5 Discrete dynamical models of dwarf spheroidal galaxies

M. den Brok, G. van de Ven, et al.

— In preparation —

ABSTRACT

We construct non-spherical dynamical models of four dwarf spheroidal galaxies, using a solution of the axisymmetric Jeans equations. We show how instead of the usual fitting of binned data, discrete dynamical modeling based on likelihood fits to individual stars yields significant improvements on the model parameters, in particular for the velocity anisotropy. Moreover, we show how discrete modeling allows for pruning of possible non-members from a sample of stars, and secondly, splitting different dynamical tracers by their metallicity without adding hard cuts.

5.1 Introduction

It is important to constrain the amount of mass in and shape of dark matter halos, for a number of reasons. Measuring dark matter can test our knowledge of galaxy formation and evolution (Strigari et al., 2008).

A longstanding debate is whether dark matter halos have cusps or cores. Early numerical simulations of dark matter, without the inclusion of baryons, pointed at universal dark matter profiles with a strong central cusp (Navarro et al., 1996), confirmed by recent very high resolution N-body simulations (Navarro et al., 2010).

Measurements of the dark matter density are also important for the identification of possible candidates for dark matter paritcles from annihilation signals (e.g. Abramowski et al. (2012)). As the latter process is proportional to the squared density of dark matter, it needs to be known with high accuracy.

In this paper, we model four classical Local Group dwarf spheroidals with axisymmetric models. Most authors have modeled Local Group dSphs with spherical models (but see Jardel & Gebhardt (2012)). However, dSphs are not spherical. Measurements of the ellipticities of the four dSphs in this paper point at ellipticities of $\epsilon \approx 0.3$ (Irwin & Hatzidimitriou, 1995). Moreover, it is predicted from numerical simulations that the density generating the potential – at least on large scales – is not spherical (Frenk et al., 1988; Bett et al., 2007). With the advent of high-quality publically available spectroscopic data (Walker et al., 2008), it is worth to try to loosen the constraint of spherical symmetry.

Another issue is that for some of the dSphs, there are a few hundred observations of radial velocities of stars which are probably members. In the past, people have binned these data to obtain velocity dispersion profiles. This way, one looses information in two ways: one looses information on individual velocities with errorbars, which are averaged together into a single dispersion with an errorbar, and one looses spatial information. Another consideration is that by binning the data, one always has the risk of including outliers or excluding proper members. In this paper we therefore try to develop a novel way to analyze kinematic data, based on likelihoods for individual stars.

The structure of this paper is as follows: The kinematic and photometric data and the method to solve the axisymmetric Jeans equation are described in section 5.2. Section 5.3 applies this modeling to Local Group dSphs, followed by section 5.4, which shows the improvement with discrete modeling. After that, we provide two applications: identification of non-members, and splitting of metallicity. We end with a short discussion and conclusions.

5.2 Data and modeling method

5.2.1 Photometric data

For our analysis, we need to have the luminosity profile of the galaxy written as a sum of Gaussians. We use the photometry of Irwin & Hatzidimitriou (1995). These authors count the number of stars as a function of radius, after which they fit a one-component King model to the data. The data are normalized to L_V/pc^2 by integrating the King model magnitude with the observed magnitude. There is however considerable uncertainty in this quantity. We fit a MGE expansion to the logarithmically sampled King profile. We assume that the ellipticity of each gaussian is fixed to the global ellipticity of the galaxy. The profiles of the dwarfs, together with the star counts and background-corrected star counts from Irwin & Hatzidimitriou (1995) are shown in Fig. 5.1.

5.2.2 Line-of-sight velocity data

For the analysis in this paper, we use the data from Walker et al. (2009), which were obtained as part of a survey (Walker et al., 2007) of Local Group galaxies. These data were taken with the Michigan/MIKE Fiber System (MMFS) at the Magellan 6.5 m Clay Telescope at Las Campanas Observatory. Besides radial velocities and errors, the data include measurements of two magnesium and several iron indices in the 5140–5174 Å spectral range. Walker et al. (2009) analyzed these data with the Expectation-Maximization algorithm to determine galaxy membership probabilities for each individual star. The radial velocities as a function of semi-major axis position are plotted in Fig 5.2. Walker et al. (2008) determine global proper motions of the four dSphs. Assuming no internal streaming motions, perspective rotation can give rise to velocities up to 3km/s in the outer parts of Fornax, and also for the other dwarf spheroidals, this effect is not unimportant. We correct for perspective rotation using equation 6 in van de Ven et al. (2006).

5.2.3 Axisymmetric Jeans solution

The distribution function (DF) of a relaxed stellar system in a smooth potential follows the Collisionless Boltzmann equation (Binney & Tremaine, 1987). In principle, one would like to recover the potential of a system by measuring the DF of a system. However, in practice this is almost never possible. Since velocity moments are much easier to measure, one often uses solutions of the Jeans equations(Jeans, 1915), which are derived by taking velocity moments of the Boltzmann equation. In the axisymmetric case, there are two Jeans equations, which relate the gravitational potential and the tracer density to four second moments of the velocity. This means that at least two additional assumptions have to be made to solve the Jeans equations. We follow Cappellari (2008) in assuming that the velocity ellipsoid is aligned with the axisymmetric coordinate system, and that the ratio of the radial and

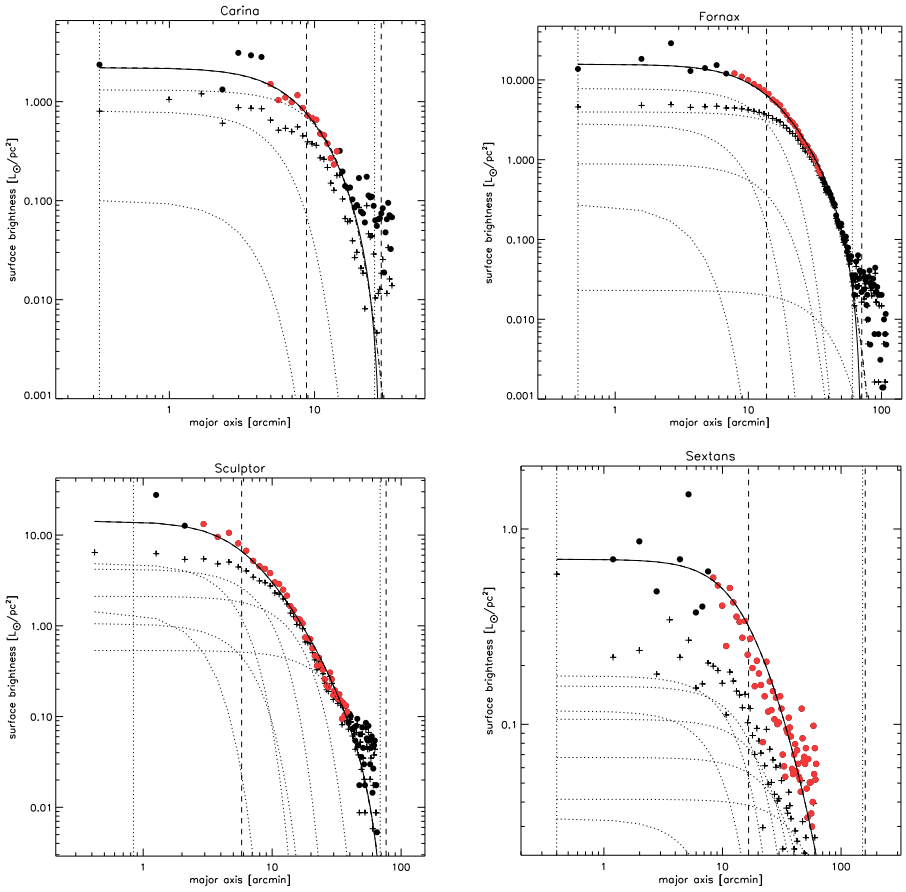

Figure 5.1 — Multi-Gaussian Expansion of the radial surface brightness profile of dSph galaxies. The solid dashed curve is the best-fit King profile from Irwin & Hatzidimitriou (1995), fitted by a MGE model with individual Gaussians shown as dotted curves and the sum as the solid curve. The crosses are the tabulated observations from Irwin & Hatzidimitriou (1995), whereas the solid circles are the observations approximately corrected for crowding effects, and in amplitude matched to the King profile based on the median of the red circles.

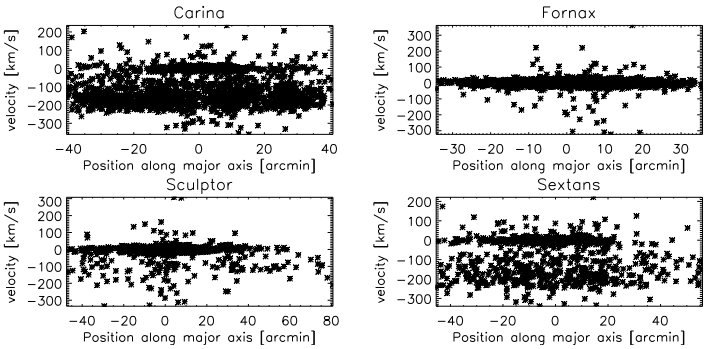

Figure 5.2 — The observed line-of-sight velocities for stars in different dwarf galaxies as a function of major axis position. All data were measured by Walker et al. (2008).

vertical second moments is constant. Furthermore, representing the potential and tracer density by Gaussians reduces the solution for the oberved velocity moment to a single integral (see Cappellari (2008)). *It should be noted that the choice of the velocity ellipsoid, although a good approximation for disky systems, may not be a good approximation for spheroidal systems of which a lot of the mass is outside the z=0 plane:* with the alignment of the velocity ellipsoid to the axisymmetric coordinate system, it actually never points to the center at positions outside the equatorial plane.

5.2.4 Implementation

We have converted the IDL JAM-code (Cappellari, 2008) to *C* and coupled it with a Markov-Chain Monte Carlo (MCMC) engine. In general we multiply the components of the luminous MGE with a 'M/L per gaussian' to obtain the mass MGE which generates the potential. This way we have the freedom of creating any kind of M/L profile we want. The velocity anisotropy can be altered for each luminous gaussian, however, in this paper, we use a global value for the anisotropy. In general we will keep the axis ratio of the mass MGE fixed to the axis ratio of the luminous MGE, but we will investigate the effect of a different flattening for the potential in section 5.3.2.

5.3 Binned data

5.3.1 Spatial Voronoi-bins

The data of W09 come with a pre-assigned membership probability. For each galaxy, we selected all stars with membership probabily greater than $P_m > 0.99$. We averaged the parameters of stars appearing twice in the list, and converted the

coordinates of the stars to a system in which the semi-major and semi-minor axes were aligned with the x and y-axes. The positions of the stars were 'pixelated' by putting them in bins with size of 1 arcmin^2. In case of multiple stars per bin, the velocity and error on the velocity were determined by a weighted mean, where the weighting was done according to the velocity uncertainty of individual stars multiplied with their relative membership probabilities. The collections of pixels were binned with the voronoi binning code of Cappellari & Copin (2003), where the signal-to-noise per pixel was defined as the number of stars divided by the square root of that number. We aimed for bins with approximately 20 stars, so we defined the target signal-to-noise as $\sqrt{20}$. Due to the low number of stars of Sextans, we aimed there for 10 stars per bin. After assignment of stars to the voronoi bins, the velocity and velocity dispersion per bin were calculated with the method given in van de Ven et al. (2006). The bins, coloured per different velocities are shown in the upper panels of Figs. 5.3–5.6.

We fit the kinematic bins with our axisymmetric Jeans models with only a small number of free parameters. We leave the inclination i free to vary between 40 and 90 degrees (90 degrees being edge-on). Lower values for the inclination are usually not allowed by the observed ellipticity. The constant anisotropy, β_z is allowed to vary between -1.0 and 0.5. Instead of a constant mass to light ratio, we allow the M/L to vary, using a similar parametrisation as the Osipkov-Merritt profile:

$$\Upsilon(r) = (\Upsilon_\infty - \Upsilon_0)\frac{r^2}{r^2 + r_b^2} + \Upsilon_0 \tag{5.1}$$

where the free parameters are Υ_0 and Υ_∞, the M/L ratios in the center and at infinity. For r_b, the break radius, we choose the geometric mean of the lowest and highest value of the sigma in our MGE expansion. This is of course a very contrived parametrisation, but it is sufficient for our current purposes.

5.3.2 Binned results

The results of fitting the binned data are shown in Figs. 5.3–5.6. The middle panel of these plots shows the best fitting model and the comparison of this model with the data. In the lower panel, we show the M/L ratio as a function of radius. It becomes now apparent that the parametrization was not so important. Except for Fornax, we confirm what spherical studies already found: the mass to light ratios appear to be too high to be explained by stellar matter only, therefore, by just looking at these plots one could conclude that there is a significant amount of dark matter present in these galaxies. The lower right panel shows the enclosed projected and deprojected mass. The value for the enclosed mass within the effective radius for Fornax is consistent with what was found by Jardel & Gebhardt (2012) (taking into account that we have assumed a slightly closer distance towards Fornax and have a applied a different cut in membership probability).
We have left the inclination as a free parameter, however, we are not able to constrain

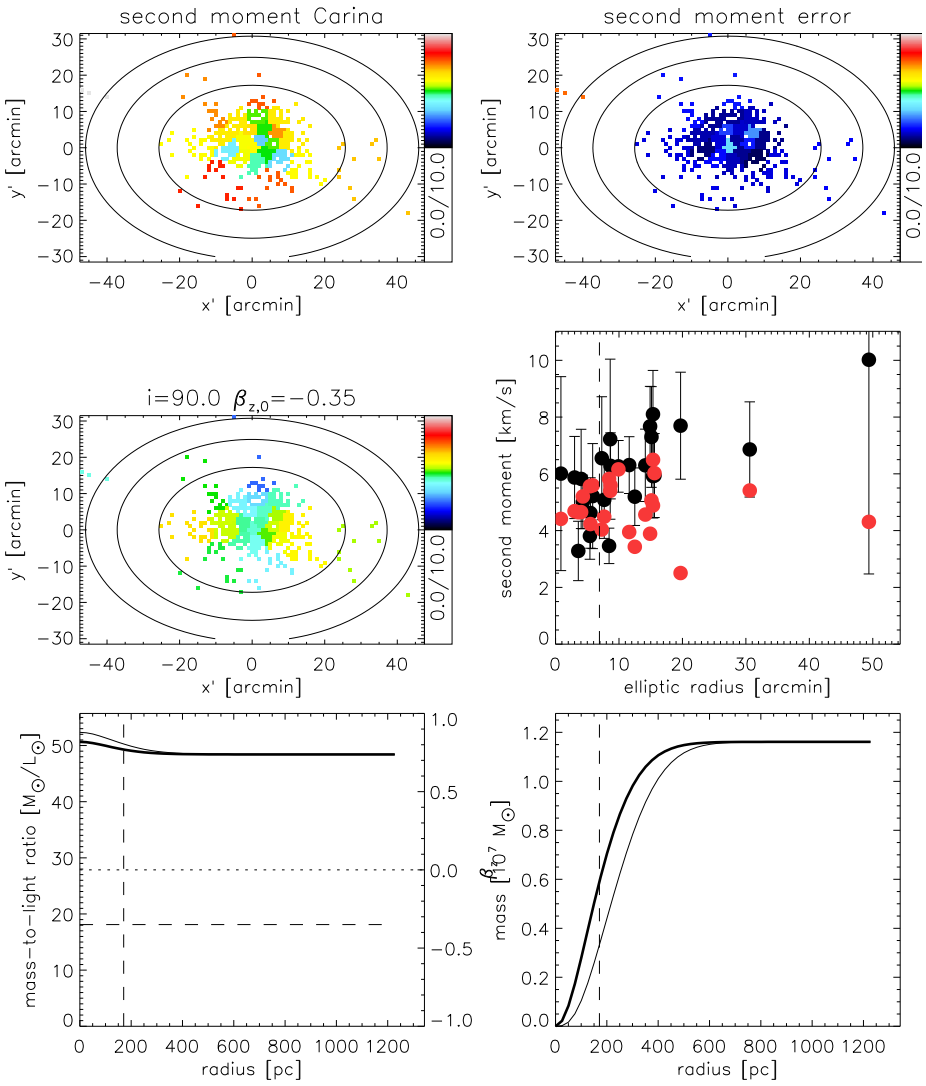

Figure 5.3 — Data and model of Carina. The upper two panels show the kinematic bins with velocity dispersion and respective errors. The left panel in the middle shows the best fit kinematic model. The right panel compares the predicted second moments with the observed second moments, sorted by ellitical radius. The lower-left panel shows the M/L profile and beta as a function of radius, whereas the lower-right panel shows the enclosed mass un-projected (thin line) and projected (thick line)

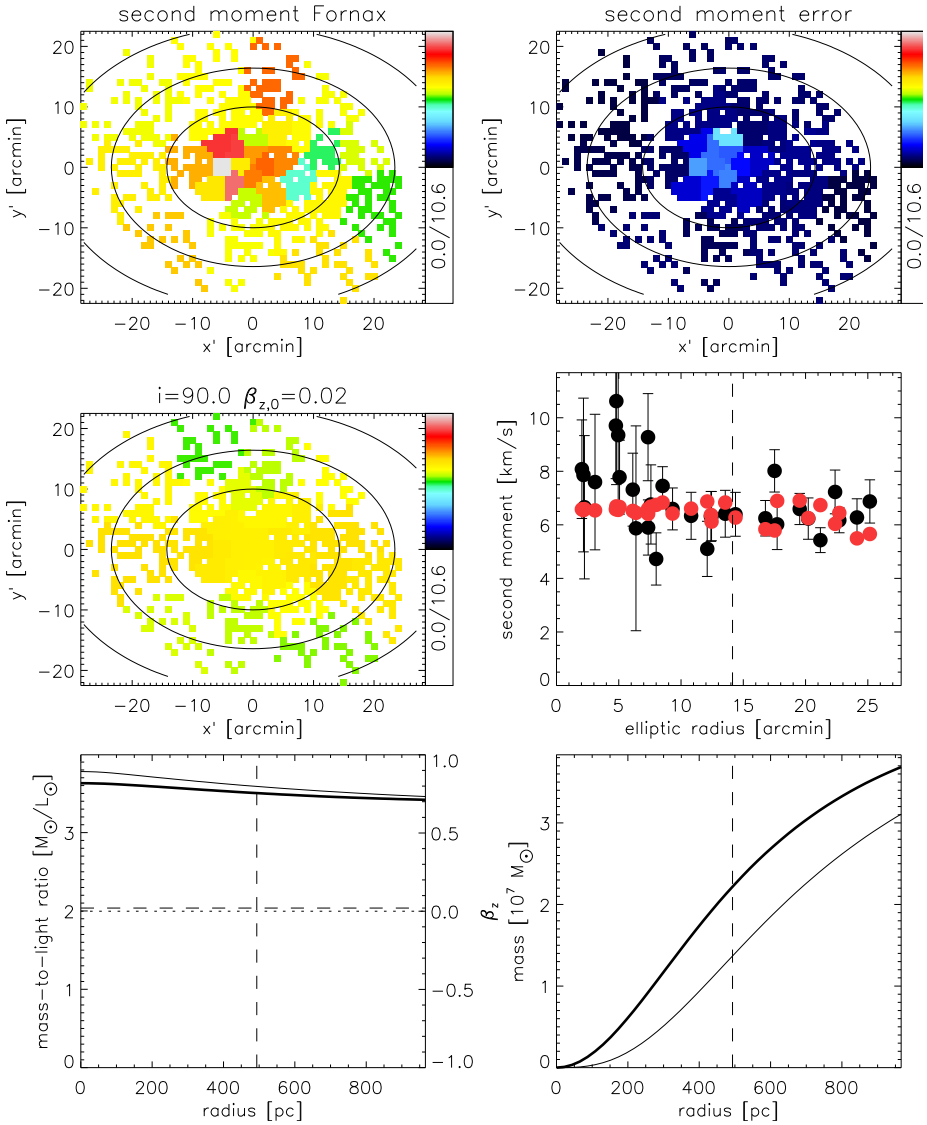

Figure 5.4 — Same as Fig 5.3 for Fornax

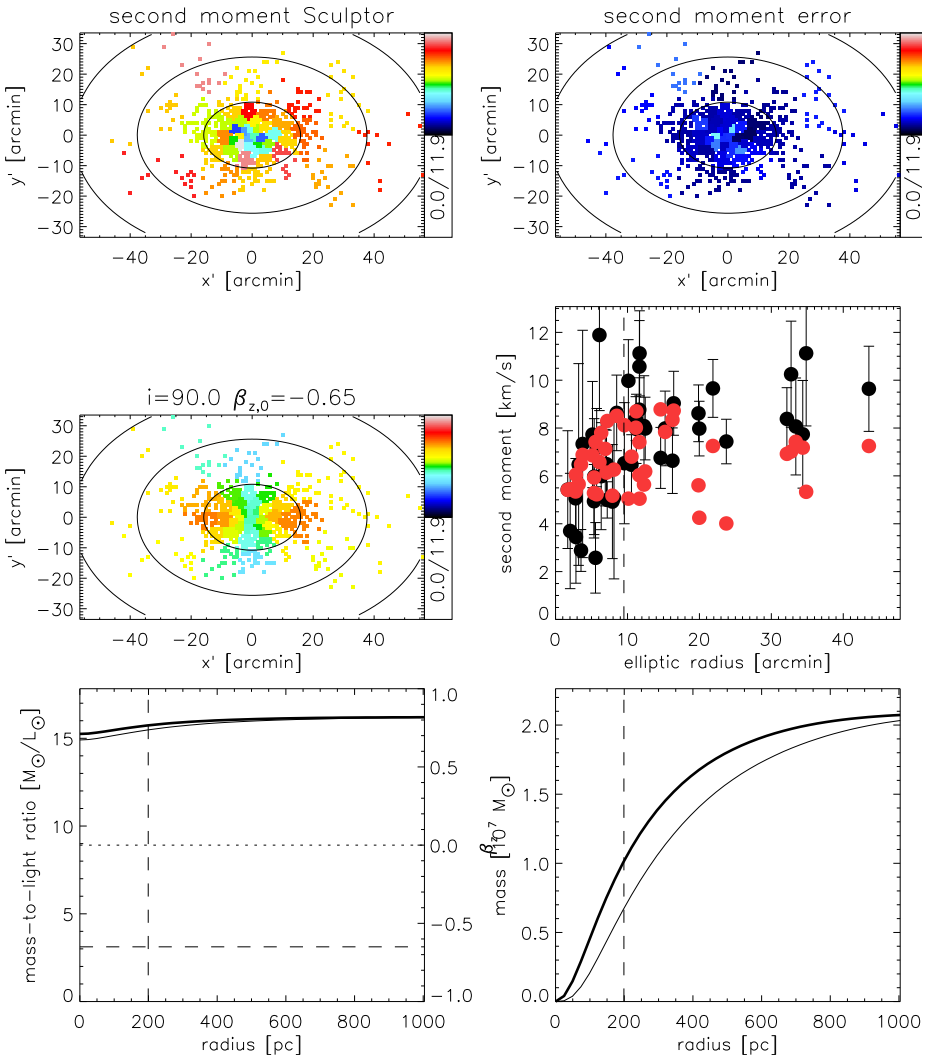

Figure 5.5 — Same as Fig 5.3 for Sculptor

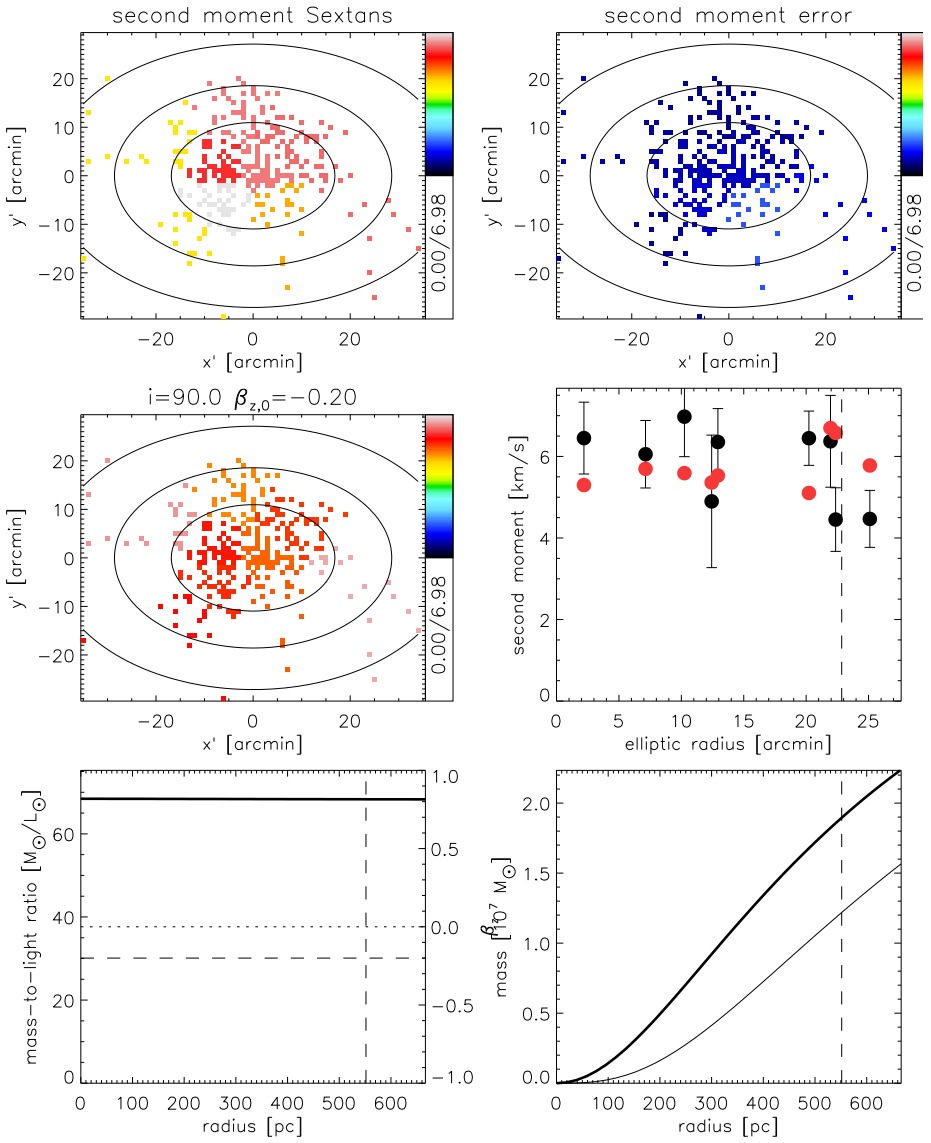

Figure 5.6 — Same as Fig 5.3 for Sextans

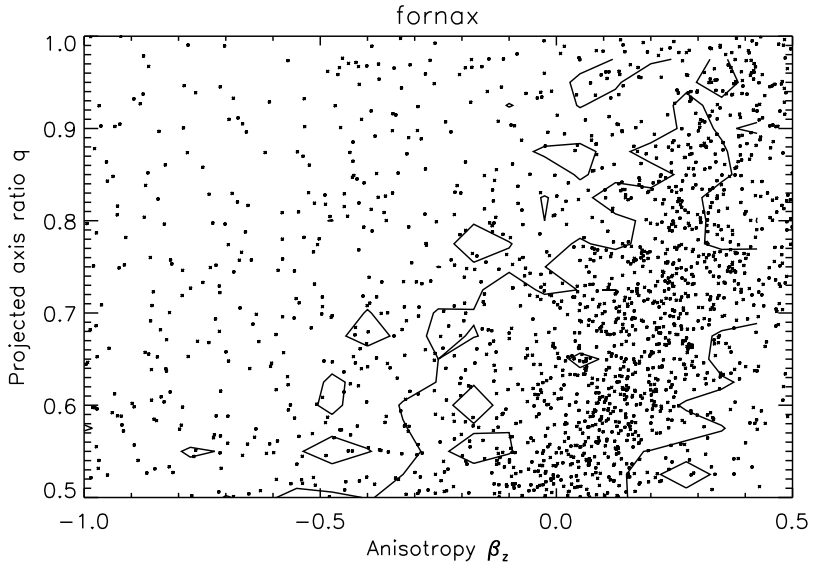

Figure 5.7 — The effect of varying the axis ratio of the potential. We show the projected output of our MCMC simulation for Fornax in the orbital anisotropy–axis ratio plane. The contours denote the region containing more than 66% of the probability. There is a clear degeneracy between the axis ratio of the potential and the anisotropy.

it. The solutions for all four dwarf galaxies tends to go to an inclination of 90 degrees. Most likely one will need proper motions to determine the inclination, as was done for globular clusters.
The anisotropies differ between dwarf galaxies. Fornax seems to be more or less isotropic, whereas the other galaxies have negative anisotropies, in particular Sculptor. Interpreting this is however very difficult. As an example, we have done a similar fit as above for Fornax, but now also leaving the axis ratio of the mass density free. From Fig 5.7 we see that the anisotropy can change substantially by chancing the shape of the potential.

5.4 Discrete data

The Jeans modeling as described above is perfectly suited for observations in which one directly measures the line-of-sight second moment as a function of position in the galaxy. However, for a number of systems, such as Milky Way globular clusters or Local Group galaxies, velocity measurements of individual stars are available. In the past, several authors have converted these measurements to line-of-sight velocity dispersions, by binning the data, radially (e.g. Battaglia et al. (2008)). It is needless

to say that one loses information by replacing the velocities of several (spatially dispersed) stars with a single velocity dispersion at a single spatial location. Besides that, the choice of the bins is often arbitrary, in the sense that it is more motivated by the signal-to-noise of the observations than by underlying physics. If we assume that our target galaxies are not rotating, Jeans modeling predicts a spread in velocity for each location in the galaxy. Assuming that the velocity distribution at a certain point is normally distributed, the probability to observe a certain velocity $v_{\rm obs}$ is:

$$P(v_{\rm obs} \in [v_{\rm o}, v_{\rm o} + dv_{\rm o}]) = \frac{dv_{\rm o}}{\sqrt{2\pi \langle v^2_{\rm los} \rangle}} \exp\left(-\frac{v^2_{\rm o}}{2\langle v^2_{\rm los} \rangle} \right)$$

Here, $\langle v^2_{\rm los} \rangle$ is a prediction of the Jeans model for the second moment at a given location. What we are after is however not the probability that the data fits the model, but the likelihood of the model given the data. In addition to that, $v_{\rm obs}$ is affected by measurements errors σ_v. Treating these errors as gaussians, we find that for a single star this is just a convolution of two Gaussians, and hence:

$$\mathscr{L}(\langle v^2_{\rm los} \rangle \,|\, v_{\rm obs}, \sigma_v) = \frac{1}{\sqrt{2\pi \left(\langle v^2_{\rm los} + \sigma^2_v \rangle \right)}} \exp\left(-\frac{v^2_{\rm obs}}{2\left(\langle v^2_{\rm los} \rangle + \sigma^2_v \right)} \right)$$

5.4.1 Results

To see to what kind of improvement this will lead to for our data, we fit again the same sets of stars (all stars with membership probability higher than 0.99), with the same prior assumptions, but now with discrete modeling. A comparison of the results is shown in Fig. 5.8. This figure shows the posterior probability density functions for both methods (similarly arbitrarily normalized for both methods) for each free variable. It is obvious that the unbinned method pushes much harder towards a certain solution than the binned method. There is even an indication that Fornax has a declination different from 90 degrees. The anisotropies have changed from (strongly) tangential toward more isotropic (or in the case of Fornax, somewhat radial)

5.5 Application: pruning possible non-members

Of course it is very nice if one has better knowledge on the uncertainties of fitted variables using discrete fitting. However, there is still more information available with discrete fitting. Here, we give two examples.

In order to accurately test a dynamical model of a dwarf galaxy, one would in principal want all stars used in the analysis to be member of the system one is studying. Adding stars to the system will give a false perspective of the amount of

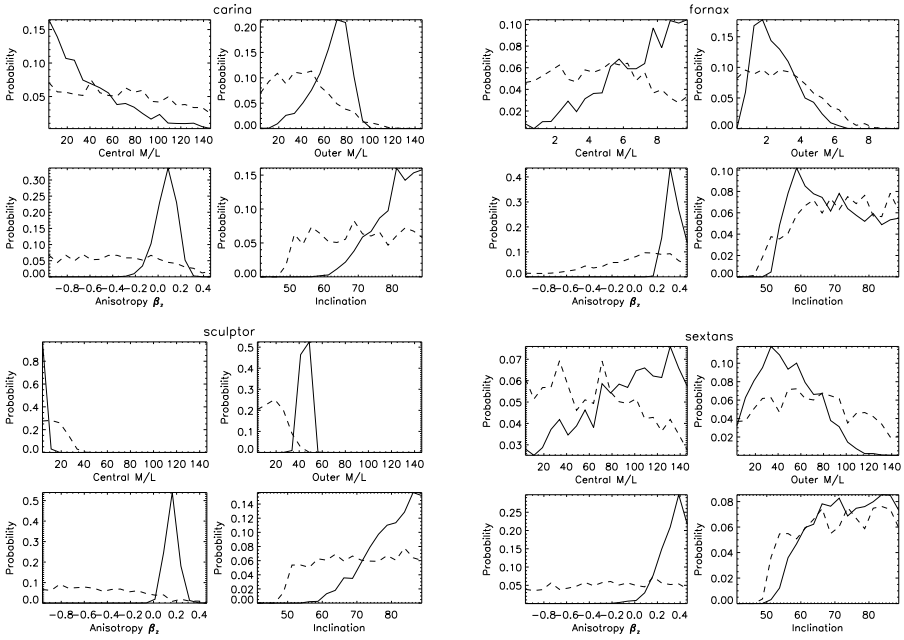

Figure 5.8 — Comparison of the posterior probability distribution functions (pdf) obtained with binned fitting and with discrete fitting. Shown are the pdfs for each of the 4 free parameters, for the binned fit (dashed line) and discrete fit (solid line). There is an obvious decrease in the width of (most) of the discretely fitted variables. Although some of the values may look inconsistent, this is mainly because certain denegeracies in the discrete modeling now seem to be broken, as is for example the case between the "inner" and "outer" M/L values. Another aspect is that most of the posterior mass is not necessarily where the likelihood is highest.

mass present, since it alters the observed second moments. However, in practice, membership has often been determined by clipping on velocities, often in combination with information about position. No need to say that this procedure is very tricky: one may remove actual member stars that are in the wings of the line-of-sight velocity distribution and interlopers that have the "right" velocities may contribute to give too low second moments. On the other hand, a few missed outliers may yield a too high M/L ratio.

Wojtak & Łokas (2007) employ an algorithm in which they iteratively calculate which objects are outliers from their current dynamical model and subsequently update their dynamical model with a new set of members. Walker et al. (2009) employ the Expectation-Maximization algorithm to separate out the distribution of Milky Way foreground stars from the stars in the dSphs. Here we try another method. Under

the assumption of a model for the Milky Way (with model we mean a distribution function of the spatial, kinematical and chemical distribution of stars) and a similar model for a dSph, we can determine the likelihood of different kinematic models for each dSph. As is, we do not have an accurate model for the Milky Way. However, given a dSph model, we can assume a set of plausible Milky Way models, and marginalize over those. To make this more formal, the total likelihood is:

$$\mathscr{L} = p(\{v_i\}_{i=1}^N|\{b_i\}_{i=1}^N, \text{dSph model, MW model}) \tag{5.2}$$

$$= \prod_{i=1}^N p(v_i|\text{dSph model})^{b_i} \cdot p(v_i|\text{MW model})^{1-b_i} \tag{5.3}$$

$$\tag{5.4}$$

where $b_i = 1$ if the star is a member and $b_i = 0$ is the star is not a member. Here,

$$p(v_i|\text{dSph model}) = \frac{1}{\sqrt{2\pi(\sigma^2 + \mathrm{d}v^2)}} \exp\left(-\frac{v_i}{2(\sigma^2 + \mathrm{d}v^2)}\right)$$

is the likelihood of the data given the dSph model, as determined from the Jeans model. Furthermore, we can write out the prior on the $\{b_i\}$:

$$p(\{b_i\}_{i=1}^N|P_{m_i}) = \prod_{i=1}^N P_{m_i}^{b_i} \cdot (1 - P_{m_i})^{1-b_i} \tag{5.5}$$

$$\tag{5.6}$$

which simplifies amazingly if all membership probabilities P_{m_i} would be same. Note how this prior punishes complicated models: if all stars are assumed to be members, $p(\{b_i\}_{i=1}^N|P_m = 1) = 1$, whereas the assumption that half of the population is non-member gives $p(\{b_i\}_{i=1}^N|P_m = \frac{1}{2}) = \left(\frac{1}{2}\right)^N$.
Since we are solving things in a Bayesian way, we are forced to assume a model for the distribution of MW star velocities. We note that this is generally dependent on the location of the dwarf galaxy on the sky, and moreover, on more complicated factors - since the observer may have had conscious and unconscious biases in selecting his stars for spectroscopic follow-up. Depending on the data, we will therefore assume either a Gaussian, exponential, or uniform distribution, and combinations of these, with plausible priors on the mean and width of these distributions and marginalize over these variables.

5.5.1 Example: Fornax

As an example, we analyse the membership of stars for Fornax. The data for Fornax are quite clean, in the sense that there are probably very few interloper stars (cf. Fig 5.2). As a first attempt, we assume for the MW a uniform distribution in velocity,

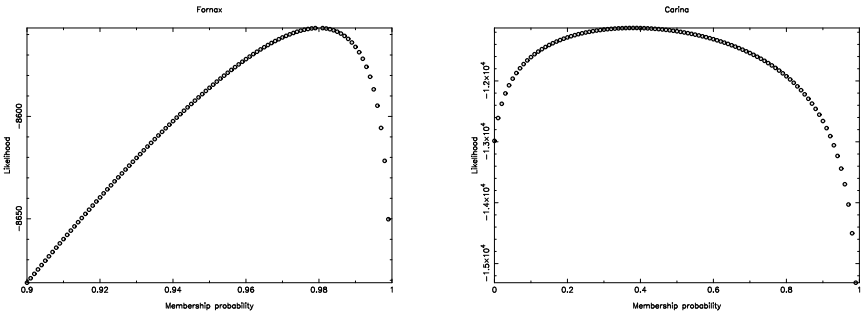

Figure 5.9 — Left: The likelihood of the combined Fornax/Milky Way model for different assumptions of the membership probabilities for individual stars. For the MW we assumed a uniform distribution with width of 700km/s. Right: Same plot, but now for Carina.

centred such that all observed velocities are possible:

$$p(v_i|\text{MW model}) = \frac{1}{2v_{\max}}$$

with $v_{\max} = 350$ km/s in this case, the likelihood for our dwarf model is:

$$p(\text{dSph model, MW model}|P_m, \{v_i\}_{i=1}^N) = \sum_{\text{all possible } b_i} \frac{p(\{v_i\}_{i=1}^N|\{b_i\}_{i=1}^N, \text{dSph model, MW model})}{p(\{v_i\}_{i=1}^N)} \times p(\{b_i\}_{i=1}^N|P_m)\, p(\text{dSph model})$$

For now, we act like we do not know the membership probabilities of the Walker et al. data, and ask ourselves if we can determine the membership purely by looking at the kinematics. As a first guess, we take P_m the same for all stars: this gives an overall membership probability for the sample. We can now evaluate the total likelihood for different membership probabilities. An example of this is given in Fig. 5.9. We see that indeed the Fornax sample is really quite clean and around 98% of the stars is probably a member. As we show in a similar plot for Carina, the situation can be quite different, however, a uniform model for the background may not be such a good approximation here. Now that we have the best fitting global P_m value, we can recalculate the probabilities that an individual star is a member of Fornax:

$$P_{m,i} = \frac{P_m P(v_i|Fornax)}{(1-P_m)P(v_i|MW)} \tag{5.7}$$

After this, one can update the models for the Milky Way foreground and for Fornax, by using the newly determined membership probabilities. As our assumed MW model in this example was a static uniform distribution, the procedure in this case is more or less equivalent to sigma clipping.

5.6 Application: splitting different kinematical tracers by metallicity

As shown by Battaglia et al. (2008), there appear to exist two distinct kinematical populations in Sculptor, one is metal poor, has a high velocity dispersion and is spread out over the galaxy, the other one is metal rich centrally concentrated and has a low velocity dispersion. Battaglia et al. (2008) analyse Sculptor by applying a hard cut in metallicity to split the sample in two parts: [Fe/H] < -1.7 and [Fe/H] > -1.5. After that, they bin the data radially and determine the line-of-sight velocity dispersions for each bin.
It is however a waste to throw away the data in the region where the two populations might overlap, especially since good line-of-sight data for dSphs is scarce. In this section, we model the metallicity distribution by two gaussians and calculate a best fitting model according to this.

We generate mock data of the Sculptor dwarf spheroidal using the following assumptions: we assume a constant mass to light ratio, which we fix to 25. Battaglia et al. (2008) determine the photometry of the metal-poor and metal-rich stellar populations of Sculptor by counting the number of blue and red horizontal branch stars as a function of radius. Subsequently, they fit a Plummer model to the metal-poor data and a Sérsic model to the metal rich data. Here, we assume the same light profiles for the metal poor and the metal rich population. We predict line-of-sight second moments for a metal poor ([Fe/H]=-1.7) and a metal rich sample ([Fe/H]=-1.3) of stars, each sample consisting of 600 stars. Each metallicity bin is smoothed in metallicity by convolving it with a Gaussian with $\sigma = 0.2$ The stars have a Gaussian distribution around the center of the galaxy, with standard deviation $\sigma = 20'$ for the metal poor population and $\sigma = 15'$ for the metal rich population. We assume slightly different velocity anisotropies for the samples: we assign $\beta_z = 0.2$ to the metal poor sample and $\beta_z = -0.98$ to the metal rich sample. This latter value for the anisotropy is very close to the edge of the prior. With these parameters, the metal rich sample has a very low velocity dispersion in the center. The metallicity distribution of the two samples and the combined sample is shown in Fig. 5.10

We assume a very simple form for the M/L profile: we leave the first and fifth Gaussian of the potential density free to vary between 0 and 50 times the luminous density, and interpolate linearly in $\Upsilon - -\log(\sigma)$-space, where σ is the width of a Gaussian in the MGE expansion.

At first, we fit the galaxy without looking at any of the metallicity information, with a global β_z and the M/L profile as above. Next we fit the galaxy taking into

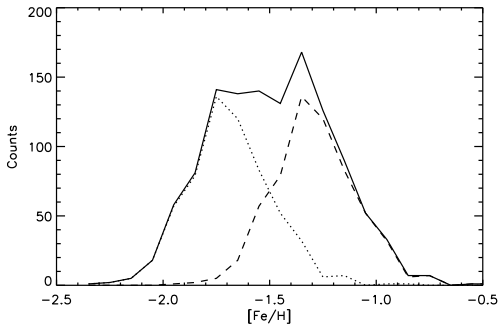

Figure 5.10 — Metallicity distribution for our simulated Sculptor. The metallicity distributions are given by two Gaussians, each with $\sigma = 0.2$.

account full knowledge of the metallicities of the stars, i.e., each model likelihood contains two values for β_z. The results of this are shown in the left four panels of Fig. 5.11. The effect of fitting a too simple model to the data manifests itself in a too high central M/L and a too low M/L in the outer parts. This is of course something that should worry people who try to constrain the slope of the potential in dSphs. The anisotropy for the single population fit is somewhere in between that of the metal rich and the metal poor tracers.

As a second experiment, we now fit again two models too the kinematic data, however, we do not have completele knowledge anymore. We assign to each star the probability of belonging to one of the metallicity samples according to the measured metallicity (we assume we know the intrinsic distribution of metallicities. This is not a strange assumption, because in practice it can be measured). The results are shown in the right four panels of Fig. 5.11, in which we compare it with the 'perfect knowledge' fit. We see that the combined fit to the data is able to recover the intrinsic anisotropies, and that we can constrain the central and outer M/Ls almost as good as in the 'perfect knowledge' case.

5.7 Discussion and conclusions

In this paper we have tried to progress the study of the matter content of Local Group dwarf Spheroidals in two ways. First, we have left the assumption of spherical symmetry, and second, we have developed a framework to analyze discrete kinematic data. We have shown that the dSphs are well fit by axisymmetric models, and that the step to discrete fitting strongly decreases the errorbars on the fitted values.

We have given two application of this work for Local Group dwarf spheroidals. By treating individual stars, it is possible to identify possible interloper stars. By carefully modeling both the foreground and the galaxy under study, it is possible to

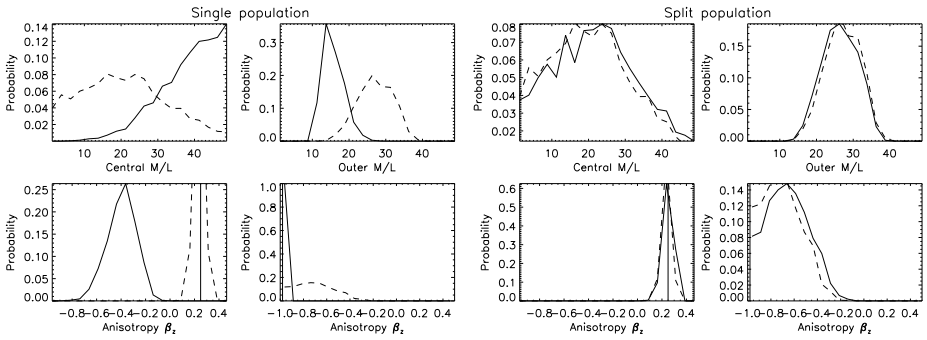

Figure 5.11 — Left: comparison of the results of fitting with two different kinematic tracer populations (of which membership of each population is exactly know – in all panels indicated by a dashed line) versus fitting of one tracer population (solid line). The addition of two distinct kinematic populations has the effect of enhancing the central M/L and lowering the outer M/L. The right panel shows how this changes when one assigns a probability of belonging to the metal rich or metal poor sample to each individual star (solid line).

improve the model of the galaxy. A second application comes from chemical tagging of stars inside a galaxy. If groups of stars form dynamically separate populations, these can be used to constrain the potential to a greater degree, as was done in Battaglia et al. (2008). It makes sense to treat this splitting in a statistical way.

Of course, Jeans modeling will not promise the existence of a physical solution – for that more advanced modeling tools like Schwarzschild modeling will stay necessary. However, due to its extremely high speed and high accuracy, modeling galaxies with discrete Jeans is a perfect precursor before starting to model galaxies with more computationally expensive other methods.