\rm Scl} = 10.1$ km s^{-1}) from Battaglia et al. (2008). We first require that the member stars are within 3σ of the systemic velocity of Sculptor, as indicated by the red solid lines in the right panel of Fig. 2.6. Furthermore we also require that they are located within $r < 0.944$ degree, indicated by the green dashed line in the same panel. We add this requirement since we are not confident that outside this radius a reliable velocity dispersion can be measured due to the low number density of (probable) Sculptor members compared to Milky Way stars.

For simplicity we first assume the line-of-sight velocity distribution is described by a Gaussian distribution with a constant velocity dispersion and zero mean velocity w.r.t

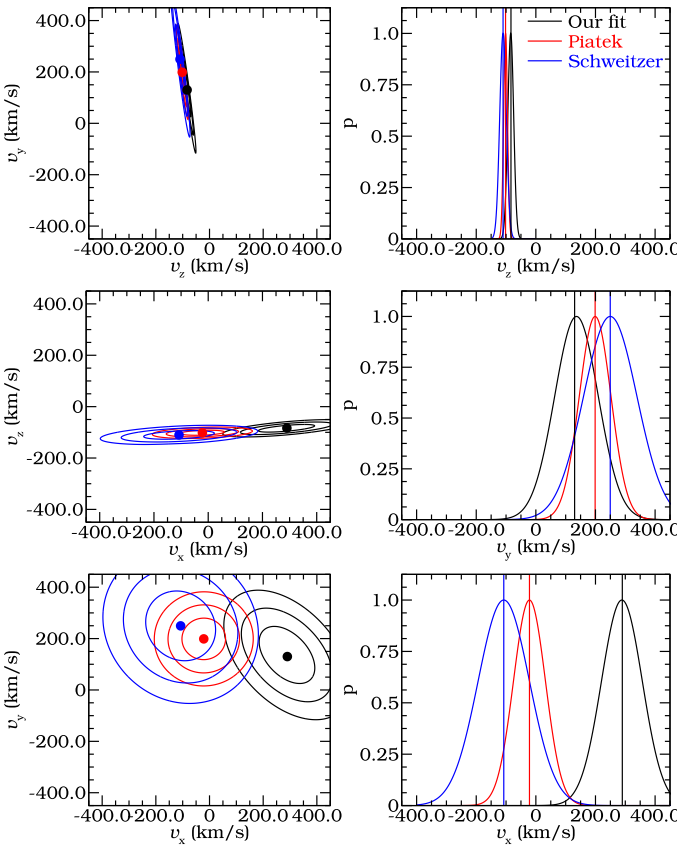


Figure 2.12: Probability distribution function (pdf) of the three velocity components of the systemic velocity of Sculptor with respect to the Galactic Standard of Rest. **Left column:** Joint pdfs with 1, 2 and 3 σ contours lines, marginalised over the other component. **Right column:** Individual pdfs marginalised over the other two components. The measurements from Piatek et al. (2006) are shown in red, while those by Schweitzer et al. (1995) are shown in blue. The vertical lines in the right panels and the dot in the left panels indicate the maximum likelihood values.

the centre of mass of Sculptor. Then the probability for $\boldsymbol{v}_{Scl,GSR}$ can be expressed as:

$$\begin{aligned} p(\boldsymbol{v}_{Scl,GSR}) &= \prod_i \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{v_{*_i,Scl}(l_i,b_i)^2}{2\sigma_i^2}\right) \\ &= \prod_i \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{1}{2\sigma_i^2}\left\{v_{*_i,hel} - \right.\right. \\ &\quad \left.\left. \boldsymbol{e}_{los}(l_i,b_i)\cdot(\boldsymbol{v}_{\odot,GSR} - \boldsymbol{v}_{Scl,GSR})\right\}^2\right] \end{aligned} \tag{2.40}$$

where $\sigma_i^2 = \sigma_{Scl}^2 + \sigma_{*_i}^2$ is the velocity dispersion of Sculptor added in quadrature with the measurement error of the velocity of star i. Although the velocity dispersion is not constant with radius, we use the global value of $\sigma_{\rm Scl} = 10.1\ \mathrm{km\,s^{-1}}$ as described previously.

The joint and marginalised probability distribution functions for the velocity components of Sculptor are plotted in Fig. 2.12 together with the 1, 2 and 3 σ contours. The maximum likelihood value is reached at $\hat{\boldsymbol{v}}_{Scl,GSR} = (v_x, v_y, v_z) = (278.5,\ 101.5,\ -81.0)\ \mathrm{km\,s^{-1}}$. These values are in agreement with Walker et al. (2008), who use a similar method. We also over plot the measurements of Piatek et al. (2006, in red) and Schweitzer et al. (1995, in blue) while the maximum likelihood value is indicated in black. Note that the uncertainty in v_z is smallest since this reflects mainly the uncertainty in the mean radial velocity of the centre of mass of Sculptor due to its high galactic latitude. The uncertainties in the other two velocity components mainly reflect the uncertainties in the proper motion measurements. Our determination of the v_y component agrees well with the various data sets, while the v_x component appears to be systematically offset. Note however, that there is overlap at the 3σ level, and the 2σ and 3σ contours for the joint v_x and v_y overlap as well. Perhaps this level of disagreement could be taken as an indication that there may be intrinsic rotation in the system. Nonetheless, we note that with this procedure we effectively have removed the observed gradient and no apparent rotation remains, whatever its origin.

Chapter 3

Model comparison of dark matter profiles in local dSphs

M. A. Breddels and A. Helmi

Submitted to A&A

Abstract

We use orbit based dynamical models to fit the 2nd and 4th moments of the line of sight velocity distributions of the Fornax, Sculptor, Carina and Sextans dwarf spheroidal galaxies. Our goal is to compare dark matter profile models of these four systems using Bayesian evidence. We consider NFW, Einasto and several cored profiles for their dark halos and present the probability distribution functions of the model parameters. When considering each system separately, we find there is no preference for one of these specific parametric density profiles. However, the combined evidence shows that is unlikely that all galaxies are embedded in the same type of cored profiles of the form $\rho_{DM} \propto 1/(1+r^2)^{\beta/2}$, where $\beta = 3, 4$. For each galaxy, we also obtain an almost model independent, and therefore accurate, constraint on the logarithmic slope of the dark matter density distribution at a radius $\sim r_{-3}$, i.e. where the logarithmic slope of the stellar density profile is -3. This is because all our best fit models essentially have the same mass distribution over a large range in radius (from just below r_{-3} to the last measured data point). This remarkable finding likely implies much stronger constraints on the characteristics that subhalos extracted from cosmological simulations should have in order to host the dSph galaxies around the Milky Way.

3.1 Introduction

According to galaxy formation theories dwarf spheroidal galaxies are believed to inhabit massive dark matter halos. Because of their large mass to light ratio these galaxies are ideal to test fundamental predictions of the ΛCDM cosmological paradigm, since it is generally considered relatively safe to neglect baryons in the construction of dynamical models.

One of the strongest predictions from ΛCDM concerns the dark matter density profile. Early simulations of dark matter halos assembled in a cosmological context showed that such a profile is accurately described by a two-sloped form, now known as NFW profile (Navarro et al. 1996, 1997). More recently Einasto profiles have been shown to provide a better fit (e.g. Springel et al. 2008; Navarro et al. 2010), in particular for satellite galaxies (Vera-Ciro et al. 2013). These predictions are made using dark matter only simulations and therefore neglect (by construction) the baryonic component. And although baryons are sub-dominant in the total potential of the system (Walker 2012a), it has been suggested that they could play a role in the evolution of dwarf spheroidal galaxies, for instance, in modifying the internal orbital structure (Bryan et al. 2012) and the overall density profile (Governato et al. 2012). The complex evolution of baryons and its non-trivial interplay with the host halo are difficult to model and not yet completely understood (see Pontzen & Governato 2012).

Another effect driving the internal dynamics of satellite halos is the tidal interaction with the main host. It can change the density profile (Hayashi et al. 2003), the geometrical shape of the mass distribution (Kuhlen et al. 2007), and also influence the kinematics of the embedded stars (Łokas et al. 2010). Unfortunately these uncertainties imply that even when the observations of the local dwarf spheroidal galaxies are not consistent with being embedded in the halos predicted from pure dark matter N-body simulations, this does not necessarily reflect a fundamental problem of ΛCDM.

Thanks to their relative proximity, information for individual stars in the dwarf galaxies satellites of the Milky Way are relatively easy to get. Sky positions are easily determined from photometry, and radial velocity measurements are possible to estimate within an error of ~ 2 km/s. Some of the datasets compiled to the date include thousands of individual members with line-of-sight velocities (Helmi et al. 2006; Battaglia et al. 2006, 2008; Walker et al. 2009a; Battaglia et al. 2011) Proper motions of individual stars are currently still too difficult to measure. Despite the fact that only three of the total of six phase space coordinates are available from measurements, it is possible to create dynamical models of these systems that can be compared to these observables.

Following the method thoroughly described in Chapter 2 we set out to model Fornax, Sculptor, Carina and Sextans with orbit-based dynamical methods (Schwarzschild modeling) assuming they are embedded in spherical halos. As extensively shown in the literature (e.g. Richstone & Tremaine 1984; Rix et al. 1997; van der Marel et al. 1998; Cretton et al. 1999; Valluri et al. 2004; van den Bosch et al. 2008; Jardel & Gebhardt 2012) this method allows to construct a non-parametric estimator of the distribution function. Among many, this method has one advantage over Jeans modeling, by not having to assume a particular velocity anisotropy profile, therefore being more general and thus less prone to biases associated to the assumptions. But even in this case there are other limitations in the modeling such as the mass-anisotropy degeneracy. In this work we use higher moments (fourth moment) of the line of sight velocity distribution

to get a better handle on this degeneracy.

To compare how different shapes for the dark matter profiles fit the data, we first need to establish a statistical framework. In this chapter we do this in a Bayesian way using the evidence (Mackay 2003). This method provides a natural way of comparing models in Bayesian inference and also makes it possible to combine the data of all the dwarf spheroidals to test for example, if all dwarf spheroidals could be embedded in a universal halo (Mateo et al. 1993; Walker et al. 2009b). Furthermore, the shape may give us hints to how the dwarf galaxy formed and the anisotropy profile may be used to distinguish between evolutionary scenarios (see e.g. Mayer 2010; Kazantzidis et al. 2011; Helmi et al. 2012).

This chapter is organized as follows. We begin in §3.2 by presenting the data and all the ingredients needed to do the model comparison. In §3.3 we present our dynamical and statistical methods. We present the results of our Schwarzschild models for the four dSph in our sample in §3.4.1, while the Bayesian model comparison is done in §3.4.2. We discuss the implications of our results in §3.4.3 and conclude in §3.5.

3.2 Data

In this section we present the data that is used for fitting our dynamical models. The radial velocity measurements of the dwarf spheroidal galaxies come from Helmi et al. (2006); Battaglia et al. (2006, 2008); Walker et al. (2009a) and Battaglia et al. (2011). We plot radius versus heliocentric velocity in Fig. 3.1 for each galaxy separately.

Figure 3.1 shows that each dSph suffers from foreground (Milky Way) contamination. To remove this contamination and reliably identify member stars we have developed a simple analytic model for the positional and kinematic distribution of both foreground and the galaxy in question (along the lines of Battaglia et al. 2008, and Chapter 4). For each particular dataset[1], we assume that the foreground has a constant surface density, and that the dSph follows a specific stellar density profile. We also assume that the velocity distribution at each radius may be modeled as sum of two Gaussians. The Gaussian describing the foreground has the same shape at all radii, while that of the stars associated with the dwarf can have a varying dispersion with radius. Their relative amplitude also changes as function of distance from the dwarf's centre. This model results in a determination of the relative contribution of member-to-non-member stars as a function of velocity and radial distance R.

Based on this model we calculate the elliptical radius at which the ratio of dSph to foreground is 3:1 (without using any velocity information). We remove all stars outside of this radius from the dataset. A particular star included in more than one dataset is removed only when it it satisfies the condition for all sets, for instance a star outside $R_{e,\mathrm{cut,\ Batt}}$, but inside $R_{e,\mathrm{cut,\ Walker}}$ will not be discarded. This simple clipping in elliptical radius cleans up part of the foreground contamination. For completeness, the radii for all datasets cleaned up in this way are presented in Table 3.2, as well as the fit to the foreground model. The number of stars and the sources are listed in Table 3.1.

For the resulting dataset, we compute the second and fourth moment of the radial velocity as a function of circular radius as follows[2]. We first define radial bins such

[1] For a given dSph there may be multiple datasets, and we treat each independently because their sampling might be different.

[2] Elliptical radii are only used for the clipping, for the rest of the analysis we use the circular radius

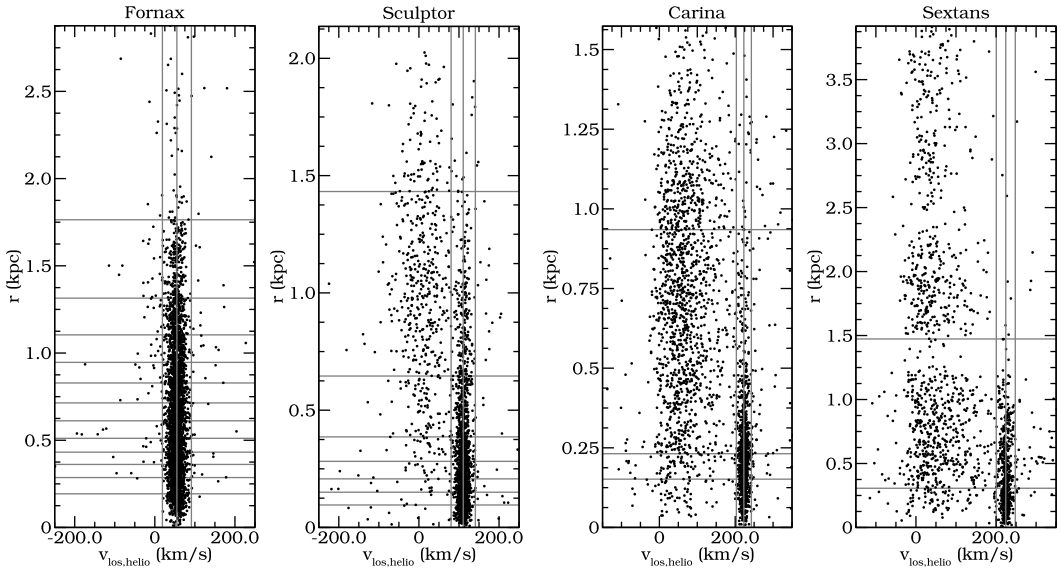


Figure 3.1: Radius versus line of sight velocity for Fornax, Sculptor Carina and Sextans. The horizontal lines show the borders of the bins, the vertical lines denote the mean systemic velocity of the galaxy together with the $\pm 3\sigma$ region.

Name	$N_{\rm Batt}$	$N_{\rm Walker}$	$N_{\rm member}$
Fornax	$945^{(1)}$	$2633^{(5)}$	2936
Sculptor	$1073^{(2)}$	$1541^{(5)}$	1685
Carina	$811^{(3)}$	$1982^{(5)}$	885
Sextans	$792^{(4)}$	$947^{(5)}$	541

Table 3.1: Number of stars in the kinematic samples used in this chapter. Sources: [(1)]Battaglia et al. (2006), [(2)]Battaglia et al. (2008), [(3)]Helmi et al. (2006); Koch et al. (2006); Starkenburg et al. (2010), [(4)]Battaglia et al. (2011), [(5)]Walker et al. (2009a)

Name	$R_{e,\rm max,Batt}$ (kpc)	$R_{e,\rm max,walker}$ (kpc)	$\mu_{\rm MW}$ (km/s)	$\sigma_{\rm MW}$ (km/s)	$\mu_{\rm dwarf}$ (km/s)	$\sigma_{\rm dwarf}$ (km/s)
Fornax	1.82	2.21	41.1	38.9	55.1	12.1
Sculptor	1.37	1.65	17.9	47.4	110.6	10.1
Carina	0.86	0.96	70.9	62.5	222.9	6.6
Sextans	1.86	1.65	67.5	74.5	224.3	7.9

Table 3.2: Parameters of the foreground plus dwarf galaxy model used for determining membership, as well as for deriving the radial profiles for the second and fourth velocity moments for each dSph.

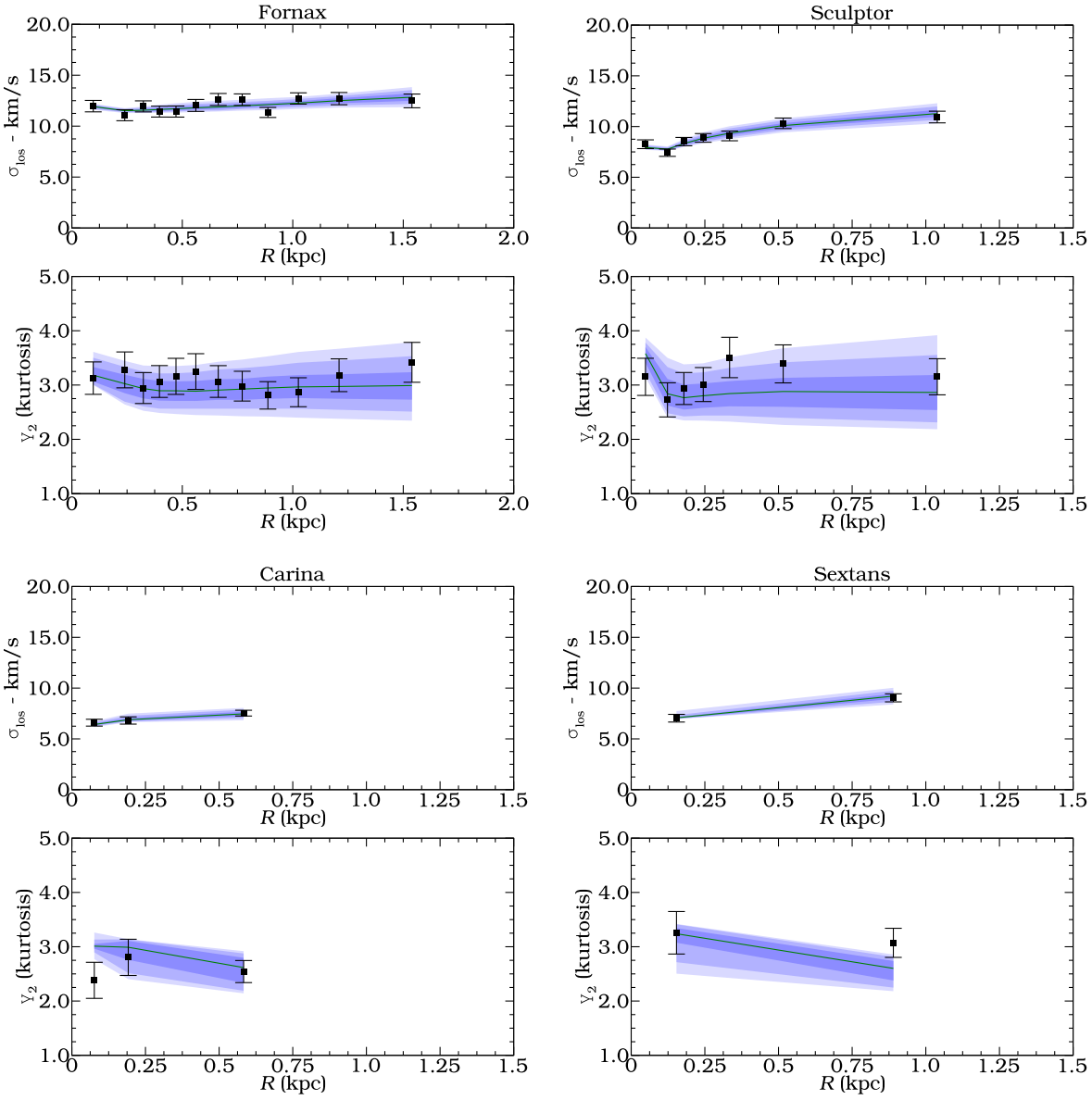


Figure 3.2: Line of sight velocity moments for Fornax, Sculptor Carina and Sextans. For each galaxy we show the velocity dispersion and the kurtosis. The black dots show the mean, and the error bars the 1σ error. The blue regions show the confidence interval for the NFW fit, similar as in Chapter 2.

Name	distance (kpc)	profile	scale radius (kpc)	L_V $\times 10^5 L_\odot$
Fornax	$138^{(1)}$	Plummer[2]	0.79	$100^{(2)}$
Sculptor	$79^{(3)}$	Plummer[3]	0.30	$10^{(3)}$
Carina	$101^{(1)}$	Exponential[4]	0.16	$2.4^{(4)}$
Sextans	$86^{(1)}$	Exponential[4]	0.39	$4.37^{(4)}$

Table 3.3: Distances, type of photometric profile used, scale radius and stellar luminosity used for the dynamic models. Sources: (1) Mateo (1998), (2) Battaglia et al. (2006), (3) Battaglia et al. (2008), (4) Irwin & Hatzidimitriou (1995)

that each bin has at least 250 stars in the velocity range $v_{\rm sys} - 3\sigma_v, v_{\rm sys} + 3\sigma_v$. If the last bin has less that 150 objects, the last two bins are merged. After this, we fit our parametric model for the galaxy plus foreground for each radial bin, to derive new velocity dispersions. Then for each bin we do a 3σ clipping on the velocity using the new velocity dispersion, and from this selection we calculate the second and fourth moments. The errors on the moments are computed using Eqs. (17) and (19) in Chapter 2. The second moment and the kurtosis[3] are shown in Fig. 3.2 for each galaxy, where the black dot corresponds to the mean, and the error bars indicate the 1σ error bar. The blue region shows the confidence interval for the NFW model found in §3.4.1.

For the photometry we use analytic fits given by various literature sources as listed in Table 3.3. Although the stellar mass is sub-dominant in the potential, we do include its contribution in the dynamic models and fix $M/L_V = 1$, as in Chapter 2.

3.3 Methods

3.3.1 Dynamical models

Our aim is to compare different models to establish what type of dark matter profile best matches the kinematical data of local dSph galaxies. Here we consider the following profiles to describe the dark matter halos of the dwarfs in our sample:

$$\rho(r) = \frac{\rho_0}{x\,(1+x)^2}, \qquad \text{NFW} \tag{3.1}$$

$$\rho(r) = \frac{\rho_0}{(1+x^\gamma)^{\beta/\gamma}}, \qquad \text{(cored) } \beta\gamma\text{-profile} \tag{3.2}$$

$$\rho(r) = \rho_0 \exp\left(-\frac{2}{\alpha'}\left(x^{\alpha'} - 1\right)\right), \qquad \text{Einasto} \tag{3.3}$$

where $x = r/r_s$ and r_s is the scale radius. Each model has at least two unknown parameters r_s and ρ_0. As we did in Chapter 2, we transform these two parameters to r_s and $\mathrm{M_{1kpc}}$ (the mass within 1 kpc). As discussed in the Introduction, the NFW and Einasto models are known to fit the halos dark matter distributions extracted from cosmological N-body simulations. On the other hand, we explore the $\beta\gamma$ models to test

[3] The kurtosis is defined as $\gamma_2 = \mu_4/\mu_2^2$, where μ_4 is the fourth and μ_2 is the second moment of the line of sight velocity distribution.

Name	Fixed parameters	Free parameters
NFW	-	M_{1kpc}, r_s
core13	$\beta = 3, \gamma = 1$	M_{1kpc}, r_s
core14	$\beta = 4, \gamma = 1$	M_{1kpc}, r_s
core23	$\beta = 3, \gamma = 2$	M_{1kpc}, r_s
core24	$\beta = 4, \gamma = 2$	M_{1kpc}, r_s
einasto.2	$\alpha' = 0.2$	M_{1kpc}, r_s
einasto.4	$\alpha' = 0.4$	M_{1kpc}, r_s

Table 3.4: Model names and their characteristic parameters of the various dark matter density profiles explored.

the possibility of a core in the dark halo. Note that, in comparison to the NFW profile, the Einasto model has one extra parameter (α'), but here we consider only two values for $\alpha' = 0.2, 0.4$ to cover the range suggested by Vera-Ciro et al. (2013). On the other hand, the $\beta\gamma$ profiles have two extra parameters, but we limit ourselves here to two different outer slopes ($\beta = 3, 4$) and two different transition speeds between the inner and the outer slopes ($\gamma = 1, 2$). Note that the $\beta\gamma$ models have a true core only for $\gamma > 1$, however in all cases the central logarithmic slope vanishes, $d \log \rho / d \log r = 0$. However, we loosely refer to these models as cored in what follows. Note that, with these choices, all of our profiles ultimately have just two free parameters. The list of models explored and their parameters are summarized in Table 3.4.

The orbit-based dynamical (Schwarzschild) models of each dwarf galaxy are obtained as follows (see Chapter 2 for a more detailed description). For each of the dark halo profiles, with its own set of parameters, we integrate a large number of orbits in the respective gravitational potential (including also the contribution of the stars). We then find a linear combination of these orbits that fits both the light and the kinematics. The orbital weights found in this way have a physical meaning and can be used to obtain the distribution function of the system. As data we have the line of sight velocity moments (second and fourth depicted in Fig. 3.2), and the light profile (Table 3.2). The best fit models (which give us the values of the parameters for a specific dark matter halo profile) are those that minimize the $\chi^2 = \chi^2_{\rm kin} + \chi^2_{\rm reg}$, under the condition that the orbital weights are positive, and that the observed light distribution is fit to better than 1% at each radius. Here $\chi^2_{\rm kin} = \sum_k (\mu_{2,k} - \mu^{model}_{2,k})^2 / {\rm var}(\mu_{2,k}) + \sum_k (\mu_{4,k} - \mu^{model}_{4,k})^2 / {\rm var}(\mu_{4,k})$. The $\chi^2_{\rm reg}$ is a regularization term to make sure that the solution for the orbit weights leads to a relatively smooth distribution function. In Chapter 2 we calibrated the amplitude of this term for Sculptor. To have the regularization term for the other dwarfs of the same relative strength, we note that $\chi^2_{\rm reg} \propto 1/N$, where N is the number of members with radial velocities, since the $\chi^2_{\rm kin}$ term also scales as $1/N$. Therefore, normalizing its amplitude to that of Sculptor we get $\chi^2_{\rm reg,\ dwarf} = \chi^2_{\rm reg,\ Scl} \times N_{\rm Scl}/N_{\rm dwarf}$.

3.3.2 Bayesian model comparison

Background on Bayesian model comparison may be found in Mackay (2003). For completeness we discuss it here briefly, but we assume the reader is familiar with the basics of Bayesian inference.

Given the data data and assuming a model M_i, the posterior for the parameters θ_i

of this model is:

$$p(\theta_i|\text{data}, M_i) = \frac{p(\text{data}|\theta_i, M_i)p(\theta_i|M_i)}{p(\text{data}|M_i)}. \tag{3.4}$$

The normalization constant $p(\text{data}|M_i)$, also called the evidence, is of little interest in parameter inference, but is useful in Bayesian model comparison. To assess the probability of a particular model given the data, we find

$$p(M_i|\text{data}) = \frac{p(\text{data}|M_i)p(M_i)}{p(\text{data})}, \tag{3.5}$$

where we see the evidence is needed. In this case $p(\text{data})$ is the uninteresting normalization constant, as it cancels out if we compare two models:

$$\frac{p(M_i|\text{data})}{p(M_j|\text{data})} = \frac{p(\text{data}|M_i)}{p(\text{data}|M_j)}\frac{p(M_i)}{p(M_j)} = B_{i,j}\frac{p(M_i)}{p(M_j)}, \tag{3.6}$$

where $B_{i,j}$ is called the Bayes factor. If we take the priors on the different models to be equal (i.e. $p(M_i) = p(M_j)$), the ratio of the evidence (the Bayes factor $B_{i,j}$) gives the odds ratio of the two models given the data data.

Using these results we can perform model comparison between dark matter density profiles, i.e. $\mathcal{M} = \{M_{\text{nfw}}, M_{\text{Einasto}}, ...\}$, and calculate for instance the odds that a given galaxy is embedded in an NFW profile compared to an Einasto model, $B_{\text{NFW,Einasto}}$.

Not only can we do model comparison on a single object, but we may also test if our objects share a particular model (e.g. they are all embedded in NFW halos). If our dataset data consists of the observations of two galaxies, i.e. $\text{data} = \text{data}_1 \cup \text{data}_2$ and assuming the datasets are uncorrelated and independent, we obtain:

$$\frac{p(M_i|\text{data})}{p(M_j|\text{data})} = \frac{p(\text{data}_1|M_i)p(\text{data}_2|M_i)}{p(\text{data}_1|M_j)p(\text{data}_2|M_j)}\frac{p(M_i)}{p(M_j)} = B_{i,j,1}B_{i,j,2}\frac{p(M_i)}{p(M_j)} \tag{3.7}$$

where each factor $p(\text{data}_k|M_i)$ should be marginalized over its (own) characteristic parameters. From Eq. (3.7) we can see that the odds ratio of the models and Bayes factor from different measurements can be multiplied to give combined evidence for a particular model.

Behind each $p(M_i|\text{data})$ is a set of orbit based dynamical (Schwarzschild) models, obtained as described above. For each of the models we calculate the evidence. Later on we compare each model's evidence to that of an NFW profile, i.e. we compute the Bayes factor $B_{i,\text{NFW}}$, where i can be e.g. Einasto. By definition $B_{\text{NFW,NFW}} = 1$, and again assuming equal priors on the different models, the Bayes factor equals the odds ratio of the models, such that for $B_{i,\text{NFW}} > 0$, model i is favored over an NFW profile.

3.4 Results

3.4.1 Schwarzschild models

As a result of our Schwarzschild modeling technique, we obtain a two dimensional probability density function (pdf) of the two parameters, $\text{M}_{1\text{kpc}}$ and r_s, for each galaxy and for each dark matter halo profile. In Fig. 3.3 we plot the pdf for the cored models and

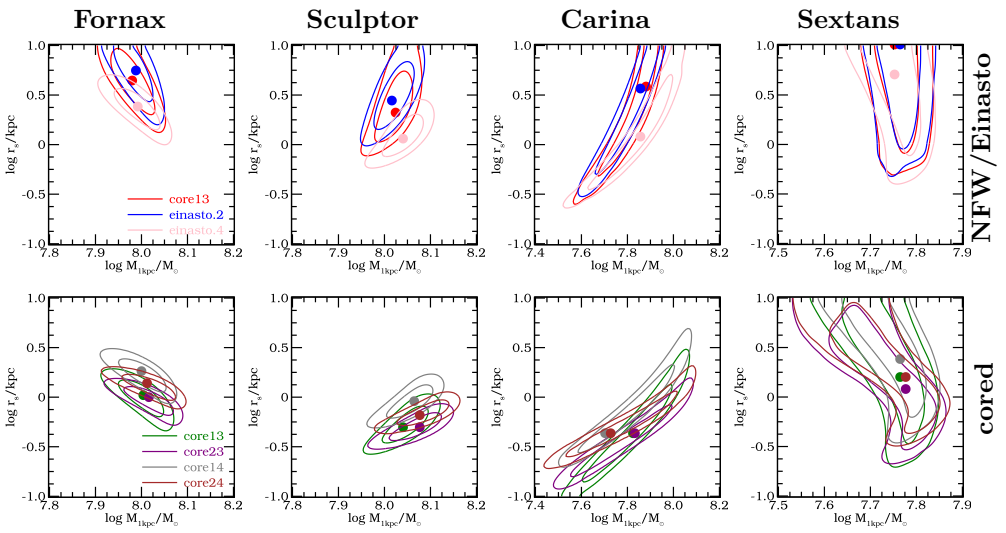


Figure 3.3: Pdf for the two free parameters characterizing the dark halo profiles for each dSph galaxy obtained using Schwarzschild modeling. The top row shows the pdfs with NFW/Einasto models, the bottom panel those for all cored models explored. The contours show the 1 and 2σ confidence levels (the 3σ contour is not shown to avoid crowding the image).

the NFW and Einasto models separately for each galaxy. The colored dots correspond to the maximum likelihood for each of the corresponding models as indicated by the legend. The contours show the 1 and 2σ equivalent confidence intervals (the 3σ contour is not shown for clarity). For both Fornax and Sculptor the parameters for all profiles are relatively well determined, while for Carina and Sextans this is less so. This can be attributed to the difference in sample size (and hence to the smaller number of members) in these systems, which has translated into fewer bins where the moments can be computed (see Fig. 3.2). In general for all four galaxies the scale radius for the cored profiles is found to be smaller than that for the NFW/Einasto profiles. We come back to this point in section 3.4.3.

Our model's masses at $r_{1/2}$, the 3d radius enclosing half of the stellar mass, are compatible with those of Wolf et al. (2010). However, our results for Fornax do not agree with those of Jardel & Gebhardt (2012). These authors prefer a cored profile with a much larger scale radius, and their enclosed mass is smaller in comparison to Wolf et al. (2010). We note that this might be partly related to the fact that the amplitude of their line of sight velocity dispersion profile (see their Fig. 2) is slightly lower than what we have determined here.

In Fig. 3.4 we overlay on the kinematic observables the predictions from the best fit Schwarzschild models. We note that all models provide very similar and virtually indistinguishable fits, especially for the 2nd moment. Some slight differences are apparent in the kurtosis, but in all cases, the differences are smaller than the error bars on the moments.

In general we find all anisotropy profiles to be roughly constant with radius and

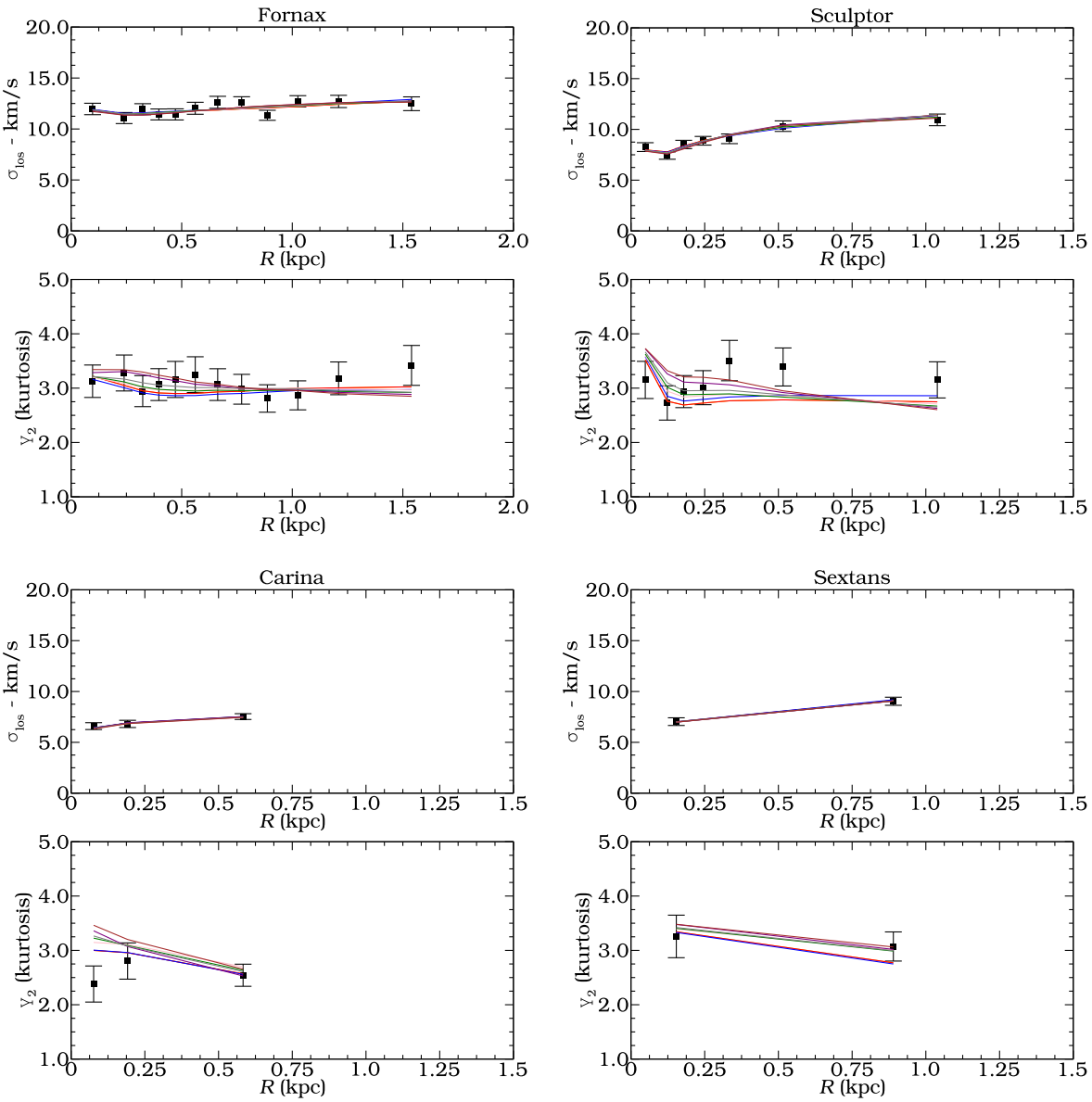


Figure 3.4: Similar to Fig. 3.2, except now we show the different best fit models for the various density profiles explored, which are indicated with different colors (the color scheme is the same as in Fig. 3.3).

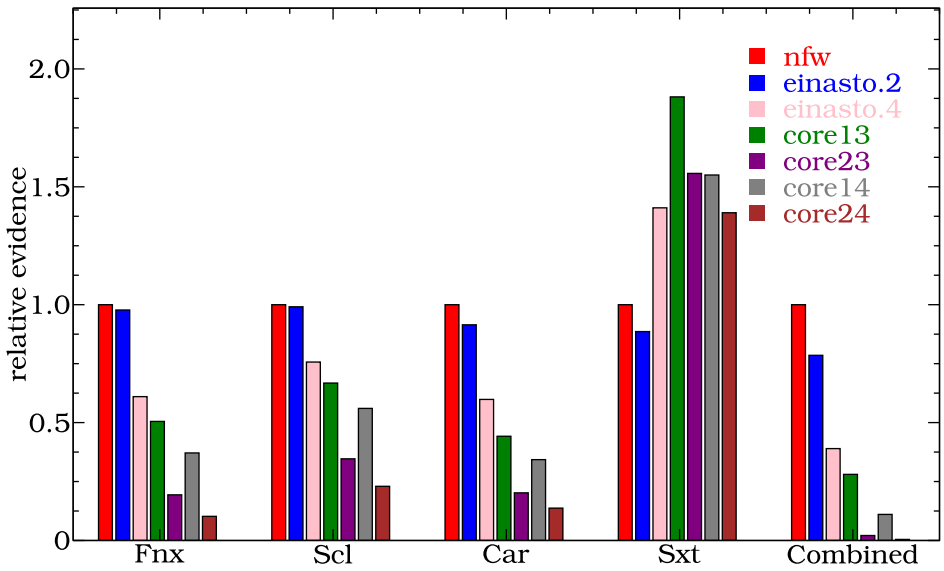


Figure 3.5: Evidences for all models listed in Table 3.4, relative to the NFW case. The last column shows the combined evidence for all galaxies together, and shows that the core23 and core24 are strongly disfavored.

slightly tangentially biased on average. We do not find significant differences between the profiles for cored and NFW models (the reason for this will become clear in §3.4.3). Fornax' s anisotropy $\beta \sim -0.2 \pm 0.2$, while Sculptor and Carina have on average $\beta \sim -0.5 \pm 0.3$. For Sextans the anisotropy cannot be determined reliably, $\beta \sim -0.3 \pm 0.5$. These values are compatible with those of Walker et al. (2007), which were derived using the spherical Jeans equation assuming a constant velocity anisotropy profile.

3.4.2 Bayesian comparison of the models

We compute the evidence relative to the NFW using Eq. (3.6) by integrating over the parameters (in our case the scale radius and the mass) the pdfs shown in Fig. 3.3. We do this for each dwarf galaxy and for all the models listed in Table 3.4. The different Bayes factors are shown in Fig. 3.5. Each set of bars shows the Bayes factors for the given dSph galaxy ($B_{i,NFW,k}$), while the last set shows the combined result ($B_{i,NFW,\mathrm{comb}} = \prod_k B_{i,NFW,k}$). We note that an odds ratio between 1:2 till 1:3 is considered "Barely worth mentioning" (Jeffreys 1998), and only odds ratios above 1:10 are considered "strong" evidence.

For each galaxy there is hardly any evidence for or against an Einasto profile (with $\alpha' = 0.2, 0.4$) compared to NFW. This is not unexpected since these profiles are quite similar over a large region (Vera-Ciro et al. 2013). Also in the case of the combined evidence the NFW and Einasto are hard to distinguish. Comparing the NFW or Einasto profiles for individual galaxies to the cored models, one cannot strongly rule out a particular model. For Fornax, Sculptor and Carina, the $\gamma = 2$ models (where the transition speed is fast) appear to be less likely, but this is not the case for Sextans. However, when

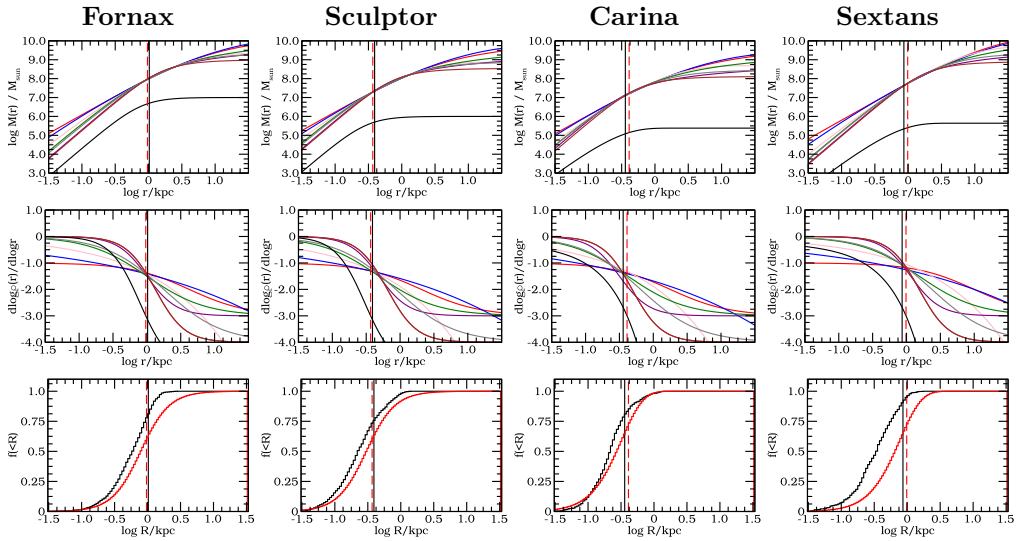


Figure 3.6: Top row: Enclosed mass as a function of radius for the different dark matter density profiles, with the stellar component in black. **Middle row**: Logarithmic density slope as a function or radius, where the black curve corresponds again to the stellar component. The red dashed line indicates r_{-3}, the radius at which the light profile has a logarithmic slope of -3, while the black line indicates $r_{1/2}$, the radius at which half of the stellar mass in enclosed (in 3d). **Bottom row**: Cumulative density distribution of the (2d) radial distribution of the data (black), and the light (red) showing the kinematic data is sampled more concentrated towards the center.

we look at the combined evidence, i.e. we explore whether all dwarfs are embedded in the same halos, such $\gamma = 2$ models are clearly disfavored.

The results for Sculptor may be compared to those of Chapter 2. In that chapter we found that the maximum likelihood value for the central slope of the density profile corresponded to a cored model. Since the evidence is the integral of the pdf, and not directly related to the maximum likelihood (except for a Gaussian distribution), we should not be surprised to find a slightly stronger evidence for the NFW case here. In any case, the differences between the models are minor as shown graphically in Fig. 3.4, and the evidence and the maximum likelihood (marginally) favoring different models can be attributed simply to not being able to distinguish amongst these.

3.4.3 A robust slope measurement

We now inspect in more detail the shape of the mass distributions found for the various best fitting models. We are interested in exploring why the differences between the various models as small as apparent in Fig. 3.4.

The top row of Fig. 3.6 shows the enclosed dark matter mass for the best fit models (indicated by the solid dots in Fig. 3.3) for each galaxy separately. We use the same color coding as in Fig. 3.3, and also include the stellar mass in black. The red-dashed vertical lines denote r_{-3}, the radius at which the light density profile has a logarithmic slope

of -3, while the black line indicates $r_{1/2}$. This remarkable figure shows that for each galaxy there is a region where the mass distributions are truly almost indistinguishable from one another. The different profiles, each characterized by its own functional form, scale radius r_s and mass $M_{1\mathrm{kpc}}$, conspire to produce a unique mass distribution. This region extends from slightly below r_{-3} to approximately the location of the outermost data point (see bottom panel). Here $M(r) \propto r^x$, where x ranges from 1.65 for Fnx, to 1.9 for Sextans.

In the middle row of Fig. 3.6 we plot the logarithmic slope of the dark halo density distribution, where the black line denotes the stellar density. Near the position where the logslope of the stellar density is -3, all the best fit dark matter density profiles seem to reach a similar logslope, although the value of the slope varies from galaxy to galaxy. The radius where the logslopes coincide lies, as expected, inside the region where the mass distribution is well determined, since both quantities are related through derivatives.

To illustrate the distribution of the kinematic sample with respect to the light, we plot in the bottom row of Fig. 3.6 the cumulative 2d radial distribution of the kinematic data in black. The cumulative 2d radial distribution for the light is plotted as the red histogram. All kinematic datasets are more concentrated than the light, but no clear trend is visible between the distribution of the kinematic sample with respect to the light, and the exact location where the logslope of most accurately determined.

The existence of a finite region where the mass is more accurately determined has also been observed in the literature in works using MCMC in combination with Jeans modeling. For example, it is visible in e.g. the right panel of Fig 1. in Wolf et al. (2010), Fig. 18 in Walker (2012b), and Fig. 10 in Jardel et al. (2013) for Draco, in the case of a non-parametric density distribution with Schwarzschild models.

The analysis of Wolf et al. (2010) used the light weighted average of the velocity dispersion to relate the radius at which the logslope of the light is -3, or the half light radius, to the point where the mass is accurately (being independent on the anisotropy) and precisely (showing the least uncertainty) determined. Our findings go beyond this result. They suggest that whatever dynamical model or method is explored, there is a better set of parameters to describe the mass distribution of dSph galaxies. Let r_{-3} be the radius at which the logslope of the (3d) light distribution is -3. Since the mass is accurately determined in this region, a natural parameter would be $M_{-3} = M(r_{-3})$. And since also the logslope at this radius is accurately determined, the next parameter should be $\kappa_{-3} = \frac{\mathrm{d}\log\rho}{\mathrm{d}\log r}|_{r=r_{-3}}$. For any general model, if the values of β and γ are fixed, this effectively makes r_s a function of κ_{-3}.

Fig. 3.7 shows the pdf for the M_{-3} and κ_{-3} parameters for both the NFW and core13 models for each galaxy, assuming a flat prior on these parameters in the domain shown in this figure, except for the NFW profile which we limit to $\kappa_{-3} = -1.05$, since for $\kappa_{-3} \geq -1$ the scale radius is unphysical. As can be seen from the pdf while there is still uncertainty associated to the logslope at this radius, the value is nearly model independent and therefore we believe this value to be accurate, especially for Fornax and Sculptor ($\kappa_{-3} = -1.4 \pm 0.15$ and $\kappa_{-3} = -1.3 \pm 0.12$ respectively for the NFW model). Note also that some uncertainties might arise because the kinematics are not sampled exactly according to the light.

These results also help us understand why we found that the scale radii of the best fitting NFW profiles always to be larger than those of the cored models (see Fig. 3.3).

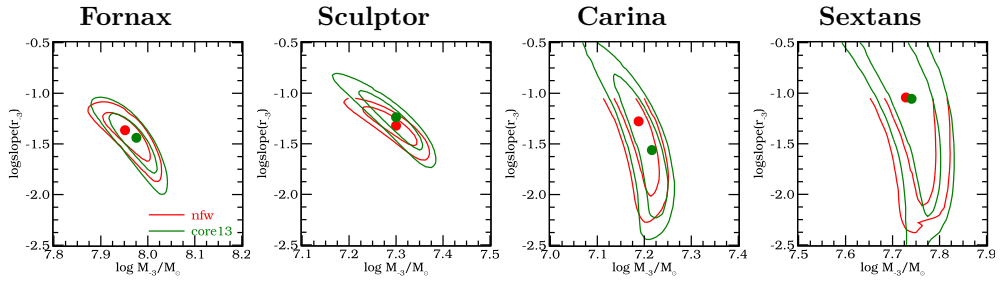


Figure 3.7: Similar to Fig. 3.3, except now using M_{-3} and r_{-3} as parameters. Note that the contours for the NFW cannot go beyond $\kappa(r_{-3}) \geq -1$.

For the NFW, we have

$$\kappa(r) = \frac{\mathrm{d}\log\rho(r)}{\mathrm{d}\log r} = \frac{-2r}{r+r_s} - 1, \tag{3.8}$$

which can be easily solved for r_s:

$$r_{s,\mathrm{nfw}} = -r\frac{\kappa+3}{\kappa+1}. \tag{3.9}$$

A similar solution can be found for the other parametric models, for instance the $\gamma\beta$ model gives:

$$r_{s,\gamma\beta} = r\left(\frac{-\kappa}{\beta+\kappa}\right)^{-1/\gamma}. \tag{3.10}$$

If we now require that the slopes are the same at r_{-3} for the NFW and core13 models, we find

$$\frac{r_{s,\mathrm{nfw}}}{r_{s,\mathrm{core13}}} = \frac{\kappa}{\kappa+1}, \tag{3.11}$$

which is > 1 for $\kappa < -1$, explaining why the cored profiles have smaller scale radii than the NFW profile, i.e. to get the same logslope at the same location for the cored models, their scale radius needs to be smaller than that of the NFW profile. A similar result holds for the other cored models.

3.5 Conclusions

In this chapter we have presented a comparison of dynamical models using different dark matter profiles for four dwarf spheroidal galaxies in the Local Group, namely Fornax, Sculptor, Carina and Sextans. The model comparison was done using Bayesian evidence. We have found that no particular model is significantly preferred, and that all four dwarf spheroidals are compatible with either NFW/Einasto or any of the explored cored profiles. Only Sextans shows a slight preference for cored models, but not with high odds. Nonetheless, we find that it is very unlikely that all four dwarf spheroidals are each embedded in a cored dark matter halo of the form $\rho_{DM} \propto 1/(1+r^2)^{\beta/2}$, with $\beta = 3, 4$.

Our best fit models however, conspire to produce the same mass distribution over a relatively large range in radii, from r_{-3} up to the last measured data point (which is often

close to the nominal tidal radius obtained from fitting the light profile). This $M(r) \sim r^x$, with $x = 1.65 - 1.9$, is similar to that suggested by Walker et al. (2009b) albeit with a slightly steeper exponent. Another (related) quantity that is robustly determined and independent of the assumed dark matter density profile, is the logslope of the density distribution at r_{-3}. We find for the dwarfs in our sample, that this slope ranges from $\kappa_{-3} \sim -1.4$ at $r_{-3} = 0.96$ kpc for Fornax, to $\kappa_{-3} \sim -1.1$ at $r_{-3} = 0.98$ kpc for Sextans.

These findings can be seen as an extension of the results of Wolf et al. (2010), who showed that the mass at r_{-3} can be determined very accurately in a model independent fashion. These authors demonstrated that this result might be understood from the Jeans equation. Although we do not have yet a solid mathematical explanation for our new findings, we suspect that this might be obtained using the virial theorem, which is effectively another, yet independent moment of the collisionless Boltzmann equation.

In the near future, we will apply Schwarzschild modeling to the same data but instead of the moments, we will use the discrete individual measurements directly. This approach should allows us to get the most out of the data, since no information is lost. When binning, one loses spatial resolution, but also the higher moments of the line-of-sight velocity distribution are not included in the fitting procedure because of their large and asymmetric errors. Furthermore, the use of the full line-of-sight velocity distribution should improve the precision of the anisotropy profile, which may be an interesting quantity to discriminate formation scenarios.

This moments-to-discrete modeling step must be carried out before deciding if and how much more data is needed to discriminate among various dark matter density profiles. Nonetheless, we have learned here that the functional form of the mass distribution may be determined over a large distance range, even when only a few hundred velocity measurements are available (as in the case of Sextans). However, the uncertainty on the value of the exact slope of the density profile at e.g. r_{-3} is driven by the sample size.

An obvious next step is to establish if the subhalos extracted from cosmological simulations have the right characteristics to host the dSph of the Milky Way, now that not only the mass, but also its functional form (1st and 2nd derivatives), of their dark halos have been determined reliably.

Acknowledgments

We are grateful to Giuseppina Battaglia, Glenn van de Ven and Remco van den Bosch for discussions that led to the work presented here. We acknowledge financial support from NOVA (the Netherlands Research School for Astronomy), and European Research Council under ERC-StG grant GALACTICA-240271.

Bibliography

Battaglia, G., Helmi, A., Tolstoy, E., Irwin, M., Hill, V., & Jablonka, P. 2008, ApJ, 681, L13

Battaglia, G., Tolstoy, E., Helmi, A., Irwin, M., Parisi, P., Hill, V., & Jablonka, P. 2011, MNRAS, 411, 1013

Battaglia, G., et al. 2006, A&A, 459, 423

Bryan, S. E., Kay, S. T., Duffy, A. R., Schaye, J., Dalla Vecchia, C., & Booth, C. M. 2012, arXiv: 1207.4555

Cretton, N., de Zeeuw, P. T., van der Marel, R. P., & Rix, H. 1999, ApJS, 124, 383

Governato, F., et al. 2012, MNRAS, 422, 1231

Hayashi, E., Navarro, J. F., Taylor, J. E., Stadel, J., & Quinn, T. 2003, ApJ, 584, 541

Helmi, A., et al. 2006, ApJ, 651, L121

Helmi, A., Sales, L. V., Starkenburg, E., Starkenburg, T. K., Vera-Ciro, C. A., De Lucia, G., & Li, Y.-S. 2012, ApJ, 758, L5

Irwin, M., & Hatzidimitriou, D. 1995, MNRAS, 277, 1354

Jardel, J. R., & Gebhardt, K. 2012, ApJ, 746, 89

Jardel, J. R., Gebhardt, K., Fabricius, M. H., Drory, N., & Williams, M. J. 2013, ApJ, 763, 91

Jeffreys, H. 1998, Theory of Probability, Third Edition, International series of monographs on physics (Clarendon Press)

Kazantzidis, S., Łokas, E. L., Callegari, S., Mayer, L., & Moustakas, L. A. 2011, ApJ, 726, 98

Koch, A., Grebel, E. K., Wyse, R. F. G., Kleyna, J. T., Wilkinson, M. I., Harbeck, D. R., Gilmore, G. F., & Evans, N. W. 2006, AJ, 131, 895

Kuhlen, M., Diemand, J., & Madau, P. 2007, ApJ, 671, 1135

Łokas, E. L., Kazantzidis, S., Klimentowski, J., Mayer, L., & Callegari, S. 2010, ApJ, 708, 1032

Mackay, D. J. C. 2003, Information Theory, Inference and Learning Algorithms, first edition edn. (Cambridge University Press)

Mateo, M., Olszewski, E. W., Pryor, C., Welch, D. L., & Fischer, P. 1993, AJ, 105, 510

Mateo, M. L. 1998, ARA&A, 36, 435

Mayer, L. 2010, Advances in Astronomy, 2010

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

—. 1997, ApJ, 490, 493

Navarro, J. F., et al. 2010, MNRAS, 402, 21

Pontzen, A., & Governato, F. 2012, MNRAS, 421, 3464

Richstone, D. O., & Tremaine, S. 1984, ApJ, 286, 27

Rix, H., de Zeeuw, P. T., Cretton, N., van der Marel, R. P., & Carollo, C. M. 1997, ApJ, 488, 702

Springel, V., et al. 2008, MNRAS, 391, 1685

Starkenburg, E., et al. 2010, A&A, 513, A34

Valluri, M., Merritt, D., & Emsellem, E. 2004, ApJ, 602, 66

van den Bosch, R. C. E., van de Ven, G., Verolme, E. K., Cappellari, M., & de Zeeuw, P. T. 2008, MNRAS, 385, 647

van der Marel, R. P., Cretton, N., de Zeeuw, P. T., & Rix, H.-W. 1998, ApJ, 493, 613

Vera-Ciro, C. A., Helmi, A., Starkenburg, E., & Breddels, M. A. 2013, MNRAS, 428, 1696

Walker, M. G. 2012a, arXiv:1205.0311

—. 2012b, ArXiv e-prints

Walker, M. G., Mateo, M., & Olszewski, E. W. 2009a, AJ, 137, 3100

Walker, M. G., Mateo, M., Olszewski, E. W., Gnedin, O. Y., Wang, X., Sen, B., & Woodroofe, M. 2007, ApJ, 667, L53

Walker, M. G., Mateo, M., Olszewski, E. W., Peñarrubia, J., Wyn Evans, N., & Gilmore, G. 2009b, ApJ, 704, 1274

Wolf, J., Martinez, G. D., Bullock, J. S., Kaplinghat, M., Geha, M., Muñoz, R. R., Simon, J. D., & Avedo, F. F. 2010, MNRAS, 406, 1220

Chapter 4

Discrete orbit based dynamical models of the Sculptor dSph galaxy

Maarten A. Breddels and Amina Helmi

Abstract

We fit the full line of sight velocity distribution of Sculptor with orbit-based dynamical models. Unlike previous work based the moments of the line of sight velocity distribution, we use the individual measured velocities. When we test our method on a Mock dataset, we find that we can recover well the distribution function and the parameters of the underlying potential. To estimate the effects of sampling, we test the method both using the expectation value of the log likelihood of the data and on a discrete dataset of 2,000 stars, i.e. the size available for Sculptor, as well as one that is 5× larger. This enables us to show we accurately estimate the parameters of the potential and without bias with 6% (3%) uncertainty in the mass for a kinematic dataset of 2,000 (10,000) stars. When we apply our method to Sculptor, assuming a Navarro, Frenk & White dark matter profile, we find a mass of $M_{1kpc} = 10^{8.05\pm0.024} M_\odot$ and $r_s = 5.1^{+2.6}_{-1.7}$ kpc which are consistent, though slightly larger, than the estimates presented in Chapter 2 and 3. Interestingly, the distribution function of our best fit model appears to be bimodal, a result that is also present when the dataset is modeled using its moments and with different dark matter halo profiles. We show that an association to the metal-poor and metal-rich components known to exist in Sculptor is plausible. Hence this constitutes a demonstration of the full power of Schwarzschild's method, because the multi-component nature of Sculptor is not explicitly assumed in our modeling.

4.1 Introduction

The dwarf spheroidal galaxies, satellites of the Milky Way, are the most dark-matter dominated systems known to-date. Their very large mass-to-light ratios (in the range 10 - 1000s, Wolf et al. 2010; Walker et al. 2009b), make them ideal to constrain the nature of dark matter, through estimates of their mass content and distribution, in a way that they provide important tests to the ΛCDM cosmological concordance model. The satellites of the Milky Way are nearby which permits accurate measurements of the line of sight velocities for hundreds to thousands of individual stars (e.g. Battaglia et al. 2008; Walker et al. 2009a), and such datasets are thus mature to call for sophisticated dynamical modeling.

Such dynamical models allow the inference of intrinsic properties from projected quantities like sky positions and line of sight velocities. Not only mass content, but for example, orbit-based dynamical models also constrain the galaxy's potential and its distribution function. Although much information about the line of sight velocity distribution (LOSVD) is contained in its first four moments, working with these has some disadvantages. Most importantly it involves binning of the data, and this leads to loss of spatial information since all stars in a (usually radial) bin are grouped together. Furthermore higher moments tend to have large and non-gaussian errors, complicating the analysis. It would be more natural to model the full LOSVD to then evaluate the likelihood in a straightforward way. However, not all dynamical modeling methods provide these, and in particular the widely used Jeans models only fit, nearly by definition, the moments of (rather than) the LOSVD itself.

In principle all methods that aim at the determination of the underlying distribution function, be it parametrically or using orbits as building blocks, should be able to compute the LOSVD, and in this Chapter we explore the latter, i.e. Schwarzschild models. The basic steps of Schwarzschild's method consist in integrating a set of orbits in a given potential, calculating the predicted observables for each orbit, and then weighing the orbits (with non-negative weights) to obtain a model that fits the observed data. This approach guarantees that the distribution function (which is reflected in the orbit-weights) is non-negative. This method was originally used by Schwarzschild (1979) to demonstrate that a self consistent solution in dynamic equilibrium exists for a triaxial system, but was only implemented to reproduce the density distribution. The method was later extended to include kinematic constraints (Richstone & Tremaine 1984; Pfenniger 1984). Since then many codes have been developed (e.g. Richstone & Tremaine 1984; Rix et al. 1997; van der Marel et al. 1998; Cretton et al. 1999; Valluri et al. 2004; van den Bosch et al. 2008). Initially only the lowest moments of the line of sight velocity distribution (mean velocity and velocity dispersion) were employed, but better data have led to the inclusion of higher moments in the fits. While employing the moments allows one to use linear or quadratic programming to find the orbit-weights, also likelihood methods for discrete datasets have been developed (e.g. Merritt & Tremblay 1993; Wu & Tremaine 2006; Chanamé et al. 2008). A great advantage of Schwarzschild modelling is that it is non-parametric in the distribution function, and thus it does not require for example the specification of the anisotropy profile, which is in fact an outcome of the model. Previous works modeling the dynamics of dSph using this method include Jardel & Gebhardt (2012); Jardel et al. (2012, for applications on the Fornax and Draco dSph), and Chapter 2 (or Breddels et al. 2012) and Chapter 3 (or Breddels & Helmi 2013) on Sculptor, Fornax, Carina and Sextans. Most of these works have employed moments of the LOSVD, although Jardel & Gebhardt (2012) model the full LOSVD approximating it with a kernel density estimator in radial bins, which however, does not fully exploit the discrete nature of the datasets.

Because of the obvious advantages of non-parametric methods and of the use of the full LOSVD, in this Chapter, we implement Schwarzschild's method in a fully discrete approach to apply it on dwarf spheroidal galaxies. In §4.2 we first explain how we implement our discrete Schwarzschild model numerically. We then test our method on a Mock model of a dSph galaxy in §4.3. In §4.4 we discuss the issue of modeling the foreground contamination, while in §4.5 we

apply our method to the Sculptor dwarf, and finally conclude in §4.6.

4.2 Dynamical model

4.2.1 Generalities

The phase-space structure of a galaxy can be specified by its distribution function (hereafter df) $f(\boldsymbol{x}, \boldsymbol{v})$, where $\boldsymbol{x}$ and $\boldsymbol{v}$ are the position and velocity coordinates respectively. The probability of finding a star in the volume $\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}$ is given by $f(\boldsymbol{x}, \boldsymbol{v})\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}$. All observables may be derived from knowledge of the df. For example the normalised surface density:

$$\mu(x, y) = \int \mathrm{d}z\mathrm{d}\boldsymbol{v} f(\boldsymbol{x}, \boldsymbol{v}), \tag{4.1}$$

where z is the direction along the line-of-sight.

According to the (strong) Jeans (1915) theorem, the df of a steady-state stellar system in which almost all orbits are regular, is a function of the isolating integrals of motion (see also Binney & Tremaine 2008). Spherically symmetric systems (both in the tracer's density and the underlying potential) have only regular orbits and generally respect 4 isolating integrals of motion, the energy and the 3 components of the angular momentum vector. However, if the galaxy shows no rotation, due to symmetry, the df will depend only on the energy and the modulus of the angular momentum vector, i.e. $f(\boldsymbol{x}, \boldsymbol{v}) = f(E, L)$. Furthermore if the velocity distribution is isotropic, the df can only depend on energy and $f(\boldsymbol{x}, \boldsymbol{v}) = f(E)$.

Most dSph galaxies are so distant that currently the only phase-space coordinates that may be measured are the projected stellar positions on the sky, and the line-of-sight velocities of (a subset of) its stars. These can be used to derive the surface density $\mu_0(R)$ and the moments of the line-of-sight velocity distribution, such as:

$$\mu_0(R) = \int \mathrm{d}z\mathrm{d}\boldsymbol{v} f(E, L), \tag{4.2}$$

$$\mu_2(R) = \frac{1}{\mu_0(R)} \int \mathrm{d}z\mathrm{d}\boldsymbol{v} v_\parallel^2 f(E, L), \tag{4.3}$$

$$\mu_4(R) = \frac{1}{\mu_0(R)} \int \mathrm{d}z\mathrm{d}\boldsymbol{v} v_\parallel^4 f(E, L). \tag{4.4}$$

Here R is the projected distance on the sky from the centre of the galaxy and $v_\parallel$ the velocity along the line-of-sight, after subtraction of the centre of mass mean motion.

Although the zeroth moment $\mu_0(R)$, is very useful in comparing the photometric data to a particular distribution function (see §4.2.2) since it is defines the probability that a star should be at (projected) radial distance R from the center, the rest of the moments do not define probability density functions (pdf) themselves. Following Chanamé et al. (2008), we are interested in the pdf that describes the probability of finding a star with a particular line-of-sight velocity at a given position:

$$p(\mathrm{v}_{\mathrm{los}}|R) = \frac{1}{\mu_0(R)} \int \mathrm{d}z\mathrm{d}\boldsymbol{v}_\perp f(E, L), \tag{4.5}$$

where $\boldsymbol{v}_\perp$ indicates the velocity vector perpendicular to the line of sight. Note that since the kinematic data often is not randomly sampled in R, we are not interested in $p(\mathrm{v}_{\mathrm{los}}, R)$ but in the conditional pdf $p(\mathrm{v}_{\mathrm{los}}|R) = p(\mathrm{v}_{\mathrm{los}}, R)/p(R)$.

This equation together with Eq. (4.2) for the surface density, constitute the basis for finding the answer to the fundamental question: what is the probability of our data given a model?

In the remainder of this Chapter we deal mainly with the implementation of a method that constitutes a numerical approximation to Eq. (4.2) and Eq. (4.5), and which using orbit-based dynamical (Schwarzschild's) method allows us to establish the model that best fits the data in the sense described above.

4.2.2 From the model to the probability density function

Our implementation of the Schwarzschild method follows a similar approach to that presented in Chapter 2, which is based on Rix et al. (1997) and van den Bosch et al. (2008). The main difference is that instead of finding the moments of the LOSVD, we actually obtain the full LOSVD and the associated projected stellar density distribution. Our procedure is quite similar to Chanamé et al. (2008). It differs in that instead of computing the probability for each star, it is performance-wise more convenient to store the (model) LOSVD on a grid since for dwarf spheroidal galaxies the only kinematic information available at the moment are the line of sight velocities.

For convenience we define $l = L/L_{\max}$ the relative angular momentum (where $L_{\max}$ is the angular momentum of a circular orbit of energy E), such that $l \in [0, 1]$. This enables us to define a rectangular grid in energy and relative angular momentum. Since the Schwarzschild method is based on orbit integrations in a specified gravitational potential, the df may be seen as a sum of Dirac delta functions:

$$f(E, L) = \sum_{i,j} \hat{f}_{i,j} \delta(E - E_i) \delta(L - l_j L_{\max,i}), \tag{4.6}$$

where $\sum_{i,j} \hat{f}_{i,j} = 1$ and $\hat{f}_{i,j} \geq 0$.

To define the grid in energy and (relative) angular momentum we proceed as follows. For the energy we choose N'_E radii between a minimum and maximum radius spaced logarithmically, and take the corresponding energy of a purely radial orbit. The minimum and maximum radii we consider are 0.033 kpc and 24.492 kpc, respectively. For each energy we choose N'_l relative angular momenta spaced linearly between 0 and 1. All orbits are integrated starting from their apocentre. For simplicity we refer to a single orbit with a single index j instead of the double index i, j.

Since the location of the star is not a random observable (as is the line of sight velocity) but chosen beforehand, the probability of a set of parameters $\boldsymbol{\theta}$ that define the gravitational potential and distribution function of the galaxy, given the measured line of sight velocity v_i at radius R_i is:

$$p(\boldsymbol{\theta}|v_i, R_i) = \frac{p(v_i|R_i, \boldsymbol{\theta}) p(\boldsymbol{\theta})}{p(v_i|R_i)}, \tag{4.7}$$

and for a whole dataset, this reduces to a product over i. For parameter inference the normalization $p(v_i|R_i)$ (the evidence) is not relevant, and the prior on $p(\boldsymbol{\theta})$ needs to be specified. The likelihood term $p(v_i|R_i, \boldsymbol{\theta})$ is where the dynamical model plays a role.

For a fixed set of characteristic parameters of the specified gravitational potential, we integrate a set of orbits, and each orbit j defines a $p_j(v_i, R_i|\boldsymbol{\theta})$, such that given a set of orbit-weights w_j we have a full pdf from

$$p(v_i, R_i|\boldsymbol{\theta}) = \frac{1}{\sum_j w_j} \sum_j w_j p_j(v_i, R_i|\boldsymbol{\theta}), \tag{4.8}$$

where the first factor automatically normalizes the pdf. For the likelihood we actually need

$p(v_i|R_i,\boldsymbol{\theta})$ since, as mentioned above, the position is not a random variable:

$$p(v_i|R_i,\boldsymbol{\theta}) = \frac{1}{\int dv p(v_i,R_i|\boldsymbol{\theta})}\frac{1}{\sum_j w_j}\sum_j w_j p_j(v_i,R_i|\boldsymbol{\theta}), \tag{4.9}$$

$$= \frac{1}{p(R_i|\boldsymbol{\theta})}\frac{1}{\sum_j w_j}\sum_j w_j p_j(v_i,R_i|\boldsymbol{\theta}), \tag{4.10}$$

$$= \frac{1}{\sum_j w_j}\sum_j w_j p_j(v_i|R_i,\boldsymbol{\theta}). \tag{4.11}$$

In the numerical implementation we integrate the orbits with a fixed timestep and count each time the star is inside a cell of a grid in R and v. We choose this grid linear in R between $R = 0, R_{\rm max}$, and linear between $v = -v_{\rm max}, v_{\rm max}$ with size $N_R \times N_v$. The choices for the $R_{\rm max}, v_{\rm max}, N_R, N_v$ depend on the resolution needed or wanted and storage (and/or performance) issues. For instance choosing $v_{\rm max}$ much larger than the escape velocity of a particular model will compromise the resolution in the v-direction, while hardly any star may lie close to this value. We therefore choose to fix $v_{\rm max}$ to 40 km/s for Sculptor, which is around $4\sigma_{\rm los}$, where $\sigma_{\rm los}$ is its velocity dispersion along the line of sight. Using $N_v = 30$ gives a resolution of $\Delta v = 80/30$ km/s $= 2.6$ km/s which we found to be sufficient (also considering the average error of 2 km/s in the data). For the radial direction we choose $R_{\rm max} = 1.5$ kpc and $N_R = 30$, giving a resolution of $\Delta R = 1.5/30$kpc $= 50$ pc. Increasing the resolution leads to more shot-noise from the finite integration time, and is potentially a big issue for datasets with larger dimensionality (see McMillan & Binney 2013).

When an orbit is outside the range of the $R - v$ grid, a count is missed. To have a proper prediction for the light (or mass) distribution, we keep a separate R grid with the same resolution in the R direction as the $R - v$ grid, but where counts are added independent of the respective velocity. After the integration all grids are normalized by dividing by the number of integration points ($N_{\rm orbit} = 5000$) and the resolution size of a grid element ($\Delta v \times \Delta R$ for the $R - v$ grid) such that:

$$\sum_k^{N_v}\sum_l^{N_R} \Delta v \times \Delta R \times g^{R,v}_{j,k,l} \leq 1, \tag{4.12}$$

$$\sum_l^{N_R} \Delta R \times g^{R}_{j,l} \leq 1, \tag{4.13}$$

where $g^{R,v}_j$ and g^R_j are the probabilities on the grids that correspond to $p_j(v,R|\boldsymbol{\theta})$ and $p_j(R|\boldsymbol{\theta})$ respectively for orbit j. For orbits that are fully inside the radial and velocity range (see e.g. Fig 4.1 for a variety of examples) the equality holds. Some orbits probe regions beyond the radial or velocity range. For those the inequality holds (the counts inside the grid are smaller than the number of integration points), and although the pdfs ($g^{R,v}_j$ and g^R_j) of the individual orbits may not be normalized in these cases, care is taken such that the final pdf is properly normalized.

For every v_i and R_i we calculate the index for these grids, k_i, l_i corresponding to the index in the v and R directions. Therefore given a set of orbit-weights the numerical equivalent of Eq. 4.10 (and Eq. 4.11) is:

$$p(v_i|R_i,\boldsymbol{\theta}) = \frac{1}{\sum_j w_j \sum_k^{N_v} g^{R,v}_{j,k,l_i}}\sum_j w_j g^{R,v}_{j,k_i,l_i}. \tag{4.14}$$

Notice that in the denominator we use the $g^{R,v}$ grid and sum over velocity rather than over the g^R grid directly, this is to ensure that the numerator and the denominator are normalized in

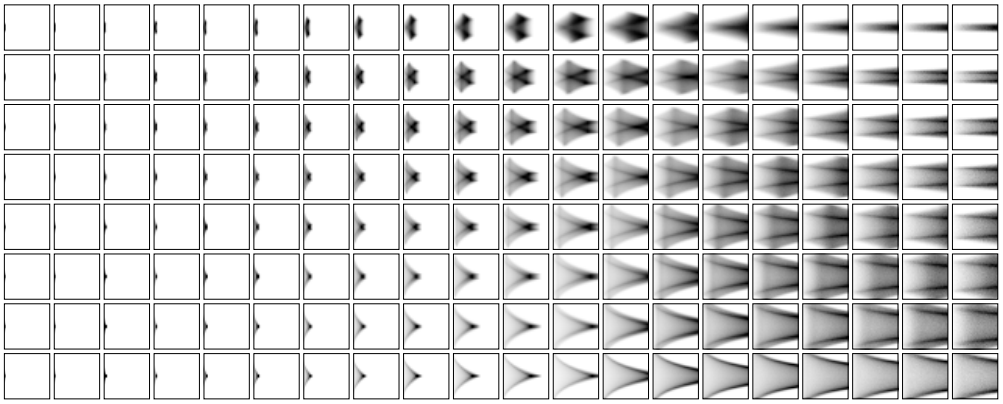


Figure 4.1: Line of sight velocity distributions: building blocks for the discrete Schwarzschild method. Each panel represents the $p_j(v_i, R_i|\boldsymbol{\theta})$ distribution associated to an orbit with energy E_i and angular momentum L_j, with R along the x-axis, and $\mathrm{v_{los}}$ along the y-axis. From left to right the binding energy is decreasing, and the angular momentum increases from bottom to top.

the same way. We can treat the photometry in a similar way as the kinematics, in which case the likelihood factors become:

$$p(R_i|\boldsymbol{\theta}) = \frac{1}{\sum_l^{N_R} \sum_j w_j g_{j,l}^R} \sum_j w_j g_{j,l_i}^R. \tag{4.15}$$

We show a graphical representation of the $g_{j,k,l}^{R,v}$ (the numerical equivalent of $p_j(v, R|\boldsymbol{\theta})$) in Fig. 4.1. Each panel represents $g_{j,k,l}^{R,v}$, i.e. one orbit (although see below for dithering), where the dark color means higher probability. For each panel the x-axis represents the radial direction and the y-axis represents the velocity axis. On the bottom left the lowest angular momentum and highest binding energy orbits are found, while the panels to the right represent orbits with lower binding energy (larger radii) and the panels to the top represent higher angular momentum orbits.

As in Chapter 2 we implement dithering by effectively replacing each $R-v$ grid by an average of a set of $R-v$ grids with neighbouring energy and angular momentum values. This leads to a more continuous distribution function instead of a small set of delta functions in energy and angular momentum space while keeping a relatively low number of degrees of freedom for the orbit-weights defining the distribution function.

4.2.3 Fitting the data

Apart from fitting the data we also include an extra regularization term, as done in Chapter 2. This effectively reduces the parameter space and yields a smoother df. This smoothness requirement is implemented by demanding the second order derivatives of the orbit-weights in the energy and angular momentum direction to be small (for more details, see §2.2.3). To convert the χ^2_{reg} of Chapter 2 into a likelihood term, we add the following term to our likelihood:

$$\mathcal{L}_{\mathrm{reg}} = e^{-\chi^2_{\mathrm{reg}}}, \tag{4.16}$$

where the λ_L and λ_E in Eq. (2.25a-2.25c) determines the relative strength of the regularization compared to the data. As in Chapter 2, we use $\lambda_L = \lambda_E/8$ and $\lambda_E = 0.1$.

Given a set of kinematic datapoints of size $N_{\rm kin}$ and a set of $N_{\rm light}$ photometric measurements, we can evaluate the likelihood of this dataset for a given gravitational potential and set of orbit-weights using Eqs. (4.14) and (4.15):

$$\mathcal{L} = \left(\prod_i^{N_{\rm kin}} p(v_i|R_i, \boldsymbol{\theta}) \right) \times \left(\prod_j^{N_{\rm light}} p(R_j|\boldsymbol{\theta}) \right) \times \mathcal{L}_{\rm reg}, \tag{4.17}$$

where $\boldsymbol{\theta}$ are the characteristic parameters of the potential and the orbit-weights, and we assume that the positions of the kinematic samples are chosen, which is generally the case for kinematic samples for dSph galaxies. For clarity we have used a different index for the two products, to emphasize that the kinematics and photometry will generally be completely different (independent) datasets.

It is not possible to find the best fitting orbit-weights using quadratic programming as done in Breddels et al. (2012), since the likelihood is not linear nor quadratic in the orbit-weights. We therefore have to resort to a general optimization routine. We have found the BFGS optimization routine from the Python package scipy to work adequately. The gradient of Eq. (4.17) can easily be computed which significantly speeds finding the optimum. Since the orbit-weights need to be non-negative, appropriate boundary conditions are required.

4.3 Testing with a mock model

In this section, we take a similar approach as in Chapter 2 to test our code. We first fix the parameters of the potential to their true values to see if we can estimate the orbit-weights and therefore recover the distribution function. After this, we test how well we can estimate the parameters of the potential. McMillan & Binney (2013) argue that finding the distribution function is possible, but that due to random numerical errors on the log likelihood, that parameter estimation of the potential is practically impossible. We do not appear to have this problem, most likely due to the lower dimensionality of the observable space in the case of data for dSph ($v_{\rm los}$ and R).

4.3.1 Dataset versus expected log likelihood

Given a kinematic dataset $\{R_i, \mathrm{v}_{\rm los,i}\}$ of size $N_{\rm kin}$ drawn from a distribution $p(\mathrm{v}_{\rm los}|R, \boldsymbol{\theta}_{\rm true})$, where $\boldsymbol{\theta}_{\rm true}$ indicates the true parameters of the potential and the true orbit-weights, the likelihood of this dataset is

$$\mathcal{L}_{\rm kin} = \prod_{i=1}^{N_{\rm kin}} p(\mathrm{v}_{\rm los,i}|R_i, \boldsymbol{\theta}), \tag{4.18}$$

$$\log \mathcal{L}_{\rm kin} = \sum_{i=1}^{N_{\rm kin}} \log p(\mathrm{v}_{\rm los,i}|R_i, \boldsymbol{\theta}). \tag{4.19}$$

We can take the expectation value of the log likelihood over the data:

$$E[\log \mathcal{L}_{\rm kin}] = \sum_{i=1}^{N_{\rm kin}} \int d\mathrm{v}_{\rm los,i} \int dR_i p(\mathrm{v}_{\rm los,i}|R_i, \boldsymbol{\theta}_{\rm true}) \log p(\mathrm{v}_{\rm los,i}|R_i, \boldsymbol{\theta}) \tag{4.20}$$

$$= \sum_{i=1}^{N_{\rm kin}} \int d\mathrm{v}_{\rm los} \int dR p(\mathrm{v}_{\rm los}|R, \boldsymbol{\theta}_{\rm true}) \log p(\mathrm{v}_{\rm los}|R, \boldsymbol{\theta}) \tag{4.21}$$

$$= N_{\rm kin} \int d\mathrm{v}_{\rm los} \int dR p(\mathrm{v}_{\rm los}|R, \boldsymbol{\theta}_{\rm true}) \log p(\mathrm{v}_{\rm los}|R, \boldsymbol{\theta}), \tag{4.22}$$

where $p(\mathrm{v_{los}}|R,\boldsymbol{\theta}_{\rm true})$ is the conditional pdf of the line of sight velocities for the true model (i.e. the data), and $p(\mathrm{v_{los}}|R,\boldsymbol{\theta})$ that of a particular model, and the integral is the negative of the cross entropy. The cross entropy is minimal when $p(\mathrm{v_{los}}|R,\boldsymbol{\theta}_{\rm true}) = p(\mathrm{v_{los}}|R,\boldsymbol{\theta})$, which corresponds to the maximum likelihood when the correct model is found. We also can see from this equation that the expection value of the log likelihood of N observations is simply N times the expected log likelihood of a single observation. This makes studying the effect of sample size on the uncertainty of the estimated model parameters simply a multiplication of the individual expected log likelihoods. Tests of the method based on a single realisation cannot detect systematic biases since the maximum likelihood never equals the true values of the parameters simply due to sample variance. On the other hand, when we calculate the expectation value, we can recognize biases directly. Tests using the expectation value of the log likelihood are similar in spirit to those performed in Chapter 2, where we considered the expected moments (rather than only those obtained from single realisations of the mock model).

A similar approach can be used for the light profile, in which case we find

$$E[\log\mathcal{L}_{\rm light}] = \sum_{i=1}^{N_{\rm light}} \int dR_i p(R_i|\boldsymbol{\theta}_{\rm true}) \log p(R_i|\boldsymbol{\theta}) \tag{4.23}$$

$$= N_{\rm light} \int dR p(R|\boldsymbol{\theta}_{\rm true}) \log p(R|\boldsymbol{\theta}), \tag{4.24}$$

where $N_{\rm light}$ is the number of photometric samples.

When we test our code using this method, $p(\mathrm{v_{los}}|R,\boldsymbol{\theta}_{\rm true})$ and $p(R|\boldsymbol{\theta}_{\rm true})$ are given since we know the underlying gravitational potential and distribution function of our mock model. Then for a particular set of characteristic parameters for the potential (reflected in $\boldsymbol{\theta}$), we adjust the orbital weights (which is also reflected in $\boldsymbol{\theta}$), to obtain $p(\mathrm{v_{los}}|R,\boldsymbol{\theta})$ and $p(R|\boldsymbol{\theta})$, until we find the maximum likelihood. Note that in this simple test, we cannot change the parameters of the potential and the orbit-weights at the same time, since changing the potential requires calculating a complete set of orbits.

4.3.2 Recovering the distribution function

As in Chapter 2 we use our mock model to test our code. Our mock model consists of a spherically symmetric stellar distribution of the Plummer form with scale radius $b = 0.3$ kpc and total mass $M_* = 10^6 \mathrm{M}_\odot$. The stars are embedded in a dark matter halo of the NFW form with $r_s = 0.5$ kpc and an enclosed mass withing 1 kpc of $\mathrm{M_{1kpc}} = 10^8\ \mathrm{M}_\odot$. We put the velocity anisotropy to a constant value of $\beta = -1$ and numerically find the distribution function as explained in Appendix 2.A.

Since we know the distribution function and the potential of our mock model (i.e. we know $\boldsymbol{\theta}_{\rm true}$), we can directly compute $p(\mathrm{v_{los}}|R,\boldsymbol{\theta}_{\rm true})$ and $\log p(R|\boldsymbol{\theta}_{\rm true})$ in Eqs. (4.22) and (4.24), where we have chosen $N_{\rm kin} = 2{,}000$, and $N_{\rm light} = 200{,}000$. Both $p(\mathrm{v_{los}}|R,\boldsymbol{\theta}_{\rm true})$ and $p(\mathrm{v_{los}}|R,\boldsymbol{\theta})$ are convolved with a gaussian of 2 km/s dispersion in the velocity direction to simulate the measurement errors. We then change the orbit-weights to optimize $E[\log\mathcal{L}_{\rm kin}] + E[\log\mathcal{L}_{\rm light}] + \log\mathcal{L}_{\rm reg}$ (i.e. Eq. 4.22 + Eq. 4.24 and the log of Eq. 4.16) using the algorithm as explained in §4.2.3. This is the same as finding the maximum likelihood of the expectation value of the log of Eq. (4.17). Note that since the regularization term does not depend on the data, its expectation value is the term itself.

We show the orbit-weights in the top left panel of Fig. 4.2 which are a good match to the true orbit-weights shown in Fig. 2.2. The true values for the velocity dispersions (top middle panel), velocity anisotropy (bottom middle panel) and mass distribution (bottom left panel) are also recovered well.

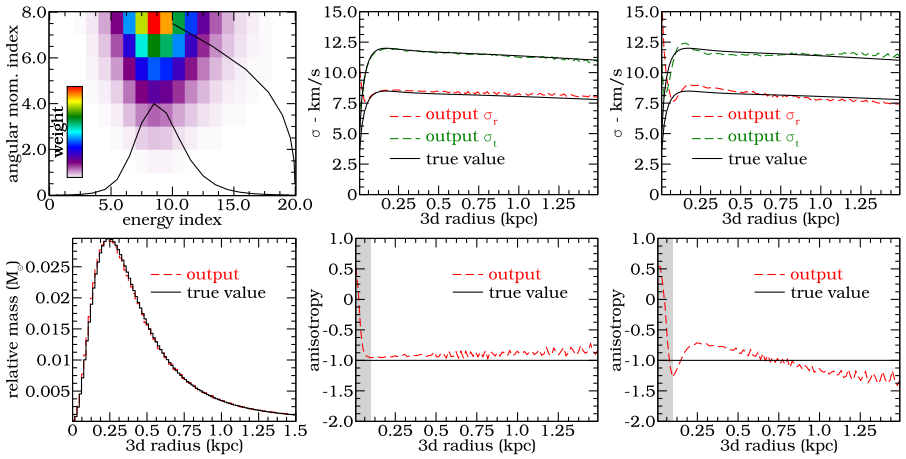


Figure 4.2: Top left: Orbit-weights found using the maximum likelihood of the expectation value of the log likelihood method for a kinematic dataset of $N_{\rm kin} = 2000$. Black lines indicate the projected orbit-weights in the energy and angular momentum direction. **Bottom left:** Mass distribution for the found solution in the red dashed line, true value in black. **Top middle:** Intrinsic velocity dispersion in the radial direction (red) and the tangential direction (green), the black lines indicate the true velocity dispersions. **Bottom middle:** Velocity anisotropy profile in red, with the true value in black. **Right column:** Similar to the middle column, except now showing the solution using a discrete sample of $N_{\rm kin} = 2000$.

Although we have shown the results for the expectation value of the log likelihood, in practice we always deal with discrete datasets. We therefore repeat the same procedure, except that we now simply use Eq. (4.17). Our more realistic (discrete) dataset is produced by drawing $2,000$ stars from the true distribution function and convolving the line of sight velocities with a 2 km/s gaussian error. The results of the fitting procedure for the velocity dispersions and anisotropy are shown on the right column of Fig. 4.2. This shows the expected sampling effects, but also shows that the trends are similar. The gray regions in the bottom middle and right panels indicate where we cannot recover the anisotropy profile reliably (see Chapter 2 for an extensive discussion).

4.3.3 Recovering the parameters of the potential

Since we have shown that we can closely recover the distribution function (or orbit-weights), we now focus on recovering the parameters of the potential. For each $\mathrm{M}_{1\mathrm{kpc}}$ and r_s we have to integrate a set of orbits as outlined in §4.2. Then for each $\mathrm{M}_{1\mathrm{kpc}}$ and r_s we find the best fitting orbit-weights, and construct the pdf of these parameters. With uniform priors on $\log r_s$ and $\mathrm{M}_{1\mathrm{kpc}}$, we show this pdf in the left column of the left part of Fig. 4.3 for a discrete realization of the mock model. The blue line (dot) indicates the maximum likelihood while the red line (dot) shows the true value. As noted in the previous paragraph, since this is only a particular sample, we do not expect the maximum likelihood to be equal to the true value, and the same holds for the velocity anisotropy. We therefore take the test a step further and check that our method indeed works when using the expectation value. The corresponding results are shown on the right part of the same Fig. 4.3. Here we see the maximum likehood value overlaps perfectly

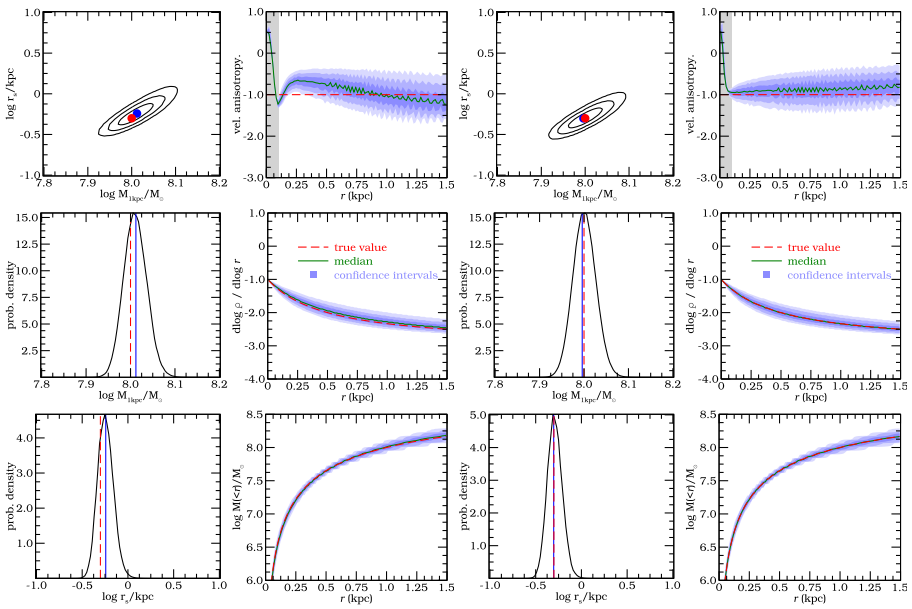


Figure 4.3: Probability distribution function for Mock Sculptor for a sample of 2,000 stars, with the two leftmost columns corresponding to the same sample of Fig. 4.2, and the rightmost columns to the expectation value. For each plot the following description holds. **Left column:** Probability distribution functions (joint and marginalised) for mass and scale parameters of the NFW dark matter halo potential recovered for mock Sculptor model. Blue dot and blue lines (left column) indicate the maximum likelihood value (of the unmarginalised pdf), while the red dot and vertical dashed lines indicate the input values for the mock Sculptor model. The green solid line indicates the median value and the blue regions (or black contour lines in the top left panel) the 68.3, 95.4 and 99.7 per cent confidence intervals. **Top right:** Recovered anisotropy profile. The grey region indicates where our method cannot recover the anisotropy (see text for details). **Middle right:** Recovered logarithmic density slope for the dark matter. **Bottom right:** Recovered enclosed mass profile.

with the true value, and the anisotropy is recovered accurately outside the grey region, except maybe for a slight deviation at large radii. If we look at the pdf of M_{1kpc} and r_s we see that the pdf is sharply cut off at the bottom left. This is due to these models having a smaller escape velocity than the maximum velocity present in the dataset, which evaluates to a likelihood of zero since these models would claim it impossible to have stars with velocities larger than the particular model's escape velocity. Since we consider a line of sight velocity error of 2 km/s, this sharp border is not that near the true value, while if we decrease the errors this sharp border approaches it.

Note that in 68% of cases, a particular sample drawn from this model will have its maximum likelihood within the 1σ contour. The uncertainty on the mass based on the expectation value of the log likelihood is $\sigma_{\log M_{1kpc}} = 0.026$, while that of the scale radius is $\sigma_{\log r_s} = 0.081$ for a sample of 2,000 stars. A natural question is what is the gain when the sample size is increased. The left panel of Fig. 4.4 shows the pdf using the log likelihood method for a kinematic dataset of 10,000 stars with errors of 2 km/s. The right panel shows the corresponding anisotropy. If the pdf

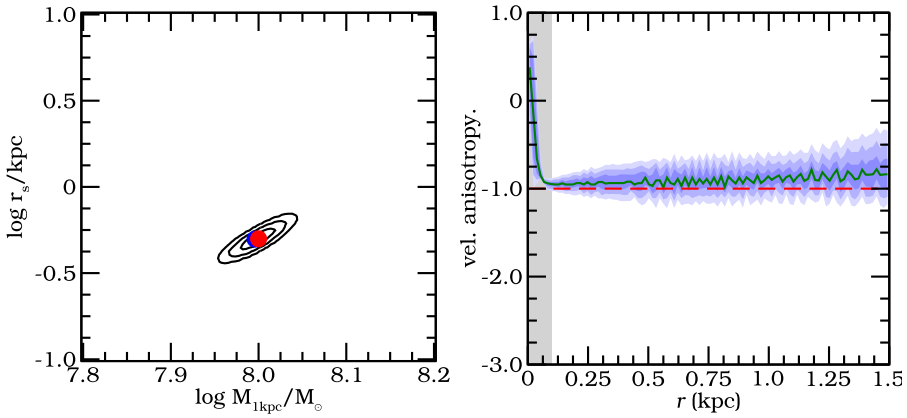


Figure 4.4: Left: Pdf for Mock Sculptor for 10,000 stars using the expectation value for the log likelihood. **Right:** Corresponding anisotropy profile.

were well described by a multivariate Gaussian, one would expect the uncertainties to decrease by a factor of $\sqrt{5} \approx 2.2$. We find however that the uncertainties are now $\sigma_{\log \mathrm{M}_{1\mathrm{kpc}}} = 0.013$ and $\sigma_{\log r_s} = 0.040$, a decrease of a factor of 2 for both parameters. Although very close to the factor of $\sqrt{5}$, we see that the comparison to a multivariate Gaussian only holds to a certain degree. The uncertainty on the velocity anisotropy also decreases significantly, which could constitute an additional valid motivation for getting kinematic datasets of this size.

4.4 A foreground model for Sculptor

In contrast to our mock model, Local Group dwarf spheroidal galaxies suffer from foreground contamination from the Milky Way. This contamination is apparent for example in Fig. 3.1 of Chapter 3 where the data from Battaglia et al. (2008); Walker et al. (2009a) are shown. The most likely contaminants may be removed using, for example, a 3σ clipping (e.g. Battaglia et al. (2008); Chapter 2), a membership probability (Walker et al. 2009c), dwarf/giant discriminators (Battaglia & Starkenburg 2012), or an interloper removal scheme based on some prejudice about the dynamical state of the galaxy (Łokas 2009). A sharp cut in any of the parameter spaces will unlikely be the best approach since members and non-members will generally overlap in any space used, be it velocity or metallicity.

In this section we therefore explore a different approach. Instead of trying to remove foreground stars from our catalogue, we develop a joint kinematic model for the foreground and for the dwarf galaxy in question, and derive a probability for each star for belonging to either component.

4.4.1 Joint kinematical model for Sculptor and the foreground

If we assume that the contamination by the Milky Way does not vary with sky position within the angle subtended by the dwarf galaxy, then the probability of a star being a member of the dwarf galaxy or being a contaminant can be expressed as,

$$p(m|R_\epsilon) + p(\neg m|R_\epsilon) = 1, \tag{4.25}$$

$$p(m|R_\epsilon)/p(\neg m|R_\epsilon) = \mu(R_\epsilon)/\mu_{0,\mathrm{fg}}, \tag{4.26}$$

where $p(m|R_\epsilon)$ is the probability of being a member of the galaxy at elliptical radius R_ϵ while $p(\neg m|R_\epsilon)$ being the probability of not being a member. The function $\mu(R_\epsilon)$ is the (normalized) surface density of the stars which, for a constant stellar mass-to-light ratio, is equal to the surface brightness. In the case of Sculptor, we assume this follows a Plummer form with a scale radius of $b = 0.3$ kpc. Note that for this model we use the elliptical radius, since the Plummer fit to the surface density is performed with elliptical instead of circular annuli. However, in the rest of this Chapter we use a circular radius since our dynamical model is spherically symmetric. The quantity $\mu_{0,\mathrm{fg}}$ is the constant surface density of the foreground and depends also on the selection function for example, on the location of stars in the color-magnitude diagram. This quantity is difficult to predict a priori, and therefore it is one of the parameters to determine with our kinematic model.

If we also take the velocity into consideration, the probability to find a star i with velocity v_i, at elliptical radius $R_{\epsilon,i}$ is:

$$p(v_i|R_{\epsilon,i}) = \overbrace{p(v_i|m,R_{\epsilon,i})}^{\text{dSph model}} \underbrace{p(m|R_{\epsilon,i})}_{\text{Eqs. 4.25}-\text{4.26}} + \overbrace{p(v_i|\neg m,R_{\epsilon,i})}^{\text{Foreground model}} \underbrace{p(\neg m|R_{\epsilon,i})}_{\text{Eqs. 4.25}-\text{4.26}} . \tag{4.27}$$

The first factor of the first term is the velocity model of our dwarf galaxy. The first factor of the second term is the velocity model of the foreground. The last factors of both terms come from Eqs. (4.25 - 4.26). Note that in the case of multiple datasets of the same galaxy, a particular star may appear more than once (through repeat observations, or because the datasets were obtained by different authors). In that case, the velocity used is equal to the mean of the various measurements. Since our goal is to have a foreground model we keep our velocity model of the dwarf galaxy simple and assume it can be well approximated by a Gaussian. We make the same assumption for the foreground velocity distribution. In total our model has six parameters, the mean and dispersion of the two Gaussians and the two constant surface densities ($\mu_{0,\mathrm{fg}}$), one for each of the datasets we work with (those from Battaglia et al. 2008; Walker et al. 2009a).

To find the best fitting foreground model, we optimize the following log likelihood function:

$$\log \prod_i^N p(v_i|R_{\epsilon,i}) = \sum_i^N \log p(v_i|R_{\epsilon,i}), \tag{4.28}$$

To check the results of our best model fit, we make a detailed comparison to the data. We take the radial bins used in Chapter 3, and use these as the limits of our elliptical bins. (These bins were found such that in each radial bin there were at least 250 members inside 3σ. Note that for the current analysis, it is not important how many stars are in each bin, and that using the radial bins to define our elliptical bins is convenient and bears no particular meaning). We count the number of stars within 3σ of the mean velocity of Sculptor (N_{inside}), and outside this 3σ (N_{outside}), where σ is the velocity dispersion of the dwarf galaxy found for the best fitting model. From our model we can calculate the fraction of members (f_{m}) and non members (f_{nm}) inside this 3σ at a given radius, since these correspond to the area under the best-fit Gaussian in or outside this given velocity range. Expressed using the terms in Eq. (4.27)

$$f_{\mathrm{m}} = \int_{v_{\mathrm{min}}}^{v_{\mathrm{max}}} dv p(v|m,R_\epsilon)p(m|R_\epsilon), \tag{4.29}$$

$$f_{\mathrm{nm}} = \int_{-\infty}^{v_{\mathrm{min}}} dv p(v|\neg m,R_\epsilon)p(\neg m|R_\epsilon) + \int_{v_{\mathrm{max}}}^{\infty} dv p(v|\neg m,R_\epsilon)p(\neg m|R_\epsilon), \tag{4.30}$$

where, in this case, we have chosen v_{min} and v_{max} corresponding to the mean velocity of Sculptor plus or minus 3σ.

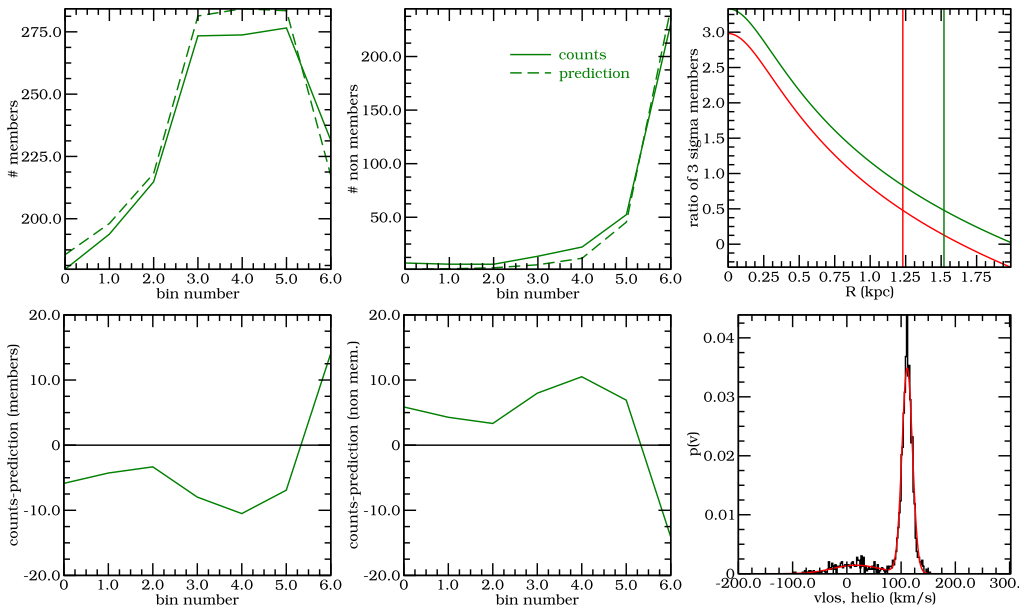


Figure 4.5: Top left: Number of "member" stars in each elliptical bin, where "members" are defined as being within 3σ of the mean velocity of Sculptor. **Bottom left:** Difference between the estimated number of members from the counts versus the predicted number based on the best fit models. **Middle column:** Similar to the left column, but now for the non-members. **Top right:** ratio of member to non members for both datasets (red is for the dataset of Battaglia et al. (2008) while green for Walker et al. (2009a)). **Bottom right:** Normalized histogram of the radial velocity of all stars in the datasets in black, and in red the best fit model.

Using the observed counts and these estimates we then solve the following equations to get an "observed" number of members and non-members within this velocity range:

$$N_{\rm inside} = f_{\rm m} N_{\rm members} + f_{\rm nm} N_{\rm non\text{-}members}, \quad (4.31)$$

$$N_{\rm outside} = (1 - f_{\rm m}) N_{\rm members} + (1 - f_{\rm nm}) N_{\rm non\text{-}members}. \quad (4.32)$$

On the other hand, we can also predict directly from the model the number of members and non-members in each bin, which are simply $N \times f_{\rm m}$ and $N \times f_{\rm nm}$, where N is the number of stars in each bin. In the top left panel of Fig. 4.5 we plot the number of "observed" members with the solid line, while the number of predicted members in each bin is given by the dashed line. The comparison for the number of non-members is plotted in the top middle panel. The differences are plotted in the panels below these. In the bottom right panel we plot the normalized histogram of the line-of-sight velocity for all the stars in Sculptor in black, with the best fit model in red. This comparison shows that we accurately describe the full velocity distribution. In the top right panel we show for each dataset separately the logarithm of the fraction of member to non-members inside 3σ of Sculptor's mean velocity and highlight the radius at which this ratio is 3:1. In the following sections we will not consider stars that are outside this elliptical radius, since we believe that the simple model we are using has its limitations, and that by introducing a conservative cut we will be less affected by biases. Note that it is not crucial that our kinematic model for Sculptor reproduces perfectly its observed velocity distribution, since we only need

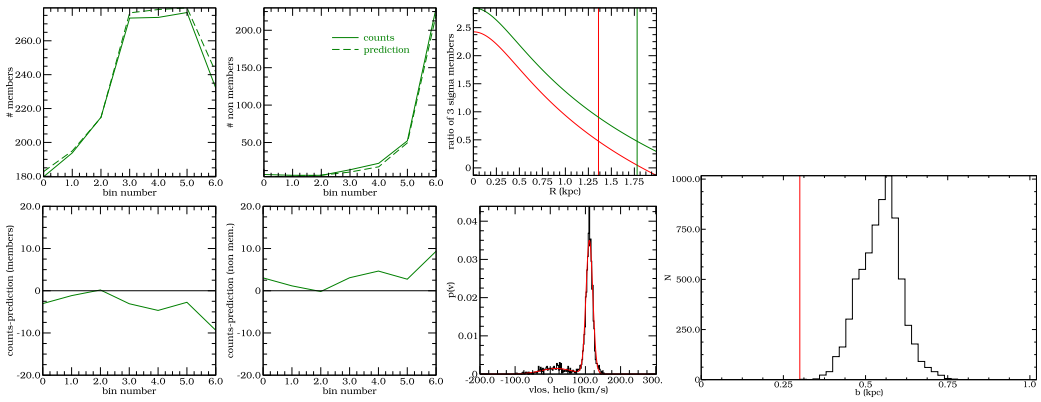


Figure 4.6: Left panels: Same as Fig. 4.5 except the best fitting model is found leaving the light profile scale radius as a free parameter. **Right panel:** Distribution of the light profile scale parameter from an MCMC analysis, showing the scale radius of the light profile obtained from the photometry is inconsistent with that found with our kinematical model.

to estimate the foreground contamination for the more sophisticated orbit based dynamical modeling that follows.

4.4.2 Foreground model issues

From Fig. 4.5 we see that at small radii the predicted number of members is larger than that based on counting, while in the last bin the model underpredicts the number of members. A possible explanation is that the analytic (Plummer) light profile is not a good match at larger radii, as inspection of Fig. 4.1 of Battaglia (2007) suggests. Since the kinematics are more uniformly sampled compared to the light, we have relatively more stars at large than at small radii compared to what would be obtained from a random sample based on the photometry. While the Plummer profile is generally a good fit to the photometric data, and a mismatch at large radii may be considered small because most of the stars are located at smaller distances, the sampling of the kinematics implies that the description of the light profile needs to be accurate at all radii. To explore the effect of the light profile on our kinematical model, we now set the scale radius of Sculptor's light profile as a free parameter, and find again the best model for the foreground and the dSph. We show the results for this model in Fig. 4.6 (similar to Fig. 4.5). With a best fitting scale radius of $b = 0.47$ kpc, instead of $b = 0.3$ kpc based on fits to photometric datasets, the number of members and non-members are predicted more accurately.

As a final test, we explore the parameter space of our seven parameters (the six mentioned above and the scale radius b), using Eq (4.28) by means of a Markov Chain Monte Carlo method. In the right panel of Fig. 4.6 we show the marginalized distribution of the scale radius of the light (b), where the red line indicates the value of 0.3 kpc, which describes the light accurately. It is clear from this figure that the scale parameter that fits the light distribution is not consistent with the distribution derived from fitting the kinematic sample. Since we cannot find a model that accurately describes the foreground, we will perform a clipping in velocity space to remove most of the foreground contimination in the next section.

4.5 Application to the Sculptor dSph galaxy

4.5.1 Data

As in Chapter 3, we use the kinematic datasets of both Battaglia et al. (2008) and Walker et al. (2009a). However, because we can not model the foreground accurately as shown in the previous section, we proceed as follows. We still make use of the foreground model (as in Chapter 3) to do a clipping in elliptical radius, and consider only those stars within an elliptical distance where the ratio of member-to-non-member is smaller than 3:1 within 3σ of the mean velocity of Sculptor. We then fit a parametric model in each bin to estimate a velocity dispersion. After this we do a 3σ clipping per radial bin. As in Chapters 2 and 3, we assume the light profile to be of Plummer form with a scale radius $b = 0.3$ kpc. Instead of evaluating the likelihood of the photometry in Eq. (4.17), we use the expection value of the log likelihood of the light as in §4.3, and Eq. (4.24) assuming that the light profile is estimated from $N = 20,000$ stars. The velocity gradient known to be present in the data is removed in the same way as in Chapter 2.

4.5.2 Parameter estimation for the NFW model

Using the cleaned up dataset, we find the orbit-weights that maximize the likelihood (Eq. 4.17) assuming an NFW potential with characteristic parameters $\mathrm{M}_{1\mathrm{kpc}}$ and r_s that are to be determined. To take velocity errors into account, we convolve our pdfs, $p(v_{\mathrm{los}}|R, \boldsymbol{\theta})$, with a Gaussian with a dispersion of 2 km/s (instead of using the individual errors and for performance reasons). The fitting procedure is essentially the same as that done for mock sculptor.

We calculate the pdf for Sculptor on a grid for a range of parameters for the NFW profile, and the findings are shown in Fig. 4.7. Compared to the results of Chapters 2 and 3 (see Fig. 2.8 and Fig 3.3) we obtain here a comparable mass and a somewhat larger scale radius, $\mathrm{M}_{1\mathrm{kpc}} = 10^{8.047\pm0.024}\ \mathrm{M}_\odot$(this chapter) versus $\mathrm{M}_{1\mathrm{kpc}} = 10^{8.004\pm0.028}\mathrm{M}_\odot$ (Chapter 2) and $r_s = 10^{0.71\pm0.18} \approx 5.1^{+2.6}_{-1.7}$ kpc versus $r_s = 10^{0.26\pm0.23} \approx 1.8^{+1.3}_{-0.8}$ kpc , while the velocity ellipsoid is slightly less tangentially anisotropic now. We have also checked that the mass M_{-3} and the logarithmic slope of the density profile γ_{-3} at r_{-3} (i.e. the radius where the logarithmic slope of the light profile is $\gamma = -3$), are entirely compatible with our previous estimates. Inspection of the moments in Fig. 4.8 shows that the preferred model found by our discrete Schwarzschild method is able to reproduce remarkably well both the velocity dispersion and kurtosis profiles and does so better than the method used in Chapter 2 (see Fig. 2.7).

The reasons behind the slight differences in the results provided by the methods are worth investigating. A possible explanation may lie in the influence of the discrete fitting, which perhaps allows for more flexibility in finding a model that fits the full LOSVD nearly on a point by point basis. Another difference with respect to Chapter 2 is in the fitting of the light profile, which we previously required to be matched to within 1%, while in this Chapter it is treated statistically, i.e. in the same way as the kinematic samples.

The mass estimates of Strigari et al. (2008) and Walker et al. (2009b) are compatible with our mass profile as shows in the lower right panel of Fig. 4.7. The mass estimate of Wolf et al. (2010) however is on the high side compared to that derived by the method presented here.

4.5.3 The distribution function

For the best fit model found in the previous section, and indicated by the black dot in Fig. 4.7, we show in the left panel of Fig. 4.9 the orbit-weights that define the distribution function. We note that this distribution seems to be bimodal, and appears to have two distinct components: one at low angular momentum (near energy index 8, angular momentum index 2), and a second one at higher angular momentum (near energy index 7, angular momentum index 7). To establish the robustness of this finding and for comparison, we have included in the same figure the distribution

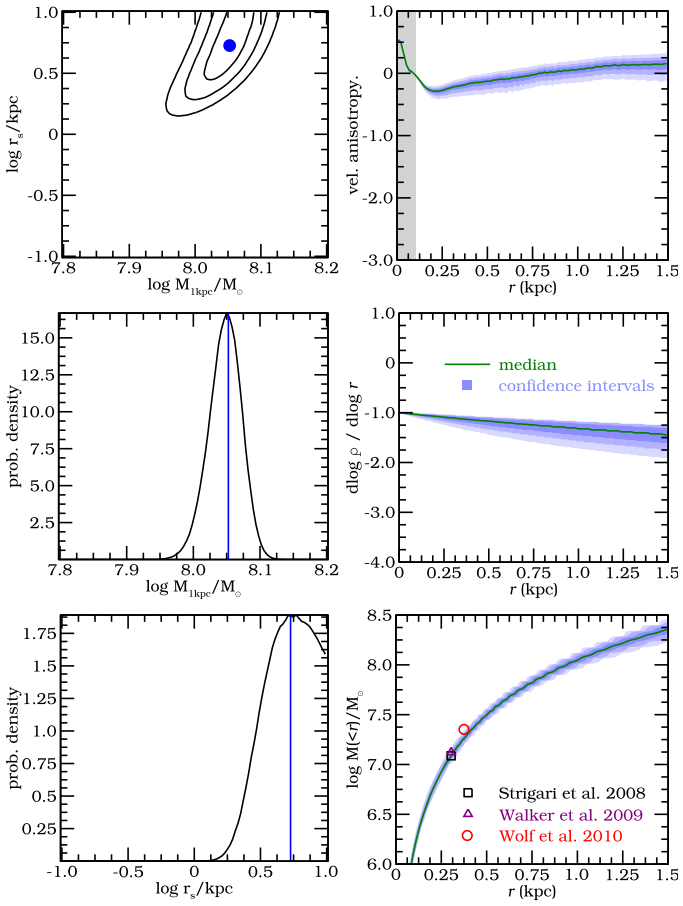


Figure 4.7: Left column: Probability density functions (joint and marginalised) for mass and scale parameters of the NFW dark matter halo potential recovered for Sculptor. Blue dot and blue lines (left column) indicate the maximum likelihood value (of the unmarginalised pdf). The green solid line indicates the median value and the blue regions (or black contour lines in the top left panel) the 68.3, 95.4 and 99.7 per cent confidence intervals. **Top right:** Recovered anisotropy profile. The grey region indicates where we cannot recover the anisotropy. **Middle right:** Recovered logarithmic density slope for the dark matter. **Bottom right:** Recovered enclosed mass profile.

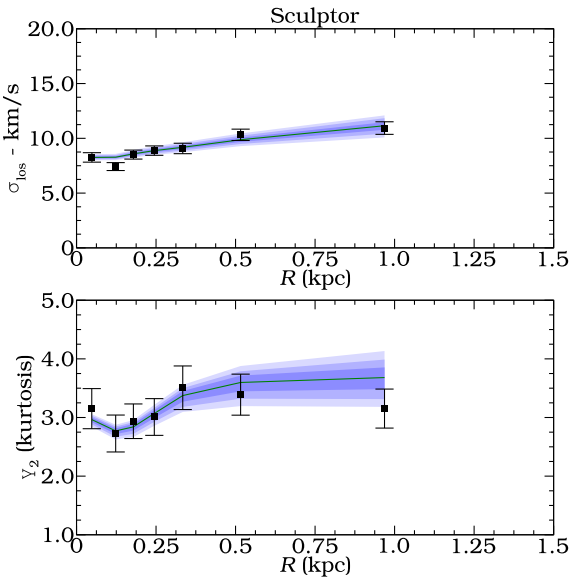


Figure 4.8: Line-of-sight velocity dispersion profile obtained in radial bins for the data by Battaglia et al. (2008) and Walker et al. (2009a) of Sculptor, taking into account the foreground contamination by the Milky Way, similar to Chapter 2. The solid curve corresponds to the (pdf weighted) median line-of-sight velocity dispersion profile from the discrete Schwarzschild models for the NFW profile, while the contours indicate the 1, 2, and 3σ uncertainties around this curve. The last bin extends to 1.3 kpc.

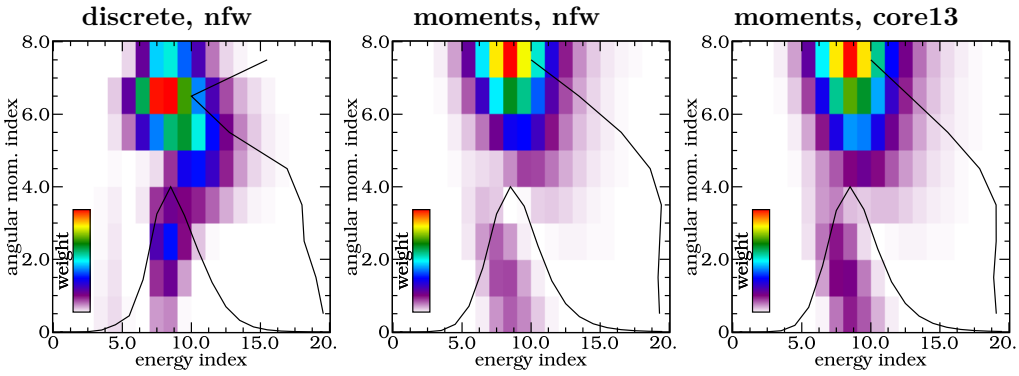


Figure 4.9: Orbit-weights for the best fit NFW model obtained using our discrete Schwarzschild method (left, see §4.5.2); the best fit NFW model obtained by fitting the moments (middle, see Chapter 3); and the best fit "core13" model also found using the moments (right, see Chapter 3). Black lines indicate the projected orbit-weights in the energy and angular momentum direction.

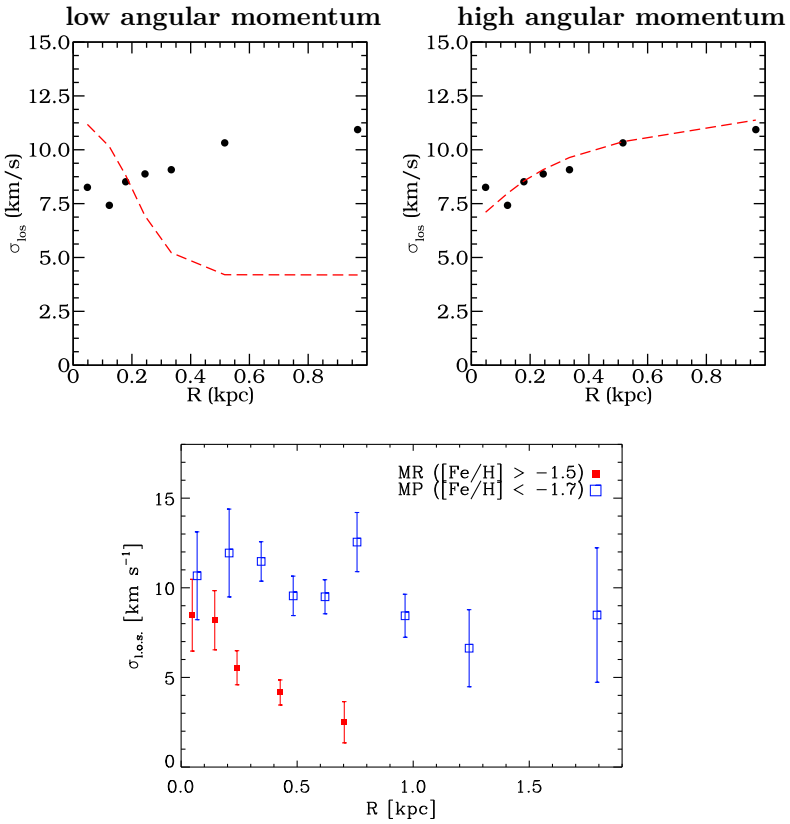


Figure 4.10: Top left: The black dots correspond to the velocity dispersion for Sculptor. The red dashed curve shows the line of sight velocity dispersion for the low angular momemtum part of the distribution function. **Top right:** Similar, but now for the high angular momentum part. **Bottom:** Velocity dispersion of the metal rich and metal poor populations in Sculptor (Battaglia et al. 2008). The velocity dispersion profile of the metal rich stars shows similarities to the component of the df with low angular momentum.

functions found in Chapter 3, for the "nfw" and the "core13" models. Although the exact location of the lower component is at a slightly different energy and angular momentum, it is reassuring that this component is found using both the discrete and the traditional Schwarzschild methods and also for a different dark matter profile.

To obtain some insight into what this second component corresponds to, we proceed as follows. We split the orbit-weights into two separate groups (that would correspond to separate distribution functions): one with angular momentum index ≤ 3 (low angular momentum) and the other with angular momentum index > 3. In Fig. 4.10 we show the line of sight velocity dispersion for the two distribution functions separately, left and right correspond to the low and high angular momentum parts respectively. The low angular momentum part shows an increasing velocity dispersion towards the centre and falls off rapidly with radius. This component is centrally concentrated and makes up 17% of the mass within 200 pc and 30% of the mass within 100 pc. Therefore it seems plausible it is related to the metal rich component in Sculptor, which

is centrally concentrated and shows a rather similar velocity dispersion profile (see Battaglia et al. 2008, and the bottom panel of Fig. 4.10). This result is quite remarkable since we have not assumed at any point the existence of multiple components in the Sculptor dwarf. Moreover it demonstrates that the metal-rich and metal-poor stars indeed are dynamically distinct, and that Sculptor is not simply a system with a radial gradient in stellar populations. Furthermore, this finding constitutes a demonstration of the full power of the Schwarzschild approach, and would not have been possible if the approach used to model the distribution function would have been parametric.

4.6 Discussion and Conclusions

In this Chapter, we have shown how a discrete orbit based dynamical can be applied to dwarf spheroidal galaxies. We first applied the method to a mock model resembling the Sculptor dwarf, both using a discrete kinematic dataset as well as using the expectation value of the log likelihood. Both tests confirm that the method can recover the distribution function, although the discrete sample of 2000 stars clearly shows the effects of limited sampling. We also showed that we can estimate the parameters of the dark matter halo potential reliably ($M_{1\mathrm{kpc}}$ within 6% and the scale radius within 18%), where the results of the expectation value of the log likelihood analysis shows these estimates are not biased and therefore accurate.

Local Group dSph suffer from foreground contamination from the Milky Way, and this needs to be taken into account. We have explored the possibility of including the modeling of the foreground in our discrete maximum likelihood approach. This would allows to refrain from performing hard cuts to establish membership. However, when we model the foreground in this way we find that the best fit model that is compatible with the light distribution fails to predict the right number of members and non-members established kinematically. When releasing some of the assumptions regarding the light profiles we can find a model that accurately describes members and non-members, but that is incompatible with the known light profile established from large photometric datasets and studies. We suspect that the Plummer profile often used is insufficiently accurate at large radii to describe the surface density of stars, which is crucial for predicting the right number (kinematic) members.

By lack of a proper foreground model, we perform a clipping in velocity space to remove the most likely non-members, and a clipping in elliptical radius to avoid including regions that include too many non-members (the outskirts). We model Sculptor using an NFW profile and find $M_{1\mathrm{kpc}} = 10^{8.05\pm0.024}$ $M_\odot$ and $r_s = 5.1^{+2.6}_{-1.7}$ kpc. Although the scale radius found is larger than that found in Chapter 2, the values meet at the 1σ confidence level, which shows that the systematic error due to modeling differences is still small compared to the uncertainties. The anisotropy although slightly closer to zero, is still negative $\beta \sim -0.2$ at intermediate radii and becomes more radial at larger radii. The distribution function for the best fit NFW potential is bimodal in energy and angular momemtum space. The two components may be split in a low and high angular momentum parts, and the properties of the low angular momentum component are similar to the metal rich component known to be present in Sculptor. It depicts a centrally rising velocity dispersion profile, and a more concentrated light distribution (Battaglia et al. 2008).

All the principles elaborated in this paper can be applied and extended to more sophisticated orbit based models. Non spherical models can be implemented by storing the LOSVD in a two-dimensional (radial and angular) grid, instead of the one-dimensional radial grid used here. Non-spherical models naturally require three integral models to adequately describe the distribution function, for both axisymmetric and triaxial models. Given the power and success of our Schwarzschild's models, it is only natural that we plan to perform such discrete orbit-based non-spherical modeling to a variety of systems in the near future.

Acknowledgements

AH and MB are grateful to NOVA for financial support. AH acknowledges financial support the European Research Council under ERC-Starting Grant GALACTICA-240271.

Bibliography

Battaglia, G. 2007, PhD thesis, Kapteyn Astronomical Institute, University of Groningen

Battaglia, G., Helmi, A., Tolstoy, E., Irwin, M., Hill, V., & Jablonka, P. 2008, ApJ, 681, L13

Battaglia, G., & Starkenburg, E. 2012, A&A, 539, A123

Binney, J., & Tremaine, S. 2008, Galactic Dynamics (Princeton University Press)

Breddels, M. A., & Helmi, A. 2013, ArXiv e-prints

Breddels, M. A., Helmi, A., van den Bosch, R. C. E., van de Ven, G., & Battaglia, G. 2012, ArXiv e-prints

Chanamé, J., Kleyna, J., & van der Marel, R. 2008, ApJ, 682, 841

Cretton, N., de Zeeuw, P. T., van der Marel, R. P., & Rix, H. 1999, ApJS, 124, 383

Jardel, J. R., & Gebhardt, K. 2012, ApJ, 746, 89

Jardel, J. R., Gebhardt, K., Fabricius, M., Drory, N., & Williams, M. J. 2012, ArXiv e-prints

Jeans, J. H. 1915, MNRAS, 76, 70

Łokas, E. L. 2009, MNRAS, 394, L102

McMillan, P. J., & Binney, J. J. 2013, ArXiv e-prints

Merritt, D., & Tremblay, B. 1993, AJ, 106, 2229

Pfenniger, D. 1984, A&A, 141, 171

Richstone, D. O., & Tremaine, S. 1984, ApJ, 286, 27

Rix, H., de Zeeuw, P. T., Cretton, N., van der Marel, R. P., & Carollo, C. M. 1997, ApJ, 488, 702

Schwarzschild, M. 1979, ApJ, 232, 236

Strigari, L. E., Bullock, J. S., Kaplinghat, M., Simon, J. D., Geha, M., Willman, B., & Walker, M. G. 2008, Nature, 454, 1096

Valluri, M., Merritt, D., & Emsellem, E. 2004, ApJ, 602, 66

van den Bosch, R. C. E., van de Ven, G., Verolme, E. K., Cappellari, M., & de Zeeuw, P. T. 2008, MNRAS, 385, 647

van der Marel, R. P., Cretton, N., de Zeeuw, P. T., & Rix, H.-W. 1998, ApJ, 493, 613

Walker, M. G., Mateo, M., & Olszewski, E. W. 2009a, AJ, 137, 3100

Walker, M. G., Mateo, M., Olszewski, E. W., Peñarrubia, J., Wyn Evans, N., & Gilmore, G. 2009b, ApJ, 704, 1274

Walker, M. G., Mateo, M., Olszewski, E. W., Sen, B., & Woodroofe, M. 2009c, AJ, 137, 3109

Wolf, J., Martinez, G. D., Bullock, J. S., Kaplinghat, M., Geha, M., Muñoz, R. R., Simon, J. D., & Avedo, F. F. 2010, MNRAS, 406, 1220

Wu, X., & Tremaine, S. 2006, ApJ, 643, 210

Chapter 5

An analytic distribution function for a massless cored stellar system in a cuspy dark matter halo

Maarten A. Breddels and Amina Helmi

Abstract

We demonstrate the existence of distribution functions that can be used to represent spherical massless cored stellar systems embedded in cuspy dark matter halos with constant mildly tangential velocity anisotropy. In particular, we derive analytically the functional form of the distribution function for a Plummer stellar sphere in a Hernquist dark halo, for $\beta_0 = -0.5$ and for different degrees of embedding. This particular example satisfies the condition that the central logarithmic slope of the light profile $\gamma_0 > 2\beta_0$. Our models have velocity dispersion profiles similar to those observed in nearby dwarf spheroidal galaxies. Hence they can be used to generate initial conditions for a variety of problems, including N-body simulations that may represent dwarf galaxies in the Local Group.

5.1 Introduction

In the concordance ΛCDM cosmological model, galaxies are expected to be embedded in massive dark matter halos. Recently, much focus has been placed on measuring and modeling the internal dynamics of dwarf spheroidal galaxies (dSph) as these systems have very high mass-to-light ratios and appear to be dark matter dominated at all radii (see Walker 2012, for a review). Particular emphasis has been put in establishing the characteristics of the host dark matter halos and to determine whether their properties are consistent with those expected in the context of the ΛCDM model (Stoehr et al. 2002; Strigari et al. 2010). More specifically, an open question is whether the dSph satellites of the Milky Way could be embedded in density profiles that are centrally cusped such as the NFW profile (Navarro et al. 1996).

Much of this modeling work has been carried out using the Jeans equations in the spherical limit (e.g. Łokas 2001; Koch et al. 2007; Walker et al. 2007, 2009; Łokas 2009). The general goal has been to constrain the dark matter content (i.e. to estimate the characteristic parameters of given density profiles) by fitting the observed l.o.s. velocity distributions, and more specifically the 2nd and 4th moments, i.e. the dispersion and the kurtosis profiles. In Jeans modeling also the functional form of the velocity anisotropy needs to be specified. A more fundamental limitation is that the existence of a distribution function, once a solution has been found, is not guaranteed. Specifically, there is no assurance that a distribution function that is positive everywhere (i.e. that it is physical), will exist.

More recently, with the application of Schwarzschild's modeling technique, which finds a distribution function in a numerical fashion, this concern has been side-stepped (e.g. Jardel & Gebhardt (2012); Chapter 2). However, the respective distribution functions have not been given in analytic form in these works, and it may not even be plausible to find a simple expression in the most general circumstances. On the other hand, Wilkinson et al. (2002) has introduced a family of distribution functions to represent spherical stellar systems with anisotropic velocity ellipsoids that are embedded in cored dark matter halos

The An & Evans (2009) theorem provides an important constraint regarding distribution functions that may be associated to dSph. This theorem states that a system with a finite central radial velocity dispersion must satisfy that the central value of the logarithmic slope of the stellar density profile γ_0 and the central velocity anisotropy β_0 be related through $\gamma_0 = 2\beta_0$. This implies that if the light profile is perfectly cored, as often assumed, i.e. $\gamma_0 = 0$, then the velocity ellipsoid must be isotropic, *independently of the dark matter halo profile* (which should be shallower than the singular isothermal sphere). However, if the system is cold at the center, i.e. $\sigma_{r,0} = 0$, then the only constraint is that $\gamma_0 > 2\beta_0$, which in the case of cored stellar profiles is satisfied by tangentially anisotropic ellipsoids (Ciotti & Morganti 2010). Since these conditions refer to the intrinsic velocity dispersion, they do not impose strong constraints on the line-of-sight velocity dispersion ($\sigma_{\rm los}$), which is the observable, and one may obtain a flat $\sigma_{\rm los}$ profiles even if the system is intrinsically cold at the centre.

Given the extensive modeling performed assuming cuspy dark matter halos and cored stellar profiles, the natural question that arises is whether physical distribution functions that can reproduce the properties of dSph exist in such cases. For example, Evans et al. (2009) have shown that for a stellar Plummer profile with an isotropic velocity ellipsoid and a *strictly constant velocity dispersion profile*, the dark matter must follow a cored isothermal sphere. We show here that this particular result cannot be generalized, and that (tracer) cored light distributions can exist in equilibrium in cuspy dark matter halos, once the condition of constant velocity dispersion is relaxed.

In this short Chapter, we present a distribution function that represents a massless stellar system following a Plummer profile embedded in a Hernquist dark matter halo, and which has a constant anisotropy $\beta = -1/2$. We focus on this particular example as it is mathematically easier to manipulate, but also because it is observationally sound. The surface brightness profiles of dSph are well fit by Plummer models (Irwin & Hatzidimitriou 1995), and the velocity ellip-

soids derived from Schwarzschild models have radially constant, if slightly negative anisotropies (Chapter 3). Therefore, models such as that presented below can be used, for example, to generate initial conditions for an N-body simulation of a dwarf galaxy resembling a dSph satellite of the Milky Way.

5.2 Methods

5.2.1 Generalities

The distribution function of a spherical system in equilibrium can depend on energy E, and if the velocity ellipsoid is anisotropic, also on angular momentum L: $f(E, L)$. It can be shown that when the distribution function takes the form $f(E, L) = f_1(E)L^{-2\beta}$, with $\beta = cst$, then then this β is the constant velocity anisotropy of the system.

The functional form of the energy part of the distribution function can be determined through an Abel equation, as outlined in Sec. 4.3.2 of Binney & Tremaine (2008). In that case (see their Eq. 4.67), we may derive $f_1(E)$ from

$$C_\beta \frac{d}{d\Psi}(r^{2\beta}\nu) = (\frac{1}{2} - \beta) \int_0^\Psi d\epsilon \frac{f_1(\epsilon)}{(\Psi - \epsilon)^{\beta+1/2}} \tag{5.1}$$

where $\nu(r)$ is the density, $\Psi(r) = -\Phi(r) + \Phi_0$ the relative gravitational potential, $\epsilon = -E = \Psi(r) - 1/2v^2$ the relative energy, and C_β is a constant. This equation is valid for $-1/2 < \beta < 1/2$, and might be inverted using the Abel integral to obtain an analytic expression for $f_1(\epsilon)$. In the case of $\beta = -1$, an additional derivative is needed to reach the Abel integral equation form, but the distribution function may also be derived, now from

$$f_1(\epsilon) = C'_{\beta=-1} \int_0^\epsilon \frac{d\Psi}{(\epsilon - \Psi)^{1/2}} \frac{d^3(\nu/r^2)}{d\Psi^3}. \tag{5.2}$$

These expressions are completely general, but only in the case of gravitational potentials $\Psi(r)$ of simple mathematical form it is possible to invert and obtain r as function of Ψ, and to easily compute all corresponding derivatives.

The case of $\beta = -1/2$ is particularly simple and yields (see Eq. 4.71 of Binney & Tremaine 2008)

$$f_1(\epsilon) = \frac{1}{2\pi^2} \frac{d^2(\nu/r)}{d\Psi^2}\Big]_{\Psi=\epsilon} \tag{5.3}$$

If the system were self-consistent, then the density and the potential would be related through Poisson's equation (see e.g. Baes & Dejonghe 2002). However, in the case of dSph, the gravitational potential is largely determined by the dark matter, and the stars may simply be considered as tracers. In this case, the above equations are still valid but the density is that of the stars $\nu_*(r)$, while we may assume the potential to be that of the dark matter only. A priori, there is no guarantee that for example, the integral in Eq. (5.2) will converge, and that a physical solution, i.e. a positive distribution function leading to a stable system, will exist.

5.2.2 Plummer stellar sphere in a Hernquist dark halo, $\beta = -1/2$

For mathematical convenience we assume that the gravitational potential is given by the Hernquist model (Hernquist 1990), and as explained above, it is meant to describe the (dominant) contribution of the dark matter. Although this model is not cosmologically motivated, its density profile has the same r^{-1} limiting behavior in the inner regions as the NFW model. On the other hand, it has a finite mass M, and a steeper fall off at large radii (as r^{-4} instead of r^{-3}).

This is also why it is often used in the literature to set up N-body simulations. The gravitational potential for the Hernquist model is

$$\Psi(r) = \Psi_0 \frac{1}{1 + r/b}, \tag{5.4}$$

where b is the scale radius, and $\Psi_0 = GM/b$. For the stars we assume a Plummer profile

$$\nu_*(r) = \frac{3}{4\pi a^3} \frac{1}{(1 + r^2/a^2)^{5/2}}, \tag{5.5}$$

where a the Plummer scale length. Note that there are two characteristic lengthscales in the problem, namely a and b, and we will relate these using a dimensionless parameter $\alpha = a/b$, and we expect that in general $\alpha \leq 1$.

Using Eq. (5.4) we may thus express $r = r(\Psi)$ for the Hernquist profile as

$$r = b \frac{1 - \tilde{\Psi}}{\tilde{\Psi}}, \qquad \tilde{\Psi} = \Psi/\Psi_0. \tag{5.6}$$

The energy part of the distribution function $f_1(\epsilon)$ may now be computed explicitly from Eq. (5.3), using Eqs. (5.5) and (5.6) and after taking the corresponding derivatives, we find

$$f_1(\epsilon) = \frac{3}{8\pi^3 (GMa)^2} \frac{\alpha^4 \tilde{\epsilon}^4}{(1 - \tilde{\epsilon})^3 (1 - 2\tilde{\epsilon} + (1 + \alpha^2)\tilde{\epsilon}^2)^{9/2}} p(\tilde{\epsilon}) \tag{5.7}$$

where $\tilde{\epsilon} = \epsilon/\Psi_0$, $0 \leq \tilde{\epsilon} \leq 1$, and $p(\tilde{\epsilon})$ is the following polynomial

$$\begin{aligned} p(\tilde{\epsilon}) = 30 - 108\tilde{\epsilon} + (132 - 5\alpha^2)\tilde{\epsilon}^2 + 24(-2 + \alpha^2)\tilde{\epsilon}^3 \\ -3(6 + 11\alpha^2)\tilde{\epsilon}^4 + 2(6 + 7\alpha^2 + \alpha^4)\tilde{\epsilon}^5. \end{aligned} \tag{5.8}$$

Figure 5.1 shows the functional form of the distribution function $f_1(\epsilon)$ for different values of α, namely $\alpha = 0.01, 0.1$ and 1, that is, for different degrees of embedding of the stars in the dark matter halo. The distribution function is well-behaved, it is continuous and positive everywhere and has a positive slope, indicating that it is stable to radial modes (see Sec. 5.5 of Binney & Tremaine 2008).

We have checked that the density profile obtained by integrating this distribution function over velocity space returns the Plummer functional form. The left column of Fig. 5.2 shows the velocity dispersion profiles in the radial (solid) and tangential (dashed) directions for different values of α, and makes explicit the dependence of the internal kinematics on the degree of embedding of the stars in the dark halo. For small α (top and middle left panels) the velocity dispersion profile is relatively flat for $r > 0.1r/a$. Since the properties of the halo are fixed by the mass M and the scale b, we note that the velocity dispersion has a smaller amplitude for smaller values of α, as expected.

In the right column of Fig. 5.2 we have plotted the resulting l.o.s. velocity dispersion profiles for the different values of α explored. It shows that these profiles are relatively flat with radius over the range $0 < R/a < 2$, which is similar to that probed by the observations the kinematics of stars in the dSph satellites of the Milky Way. If we set $M = 10^9 \mathrm{M}_\odot$ and $b = 2.5$ kpc, then the system with $\alpha = 0.01$ (top right panel) would have a velocity dispersion of ~ 2 km/s and $a \sim 25$ pc. On the other hand, if $\alpha = 0.1$ (top middle panel), then $a = 250$ pc, and $\sigma_{\rm los} \sim 6.3$ km/s at the centre. This case could for example, represent systems akin the Carina, Sextans or Ursa Minor dSph. A more massive halo would probably be required if the aim is to represent systems like Sculptor or Fornax, as this would allow a better matching to the central value of $\sigma_{\rm los}$.

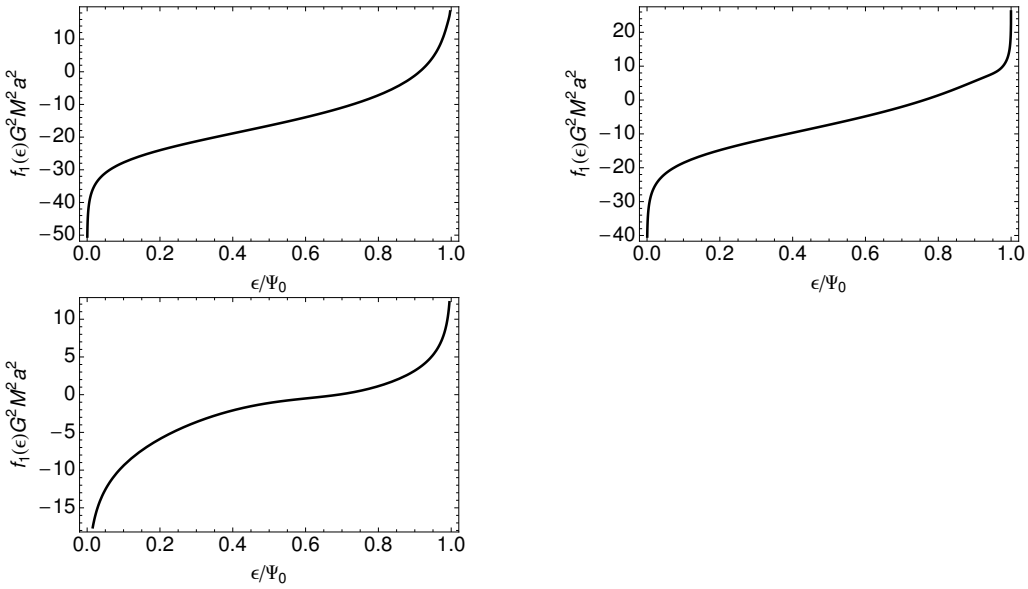


Figure 5.1: Energy part of the distribution function for a Plummer stellar sphere of scale a embedded in a Hernquist dark matter halo of mass M and scale b, and with constant velocity anisotropy $\beta = -1/2$, as given by Eq. (5.7). The various panels correspond to different degrees of embedding $\alpha = a/b$ of the stars in the (same) dark matter halo, $\alpha = 0.01$ (lop left), 0.1 (top right) and 1 (bottom left).

5.3 Conclusions

By example, we have demonstrated that a Plummer stellar system embedded in a Hernquist dark matter halo constitutes a plausible physical configuration. We have explicitly derived the form of the distribution function for the case of a tangential anisotropy $\beta = -1/2$ and for different degrees of embedding of the stars, as quantified by the ratio of scalelengths parameter α. This distribution function is positive for the values of $\alpha = 0.01 - 1$ and also leads to a system that is stable to radial modes, as $\partial f/\partial\epsilon > 0$. The line-of-sight velocity dispersion profiles characteristic of this family of distribution functions resemble those observed for dSph, and hence can be used to represent these systems. They satisfy the An & Evans (2009) theorem, namely that $\gamma_0 > 2\beta_0$, but clearly not the equality condition. We have also explored the $\beta = -1$ case, and also found an analytic physical solution, though this is more cumbersome mathematically and hence not presented here.

Acknowledgments

We are grateful to Tim de Zeeuw for encouragement and useful discussions, and to J. An & W. Evans for various interactions that led to the work presented here. We acknowledge financial support from NOVA (the Netherlands Research School for Astronomy), and European Research Council under ERC-StG grant GALACTICA-240271.

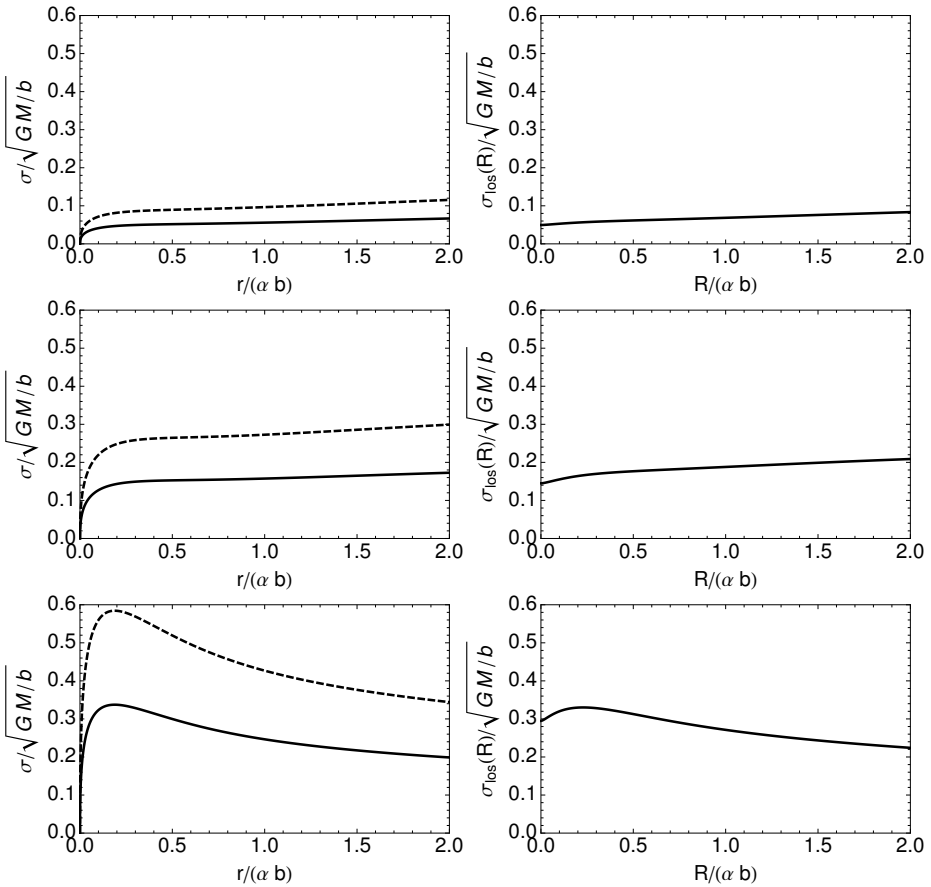


Figure 5.2: Velocity dispersion profiles associated to the distribution functions presented in Fig. 5.1 for different degrees of embedding, namely $\alpha = 0.01, 0.1$ and 1 from top to bottom. **Left column:** Intrinsic radial (solid) and tangential (dashed) velocity dispersions as function of radial distance r/a. **Right column:** Corresponding line-of-sight velocity dispersions as function of projected distance R/a. The resulting $\sigma_{\rm los}$ profiles are relatively flat, especially for small α, i.e. 0.01 and 0.1, and reach a finite value at the centre. In general, the curves bare a good resemblance to the observed velocity dispersion profiles of stars in dSph satellites of the Milky Way, as determined in e.g. Walker et al. (2009).

Bibliography

An, J. H., & Evans, N. W. 2009, ApJ, 701, 1500

Baes, M., & Dejonghe, H. 2002, A&A, 393, 485

Binney, J., & Tremaine, S. 2008, Galactic Dynamics (Princeton University Press)

Ciotti, L., & Morganti, L. 2010, MNRAS, 408, 1070

Evans, N. W., An, J., & Walker, M. G. 2009, MNRAS, 393, L50

Hernquist, L. 1990, ApJ, 356, 359

Irwin, M., & Hatzidimitriou, D. 1995, MNRAS, 277, 1354

Jardel, J. R., & Gebhardt, K. 2012, ApJ, 746, 89

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

Łokas, E. L. 2001, MNRAS, 327, L21

—. 2009, MNRAS, 394, L102

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

Stoehr, F., White, S. D. M., Tormen, G., & Springel, V. 2002, MNRAS, 335, L84

Strigari, L. E., Frenk, C. S., & White, S. D. M. 2010, MNRAS, 408, 2364

Walker, M. G. 2012, arXiv:1205.0311

Walker, M. G., Mateo, M., Olszewski, E. W., Gnedin, O. Y., Wang, X., Sen, B., & Woodroofe, M. 2007, ApJ, 667, L53

Walker, M. G., Mateo, M., Olszewski, E. W., Peñarrubia, J., Wyn Evans, N., & Gilmore, G. 2009, ApJ, 704, 1274

Wilkinson, M. I., Kleyna, J., Evans, N. W., & Gilmore, G. 2002, MNRAS, 330, 778

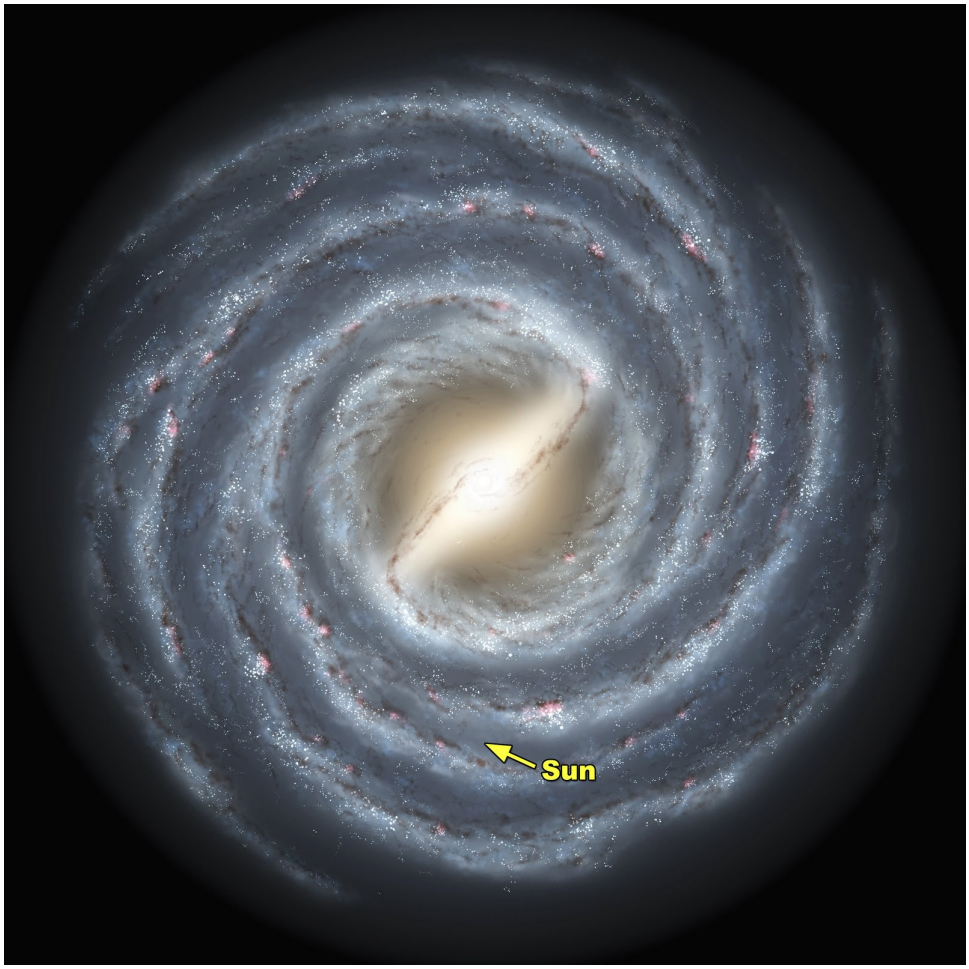
Nederlandse Samenvatting

Van ons zonnestelsel naar de Melkweg

Afstanden in de sterrenkunde worden meestal niet in kilometers of mijlen uitgedrukt. Afhankelijk van de situatie gebruiken we vaker astronomische eenheden, lichtjaren, parsec of megaparsec. In deze samenvatting zal ik lichtseconden, lichtminuten, lichtjaren en kiloparsec gebruiken. In 1 seconde kan licht een afstand afleggen gelijk aan zeven maal rond de aarde. Dat is vrij snel, maar zeker eindig. We kunnen dus afstanden ook uitdrukken in lichtseconden, de afstand die licht aflegt in 1 seconde (bijna 300 000 kilometer). Zo is bijvoorbeeld de afstand tot de maan ongeveer 1.2 lichtseconden. De afstand tot de zon is een stuk groter, iets meer dan 8 lichtminuten. Dat wil zeggen, het licht doet er dus meer dan 8 minuten over om van de zon tot de aarde te komen.

De aarde draait in bijna een cirkelbaan om de zon. De snelheid waarmee we deze baan afleggen, hangt af van de afstand van de aarde tot de zon en haar massa. Dat wil zeggen, als we de snelheid weten waarmee de aarde zich beweegt en we de afstand tot de zon weten, we de massa van de zon kunnen bepalen. Dit is een belangrijk concept. In werkelijkheid hebben de planeten in ons zonnestelsel ook nog invloed op elkaar, maar dit effect is zo klein dat we het kunnen verwaarlozen.

Onze zon bevindt zich in een sterrenstelsel, de Melkweg, en volgt een bijna cirkelbaan rond het middelpunt van ons sterrenstelsel. De afstand van de zon tot het middelpunt is ongeveer 26 000 lichtjaar. Nu is misschien ook duidelijk waarom we vaak overstappen naar andere eenheden. Als we deze afstand uitdrukken in kiloparsec, dan komen we uit op 8 kiloparsec, dat werkt prettiger. Omdat we ons ín de Melkweg bevinden kunnen we deze niet in zijn geheel afbeelden, maar we hebben wel een idee hoe het er van buitenaf gezien uit zal zien. Figuur 1 beeldt de Melkweg uit waarbij de lokatie van de zon staat aangegeven door een gele pijl. In tegenstelling tot een zonnestelsel, waar zich nagenoeg alle massa zich in het middelpunt bevindt (de zon), is de massa in een sterrenstelsel (onder andere de sterren) verdeeld over het hele sterrenstelsel. Waar in een zonnestelsel sprake is van een simpele relatie tussen massa, afstand tot het centrum en snelheid van de planeet is een vergelijkbare relatie in een sterrenstelsel minder eenvoudig. Hoewel het in dit geval minder simpel is, kunnen we hier toch mee rekenen. Als we nu gaan schatten hoe zwaar de sterren zijn en de andere materie tussen de sterren, blijkt dat de er niet genoeg materie is om de bewegingen van de sterren in de nabijheid van onze zon te kunnen verklaren, er mist 50% van de materie. Dit fenomeen vinden we vaker in de sterrenkunde. Er lijkt een systematisch tekort te bestaan aan materie om de bewegingen van sterren en gas, maar ook sterrenstelsels onderling te kunnen verklaren.



Figuur 1: Een *artist's impression* van de Melkweg, onze zon (aangegeven door de gele pijl) staat op een afstand van ongeveer 8 kiloparsec (of 26 000 lichtjaar) van het centrum van onze Melkweg. *Credit: NASA/JPL-Caltech/R. Hurt (SSC/Caltech).*

Donkere materie

Een oplossing voor dit probleem is het postuleren van donkere materie. Dit is materie die niet direct waar te nemen is, maar wel zwaartekracht uitoefent op elkaar en normale materie. Het lijkt misschien een erg simplistische oplossing voor dit probleem, maar blijkt zeer succesvol in het verklaren van bijvoorbeeld de grote schaal structuur in het universum en enkele karakteristieken van de microgolfachtergrondstraling (maar dan dwalen we af). Daarnaast beschikken natuurkundigen over modellen die het bestaan van dit soort deeltjes kunnen verklaren.

Kosmologen kunnen erg goed vooorspellingen maken over hoe de donkere materie zich verdeelt over het universum. Dit doen ze met behulp van simulaties waarbij ze de beweging van heel veel donkere materie kunnen volgen. Plaatsen waar veel donkere materie samenkomt, noemen we halos. Deze simulaties laten zien dat de verdeling van de donkere materie nagenoeg in elke halo hetzelfde is, dus universeel.

In het universum bevindt zich waarschijnlijk meer dan vijf maal meer donkere materie dan normale materie. Aangezien er in het universum zoveel meer donkere dan normale materie is, zal op plaatsen waar heel veel donkere materie samenkomt, de donkere materie de gewone materie aantrekken. De gewone materie zal dichter bij elkaar komen en onder zijn eigen zwaartekracht ineenstorten om sterren te vormen, een sterrenstelsel is daarmee geboren.

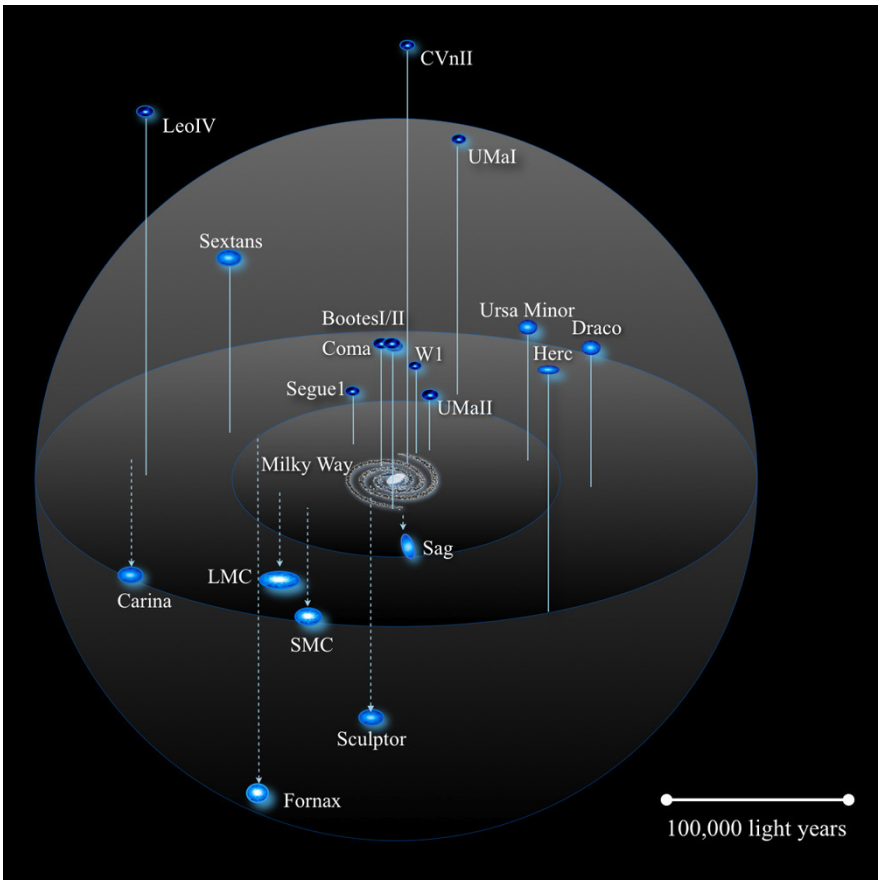
Veel onderzoek heeft zich gericht op het testen of de universele verdeling van de donkere materie ook overeenkomt met de snelheidsverdeling in sterrenstelsels. Het lastige is dat we weten dat in sommige sterrenstelsels de normale materie invloed kan hebben op de verdeling van de donkere materie. In simulaties wordt vaak alleen de donkere materie gesimuleerd, omdat het modelleren van de krachten (of processen) die gewone materie kunnen ondergaan ingewikkeld is. Dit maakt een vergelijking met simulaties zonder deze normale materie soms niet geheel betrouwbaar.

Dwergsterrenstelsels

Ons sterrenstelsel is niet alleen, het wordt omringd door kleinere sterrenstelsels (dwergsterrenstelsels) van verschillende typen en groottes (zie Figuur 2 voor een overzicht van de posities ten opzichte van de Melkweg). De meeste van deze dwergsterrenstelsels bevinden zich in een baan om ons sterrenstelsel, sommige worden opgeslokt en andere passeren simpelweg. Een bepaald type dwergsterrenstelsels, sferoïde[1] dwergsterrenstelsels (in het engels *dwarf spheroidal galaxies*), zijn bijna bolvormig en bevatten geen of weinig gas. Dit maakt het relatief makkelijk om deze sterrenstelsels met een model te beschrijven of te vergelijken. Een opname van een in dit proefschrift veel gebruikt sterrenstelsel is weergegeven in Figuur 3. Aan dit figuur is te zien dat dit sterrenstelsel weinig bijzondere kenmerken heeft en een bijna ronde vorm. Dit dwergsterrenstelsel bevindt zich op een afstand van ongeveer 80 kiloparsec. Binnen een straal van een halve kiloparsec bevindt zicht meer dan 50% van alle sterren. Dit sterrenstelsel is dus vrij klein vergeleken met de Melkweg.

Aangezien deze sterrenstelsels relatief dichtbij zijn, is het mogelijk om de snelheid van individuele sterren te bepalen. Als we het licht van een ster splitsen met bijvoorbeeld een prisma, zullen we zien dat er zich vaak donkere lijnen bevinden op bepaalde kleuren, of golflengtes. Dit komt omdat er chemische elementen in de atmosfeer van de ster zitten die het licht van die golflengte zal absorberen. Van metingen uit het laboratorium of van natuurkundige modellen weten we vaak precies op welke golflengtes deze donkere lijnen horen te zitten. Echter, de sterren bewegen zich ten opzichte van de aarde. Hierdoor zullen deze donkere lijnen iets verschuiven naar langere golflengtes als ze van ons af bewegen en naar kortere golflengtes als ze naar ons toe bewegen. Een vergelijkbaar effect bij geluid is hoorbaar wanneer een ambulance iemand inhaalt.

[1] Een sferoïde is een enigszins platgedrukte of uitgerekte bol

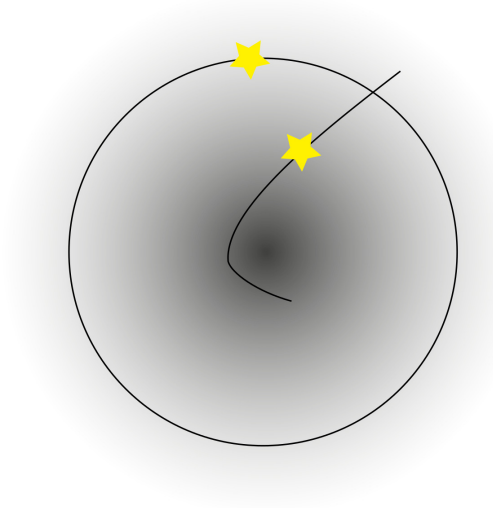


Figuur 2: Een overzicht van de omgeving van de Melkweg. De schaal rechtsonder geeft aan hoeveel 100,000 lichtjaar is (ongeveer gelijk aan 30 kiloparsec). Sculptor bevindt zich op een afstand van ongeveer 80 kiloparsec. *Credit: J. S. Bullock.*

Naar de persoon toe zal de toonhoogte hoger zijn dan wanneer de ambulance van de persoon af beweegt. Uit de verschuiving van deze golflengtes kan dus de snelheid van de sterren afgeleid worden.

Van veel van de sterren in lokale dwergsterrenstelsels zijn snelheden gemeten (voor Sculptor ongeveer 2000). Uit deze snelheden kunnen we afleiden dat er ongeveer 100 maal meer massa lijkt te zijn dan dat we zien in de vorm van sterren. In de Melkweg, rond de positie van de zon, is dit een gelijke hoeveelheid. Ook bij andere sferoïde dwergsterrenstelsels zien we vergelijkbare discrepanties tussen de snelheden van de sterren en de hoeveelheid normale materie. Omdat zich in deze sterrenstelsels dus zoveel meer donkere materie dan normale materie bevindt, kunnen we gerust zeggen dat de invloed van de sterren klein is. Het is bij deze sterrenstelsels dus zeer waarschijnlijk dat de verdeling van de donkere materie goed overeenkomt met de universele vorm die similaties voorspellen, *als de theorie klopt.*

Figuur 3: Een opname van het nabije (ongeveer 80 kiloparsec) dwergsterrenstelsel Sculptor (beeldhouwer in het Nederlands), met weinig bijzondere kenmerken en een bijna ronde vorm. *Credit: David Malin, Anglo-Australian Observatory.*



Figuur 4: Een schematische weergave van twee banen in een sterrenstelsel. De zwarte cirkel geeft de cirkelbaan aan van de ene ster. De andere lijn geeft een excentrisch baan weer. De donkere kleur geeft de donkere materie weer.

Dit proefschrift

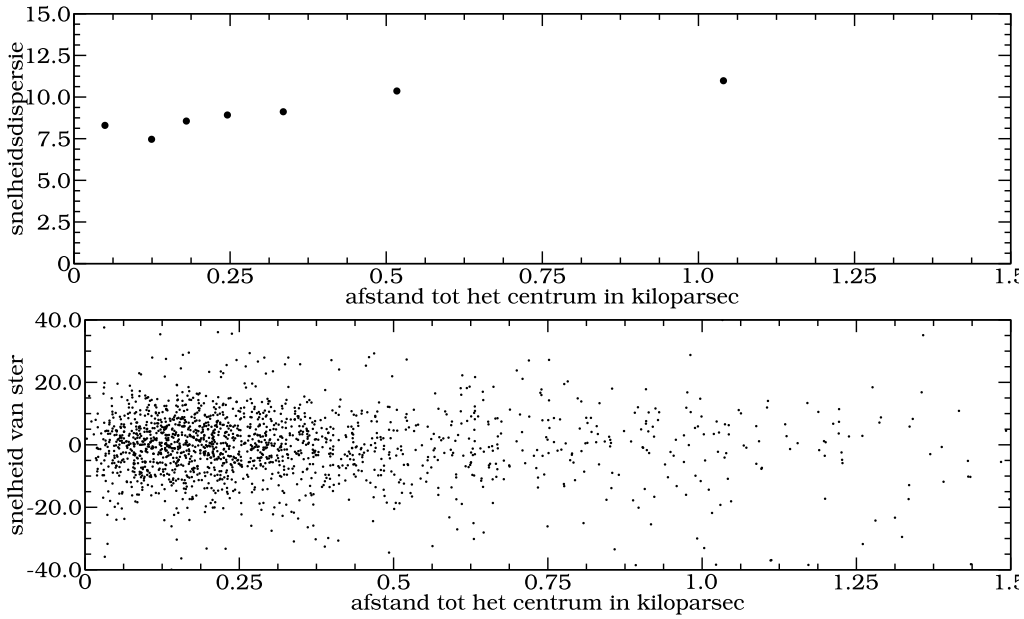
In dit proefschrift hebben we nieuwe modellen ontwikkeld om de waarnemingen van sferoïde dwergsterrenstelsels te verklaren. Met deze modellen kan getest worden wat voor verdelingen van de donkere materie passen bij de waarnemingen van een sterrenstelsel. Aangezien de donkere materie modellen hier voorspellingen over doen, willen we dit valideren met de waarnemingen in het universum. Misschien kloppen de aannames niet die we maken in onze donkere materie modellen. Het kan zelfs zo zijn dat donkere materie niet bestaat en er een andere oplossing nodig is om de hoge snelheden in deze sterrenstelsels te verklaren. Naast de informatie over de verdeling van de donkere materie, leren we ook welke banen nodig zijn om deze waarnemingen te verklaren. De samenstelling van deze banen kan ons eventueel iets vertellen over de ontstaansgeschiedenis van het sterrenstelsel.

Eerdere modellen moesten aannames maken over wat voor banen nodig waren om deze sterrenstelsels te kunnen verklaren, wat kan leiden tot verkeerde inzichten. De modellen die wij gebruiken bestaan uit een verzameling banen, niet alleen cirkelbanen, maar (bijna) alle mogelijke banen die kunnen bestaan in dit soort sterrenstelsels, zogenaamde excentrische banen. In Figuur 4 geven we schematisch 2 banen weer, een cirkelbaan en een excentrische baan. Met deze methode neem je eerst een verdeling van de donkere materie aan, deze verdeling van materie defineert het krachtenveld. In dit krachtenveld laat je een testdeeltje voortbewegen die een ster voor moet stellen en je houdt bij wat voor observaties je bij deze baan zou verwachten. Doe je dit met heel veel banen, dan kan je al deze banen als het ware bij elkaar optellen, met een bepaald gewicht, om daarmee je waarnemingen te kunnen verklaren. Dit zal niet lukken voor elke verdeling en hoeveelheid van donkere materie, als bijvoorbeeld de massa te laag is zullen de snelheden te laag zijn en vice versa.

In hoofdstuk 2 passen we deze methode toe op het sferoïde dwergsterrenstelsel Sculptor. Hoewel deze methode al lange tijd werd toegepast op andere soorten sterrenstelsels, was deze weinig toegepast op specifiek deze sterrenstelsels. Nadat wij deze methode uitvoerig hebben getest op gesimuleerde waarnemingen hebben we deze op Sculptor toegepast. De hoeveelheid donkere materie komt in grote lijnen overeen met wat anderen hebben gevonden met de andere (meestal simpelere) modellen. Dit is geruststellend, aangezien in veel gevallen deze simpele modellen je correcte antwoorden geeft over de hoeveelheid donkere materie binnen een bepaalde straal. Wat nieuw is, is dat deze baan-gebaseerde modellen je vertellen uit wat voor banen het sterrenstelsel is opgebouwd. Dit kan van belang kan zijn voor het begrijpen van het onstaan van deze sterrenstelsels. Hiernaast hebben we laten zien dat, in tegenstelling wat andere auteurs beweren, de verdeling van de donkere materie toch consistent kan zijn met wat donkere materie simulaties voorspellen.

In hoofdstuk 3 passen we dezelfde modellen toe op drie extra sterrenstelsels: Fornax, Carina, Sextans. We vergelijken veel verschillende donkere materie verdelingen, maar vinden dat er erg lastig onderscheid te maken is hiertussen. Wanneer we deze modellen met elkaar vergelijken, zien we dat hoewel ze in het binnengebied en buitengebied erg kunnen verschillen, er een straal is waarop alle modellen bijna dezelfde eigenschappen hebben. Een van die eigenschappen was al bekend, de massa binnen deze karakteristieke straal was bijna onafhankelijk van het model. Wat wij hebben gevonden is dat ook de verdeling rond dit punt statistisch niet veel verschilt per model. Waar veel onderzoek zich richt op de binnengebieden, waar erg veel dissussie over is, stellen wij voor dat het zinniger is om de vorm van de donkere materie op deze karakteristieke schaal te bekijken. Het is immers makkelijker om iets te vergelijken wat goed te meten is, dan iets wat erg lastig te meten is.

In hoofdstuk 4 verfijnen we ons model. In hoofdstuk 2 en 3 vergelijken we onze modellen met zogenaamde momenten van de snelheidsverdeling, statistieken die de vorm van de snelheidsverdeling beschrijven. Hoewel deze statistieken erg handig zijn in de analyse, is het statistisch simpeler en correcter om elke individuele ster te gebruiken. Op deze manier gaat er minder informatie verloren en wordt dus de maximale informatie uit de gegevens gehaald. Dit verschil



Figuur 5: In het onderste paneel geven we de waarnemeingen weer. Elk punt geeft de snelheid en afstand tot het centrum van een ster weer. De dispersie van de snelheidsverdeling per groep is weergegeven in het bovenste paneel. (De eenheid van de verticale as is in beide gevallen in kilometers per seconde.)

maken we duidelijker in Figuur 5. In het onderste paneel geven we de waarnemingen weer. Elk punt geeft de snelheid en afstand tot het centrum van een ster weer. Wanneer we nu sterren die op bijna dezelfde afstand tot het centrum staan groeperen, en de snelheidsverdeling van deze groep beschrijven met een statistiek (voor de geïnteresseerden: de dispersie), krijgen we het bovenste paneel. Hoewel deze statistiek een zeer goede beschijving is van de snelheidsverdeling in het onderste paneel van deze figuur, bevat dit onderste paneel toch altijd meer informatie. In hoofdstuk 2 en 3 hebben we met de gegevens van het bovenste paneel gewerkt, de statistieken die de snelheidsverdeling beschrijven, terwijl we in hoofdstuk 4 met de gegevens van het onderste paneel werken, de complete snelheidsverdeling.

Na het wederom uitgebreid testen van deze methode, passen wij deze toe op het sterrenstelsel Sculptor. Hoewel we een iets andere verdeling van de donkere materie vinden dan in de vorige hoofdstukken, is het (statistisch gezien) in overeenstemming met de vorige resultaten. Wanneer wij kijken naar het soort banen waarmee dit sterrenstelsel wordt beschreven, zien we dat de verzameling banen als het ware twee groepen vormt. Een groep met cirkelbanen en licht excentrisch banen en een tweede groep met voornamelijk meer excentrische banen. Het sterrenstelsel lijkt dus uit meerdere componenten te bestaan. Wanneer we deze twee groepen apart bekijken vinden we een grote gelijkenis met de twee componenten gevonden door middel van scheiding in chemische eigenschappen van de sterren. Met deze methode hebben we laten zien dat deze modellen de twee componenten waarschijnlijk al weten op te sporen zonder de informatie over de chemische samenstelling van de twee componenten. Het gebruiken van alle individuele sterren zal belangrijk zijn voor toekomstige ingewikkeldere modellen, welke het steeds lastiger zullen maken om te werken met de momenten van de snelheidsverdeling. Het aantonen van de werking van deze modellen is hiervoor een eerste aanzet.

Onze modellen zijn gemaakt zodat ze altijd fysisch zijn en dus kunnen bestaan in de realiteit. Desondanks ontvingen we enige skepsis of onze modellen daadwerkelijk fysisch waren en de numerieke aard van onze modellen dit misschien verborg. Om alle twijfel weg te nemen hebben we in hoofdstuk 5 een wiskundig model gepresenteerd, welke deze onzekerheid moet wegnemen. Het is een vergelijkbaar, maar minder flexibel model, wat overeenkomstigheden heeft met de modellen in hoofdstukken 2, 3 en 4. Hiermee hebben we dus aangetoond dat de modellen van hoofdstuk 2, 3 en 4 zonder twijfel een echt sterrenstelsel kunnen beschrijven. Hiernaast is een wiskundig model erg handig om simulaties van sferoïde dwergsterrenstelsels op te zetten.

Het belangrijkste hoofdstuk voor toekomstige ontwikkelingen is hoofdstuk 4. De modellen in dit proefschrift zijn allemaal sferisch, maar in Figuur 3 is te zien dat dit stelsel niet bolvormig is, maar iets afgeplat. De techniek om individuele sterren te gebruiken in plaats van de statistieken, maakt het mogelijk om deze niet-bolvormige modellen op een goede manier te kunnen vergelijken met de waarnemeingen. Hiernaast zou het interessant zijn om te resultaten over de verdeling van de donkere materie van hoofdstuk 3 te kunnen vergelijken met de voorspellingen van donkere materie simulaties, rekening houdend met eventuele effecten van de normale materie.

Acknowledgments

A thesis is more than the science done in four years. At some moment finishing the thesis is a goal in itself, and finishing it seems more like a product of suffering, persistence, motivation and discipline rather than science. Being able to share some of the struggles with colleagues and friends made it possible to actually get this far, the acknowledgments. Rense (Boomsma) his future legendary words "Whatever you do, just finish the damn thesis." was my mantra during the last few months.

The person that made my PhD possible was also the most valuable to my thesis, my promoter. Amina, thank you for dragging me through that horrible, unfair and unprofessional refereeing process. I also hope my stubbornness wasn't too annoying too you. My dutch genes make me completely ignore titles such as "Prof.", so that I refuse to accept anything from anyone unless I also believe it to be true. If I got you completely frustrated by your inability to steer me, you can share those frustrations with my mother or my girlfriend. I learned a lot from you over the years, and not just about science and writing. With the dreadful referee reports, the papers you wrote and the work we did together, I especially learned one valuable thing: Persistence rewards. I tend to give up sooner than you, and on numerous occasions you showed me that digging deeper, looking further or thinking longer usually pays off.

The computer group, Martin, Eite and Wim, without you running for instance my many models using Condor would not be possible. The homogeneous platform makes this possible, and this is one of the big strengths of the institute, and as I heard from ex-Kapteyners not standard at all institutes. I thank especially Martin, who accepted me for a small internship, 13 years ago, which marked the start of this crazy endeavor, together with Danny, who sparked my enthusiasm to go to university. Hennie, Jackie, Gineke, Christa and Lucia, you basically run the institute such that other people can do science, and you all do it well and cheerful.

German, I heard you appreciate to have your own sentence, here you go, maybe even two. Thanks for being a cool roommate, and having some decent wall decoration! Carlos, I enjoyed our travels together, especially the one in Bologna, sharing a room with me while having food poisoning without complaining, respect army guy. I hope you didn't find that Nyan-Nyan-cat tune in Santa-Barbara too gay though. Else, from my first year as a student till almost the end of my PhD you have been around, your everlasting smile was always good for my mood. And thank you Tjitske for taking over that job. Shoko, thanks for showing me what fish can taste like. Jakob, thanks for sharing your C++ passion and knowledge with me. Harish, man you made me embrace my inner geek, no matter what kind of geeky stuff I come up with, you think it is a great idea. You are a danger to my spare time, if it was up to you I think I would not have finished this thesis, but would have developed software to create a complete thermodynamic model of my house. Your enthusiasm is encouraging. Patrick en Giacomo, thanks for sharing some many coffee brakes with me, I hope the new coffee machine doesn't destroy that. Giacomo,

your almost anarchistic political views nicely balance the average hippy-lefty-communist (me included) view that is more common at the institute, although I fully disagree with you. Hans, make sure you never loose your alternative way of looking at things. During classes, I noticed you often had questions that I thought were strange at first. But you were simply thinking outside of the box. And thanks for trying to cheer me up during difficult times. Jouke, why didn't you do a PhD at Kapteyn, your politically incorrectness was missed. Thanks to all the other (ex) Kapteyners for making the institute a nice working place, in particular Ajinkya, Aleksander, Beth, Burcu, Chiara, Derek, Eva, Facundo, Gergo, Giuseppina, Hao, Hugo, Jeffrey, Johan, Koshy, Kyle, Laura, Leon, Marius, Parisa, Robyn, Stephan, Teresa and Thomas. Martin (Smith), thanks you for being so helpful, from a supervisor during my bachelor research to a tour guide in Cambridge. Glenn and Remco, thanks for helping me start up this project, and having me over in Heidelberg. Even though you probably expected this project to take a different path, I hope you like the outcome.

Pap, bedankt dat je me alles hebt laten doen wat ik maar wilde, als kind mocht ik van op het dak klimmen, met een kettingzaag spelen tot het verbouwen van groente. Ik heb hierdoor geleerd dat als je er zin in hebt bijna alles kan, dus ook een promotieonderzoek. Mam, bedankt voor het nu al 31 jaar klaar staan voor me, en mij te stimuleren om sterrenkunde te studeren. Ragna, ik had het niet beter kunnen treffen, bedankt voor je steun, luisterend oor en liefde. Ik hoop dat we het de komende jaren samen in Ezinge net zo leuk zullen hebben als de afgelopen jaren. Als laatste bedank ik al mijn familie en vrienden voor de afleiding, stimulering, zingeving en leuke tijden, in het bijzonder Alle, Bonny, Daisy, Martijn en Link, Jac en Vera, en Mario.

Maarten