DWARF SPHEROIDAL GALAXIES: KINEMATICS OF STELLAR POPULATIONS AND MASS MODELLING WITH TIDES

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

by

Uğur Ural
Department of Physics and Astronomy
University of Leicester

October 2010
Abstract

DWARF SPHEROIDAL GALAXIES: KINEMATICS OF STELLAR POPULATIONS AND MASS MODELLING WITH TIDES

Üğur Ural

Dwarf Spheroidal galaxies (dSphs) are the nearest and the most dark matter dominated galaxies in the Universe. In this thesis, we study their kinematics and dynamics using Monte Carlo methods. First, we study the kinematics of stellar sub-populations in dwarf spheroidal galaxies and present a robust Monte Carlo based method for interpretation of the sub-population data in dSphs. We apply the method to new spectroscopic data for twenty six stars in the recently-discovered Canes Venatici I (CVn I) dSph, obtained with the GMOS-N spectroscope on the Gemini North telescope. We use these data to investigate the recent claim of the presence of two dynamically different stellar populations in this system [Ibata et al., 2006]. While we find no evidence for kinematically distinct sub-populations in our sample, we also show that the available kinematic data sets in CVn I might be too small to draw robust conclusions about its sub-populations. Second, we introduce a Markov Chain Monte Carlo based method for studying the dynamics of dSphs. We perform a large number of N-body simulations of the Carina dSph, modelling it with different dark matter halo profiles in the presence of tidal interactions. We show that due to the uncertainties in the data, it is possible to find several good models that can match Carina’s observed data. We show the differences in the mass, density, compactness and orbital eccentricities that result from different split power halo profiles, as well as mass follows light models. Finally, using high resolution re-simulations of some of our best models, we show the robustness of our results and perform a more detailed analysis of the models, looking at their mass evolution, and general tidal signatures that should be looked for in future observations.
Anneannem ve babaanneme tüm sevgimle...
ACKNOWLEDGEMENTS

I cannot thank my supervisor Mark Wilkinson enough for his constant care and support. With his deep knowledge, enthusiasm for science and great patience to teach with no judgements, he gave me the perfect opportunity to make the most of my time in Leicester. After some of our very long meetings, I used to wonder when you found time to work on anything else. It was a great pleasure to be your student. Thank you very much.

I would like to thank my PhD examiners, Wyn Evans and Sergei Nayakshin, for the helpful discussions and feedback that they provided.

I was also very fortunate to receive a SPARTAN Marie Curie Fellowship from the European Commission. I was offered all the support that I needed for my personal development as a researcher. I would like to thank Anna Odzimek and Martin Barstow for all their efforts in ensuring this.

I would like to thank Justin Read, for his ideas and consideration, and the three thrilling months that I have spent at the University of Zurich. Our small project ended up being a very long but exciting one for me. It was a pleasure to be in such a friendly and scientifically vibrant environment and beautiful country. I would like to thank everyone who made me feel welcome: George Lake, Prasenjit Saha, Aurel Schneider and Peter Englmaier and especially to Silvia Garbari, Hanni Lux and Jonathan Cole who also shared their codes with me and endured the working process.

I would like to thank Andreas Koch, for his help in the metallicity calculations, his careful feedback, as well as his care and valued friendship, and Matthew Walker for his data, help, and friendly encouragement.

I would also like to thank all my other collaborators for their contributions and
feedback in the CVnI paper: Gerry Gilmore, Timothy Beers, Vasily Belokurov, Eva Grebel, Simon Vidrih and Daniel Zucker. And thank you Walter Dehnen and Joachim Staedel for your codes.

Thank you Gauthier, Gail, Magda, Laura, Vicky, Artur, Christian, Rebecca, Joao, Kristiina and Chris for the fun, for the ping pong, for your friendship, for helping me to survive. Thank you my office mates who have been great company, Gary, David, Naomi, Şirin, Patrick, Fergus.

Thank you to my extended family the Parsons, for the two and a half years that you have been there for me. Especially Mark James Parsons for your love and support as well as your bright mind and the evenings you sat on the floor with me discussing this “strange field of science”.

I would like to thank Ceyhun Kuyumcu, for supporting me selflessly and never letting me doubt myself. Oktay Urfalı, for giving me the chance to start. I am thankful to Cengiz Aslan, for his great interest and care, looking out to ease my path for the future. To Emine Yanıkdağ for the long distance calls that were always perfectly timed, just when I needed to hear a friendly voice. To Sevtap Elmas and Evrim Çalkavur whose faith gave me strength.

Most of all I would like to thank my beloved family that I miss very much. My aunt, for her love and encouragement. Both my grandmothers who passed away while I was working on this thesis, for how much they only and always saw the good in me and for the inspiration they gave me for life with the way they lived theirs. My mother, for her endless love and loyalty, and invincible hope. Finally my safe port, my guide, my father, for always reminding me who I want to be when I ”grow up”, with his kindness and love. I love you all very much.
Teşekkürler, teyzeme sevgisi ve teşvigi için. Her ikinizi de doktoramı tamamlarken kaybettığim anneannem ve babaannem; teşekkürler, her zaman ve her durumda yalnız iyı yönleriğini görmüş olduğunuz ve kendi yaşamlarınızla bana esin kaynağı olduğunuz için. Annem, teşekkürler, sonsuz sevgin, sadakatin, ve yenilmez umudun için. Ve son olarak rehberim ve güvenli limanım, babam, teşekkürler ‘büyükence’ kim olmak istediğimi hatırlamama sevgin ve şefkatinle hep yardımcı olduğun için. Hepinizi çok seviyorum.
# Contents

## 1 Introduction

1.1 Cosmological framework: Λ Cold Dark Matter model ........................................... 3  
   1.1.1 Hot, warm or cold? ......................................................................................... 7  
   1.1.2 Problems with Λ Cold Dark Matter ............................................................... 10  
1.2 Mass modelling and tides ....................................................................................... 14  
   1.2.1 Collisionless Boltzmann Equation and Jeans analysis ................................. 16  
1.3 Estimating dSph masses ......................................................................................... 17  
1.4 Dwarf spheroidal galaxies ..................................................................................... 19  
1.5 Stellar structures of the dSphs and Ultra-faint dwarfs ......................................... 20  
1.6 Outline .................................................................................................................. 24

## 2 Kinematic sub-populations in dwarf spheroidal galaxies ............................... 26

2.1 Introduction ............................................................................................................. 26  
2.2 Canes Venatici I ................................................................................................. 31  
   2.2.1 Data reduction ............................................................................................... 31  
   2.2.2 Velocities ...................................................................................................... 34  
   2.2.3 Metallicities ................................................................................................... 37
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Map of the Local Group</td>
<td>3</td>
</tr>
<tr>
<td>1.2</td>
<td>Black body spectrum</td>
<td>5</td>
</tr>
<tr>
<td>1.3</td>
<td>Cosmic Microwave Background power spectrum</td>
<td>8</td>
</tr>
<tr>
<td>1.4</td>
<td>Bullet cluster</td>
<td>10</td>
</tr>
<tr>
<td>1.5</td>
<td>Rotation curve</td>
<td>12</td>
</tr>
<tr>
<td>2.1</td>
<td>Color magnitude diagram of CVn I</td>
<td>32</td>
</tr>
<tr>
<td>2.2</td>
<td>GMOS target fields for CVn I</td>
<td>33</td>
</tr>
<tr>
<td>2.3</td>
<td>Radial distribution of CVn I stars</td>
<td>33</td>
</tr>
<tr>
<td>2.4</td>
<td>Examples of spectra for CVn I</td>
<td>35</td>
</tr>
<tr>
<td>2.5</td>
<td>Velocity histogram of CVn I members</td>
<td>39</td>
</tr>
<tr>
<td>2.6</td>
<td>Difference between GMOS and KECK data of CVn I</td>
<td>40</td>
</tr>
<tr>
<td>2.7</td>
<td>Metallicity and line of sight velocity for CVn I members</td>
<td>40</td>
</tr>
<tr>
<td>2.8</td>
<td>Detectability histograms for GMOS data of CVn I</td>
<td>47</td>
</tr>
<tr>
<td>2.9</td>
<td>Detectability histograms for KECK data of CVn I</td>
<td>47</td>
</tr>
<tr>
<td>2.10</td>
<td>Detectability depending on the sample size</td>
<td>48</td>
</tr>
<tr>
<td>2.11</td>
<td>Detectability depending on cold population size</td>
<td>51</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td>------</td>
</tr>
<tr>
<td>2.12</td>
<td>Detectability depending on the velocity errors</td>
<td>52</td>
</tr>
<tr>
<td>2.13</td>
<td>Improvements in detections using velocity and metallicity</td>
<td>57</td>
</tr>
<tr>
<td>3.1</td>
<td>Opening angle</td>
<td>66</td>
</tr>
<tr>
<td>3.2</td>
<td>Stellar and dark matter components for tests B1 and B7</td>
<td>84</td>
</tr>
<tr>
<td>3.3</td>
<td>Comparison of $d\eta = 0.1$ and $d\eta = 0.4$ tests</td>
<td>88</td>
</tr>
<tr>
<td>3.4</td>
<td>Comparison of different $\epsilon$ values</td>
<td>88</td>
</tr>
<tr>
<td>3.5</td>
<td>Comparison of $d\eta = 0.25$ and $d\eta = 0.25$ tests</td>
<td>89</td>
</tr>
<tr>
<td>3.6</td>
<td>Stability test $\chi^2$</td>
<td>89</td>
</tr>
<tr>
<td>3.7</td>
<td>Stability test initial evolution</td>
<td>90</td>
</tr>
<tr>
<td>3.8</td>
<td>Stability test long term evolution</td>
<td>91</td>
</tr>
<tr>
<td>3.9</td>
<td>High and low resolution comparison for extreme orbit tests</td>
<td>92</td>
</tr>
<tr>
<td>3.10</td>
<td>Residuals between the high resolution numerical tests</td>
<td>93</td>
</tr>
<tr>
<td>3.11</td>
<td>Particle distribution comparison for the high and low resolution tests</td>
<td>93</td>
</tr>
<tr>
<td>3.12</td>
<td>Velocity distribution comparison for the high and low resolution tests</td>
<td>94</td>
</tr>
<tr>
<td>4.1</td>
<td>Surface brightness profile for Carina</td>
<td>100</td>
</tr>
<tr>
<td>4.2</td>
<td>Projected velocity dispersion profile for Carina</td>
<td>101</td>
</tr>
<tr>
<td>4.3</td>
<td>Simulated model galaxy’s observed profiles</td>
<td>107</td>
</tr>
<tr>
<td>4.4</td>
<td>Evolution of the $\chi^2$ in MCART</td>
<td>108</td>
</tr>
<tr>
<td>4.5</td>
<td>Perigalactic distance and eccentricity for MCART orbits</td>
<td>109</td>
</tr>
<tr>
<td>4.6</td>
<td>Distribution of the initial halo masses in MCART</td>
<td>110</td>
</tr>
<tr>
<td>4.7</td>
<td>Distribution of the final halo masses in MCART</td>
<td>110</td>
</tr>
<tr>
<td>4.8</td>
<td>Density contours for MCART:1</td>
<td>112</td>
</tr>
<tr>
<td>4.9</td>
<td>Density contours for MCART:2</td>
<td>113</td>
</tr>
<tr>
<td>4.10</td>
<td>The evolution of the $\chi^2$ for MC1</td>
<td>115</td>
</tr>
<tr>
<td>4.11</td>
<td>Stellar mass distribution for MC1</td>
<td>116</td>
</tr>
</tbody>
</table>
4.12 Density and velocity profiles of the best MC1 model ............. 117
4.13 Density contours for MC1:1 ........................................ 118
4.14 Density contours for MC1:2 ........................................ 119
4.15 The evolution of the \( \chi^2 \) for MC2 ......................... 122
4.16 The distribution of stellar masses in MC2 ...................... 123
4.17 Perigalactic distance and eccentricity for MC2 ............... 123
4.18 Distribution of the final masses in MC2 ....................... 123
4.19 Density contours for MC2:1 (All accepted models) ........... 124
4.20 Density contours for MC2:2 (Best \( \chi^2 \)) .................. 126
4.21 Density contours for MC2:3 (Best \( \chi^2 \)) .................. 127
4.22 Color coded perigalactic distance eccentricity for MC2 .... 129
4.23 Density contours for MC3:1 ....................................... 132
4.24 Density contours for MC3:2 ....................................... 133
4.25 Perigalactic distance and eccentricity for MC3 ............... 133
4.26 Color coded perigalactic distance and eccentricity for MC3 . 134
4.27 Rejected halo masses and stellar scale lengths for MC3 .... 136
4.28 Density contours for MC4:1 ....................................... 138
4.29 Density contours for MC4:2 ....................................... 139
4.30 Perigalactic distance and eccentricity for MC4 ............... 139
4.31 Color coded perigalactic distance and eccentricity for MC4 . 140
4.32 Density contours for MC5:1 ....................................... 142
4.33 Density contours for MC5:2 ....................................... 142
4.34 Perigalactic distance and eccentricity for MC5 ............... 143
4.35 Color coded perigalactic distance and eccentricity for MC5 . 144
5.1 Stellar and dark matter components for \( C22_{238} \) .......... 151
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2</td>
<td>Stellar and dark matter components for $C28_7$.</td>
<td>152</td>
</tr>
<tr>
<td>5.3</td>
<td>Stellar and dark matter components for $GN1_{141}$.</td>
<td>153</td>
</tr>
<tr>
<td>5.4</td>
<td>Stellar and dark matter components for $C2_{11}$.</td>
<td>154</td>
</tr>
<tr>
<td>5.5</td>
<td>Mass loss – orbit relation for high resolution simulations.</td>
<td>155</td>
</tr>
<tr>
<td>5.6</td>
<td>Scattering plots $C22_{38}$.</td>
<td>157</td>
</tr>
<tr>
<td>5.7</td>
<td>Scattering plots for $M2_{11}$.</td>
<td>158</td>
</tr>
<tr>
<td>5.8</td>
<td>2D $v_{los}$ and $\sigma_{los}$ dispersion maps for $M2_{11}$.</td>
<td>159</td>
</tr>
<tr>
<td>5.9</td>
<td>Mean velocity along the major axis for $M2_{11}$.</td>
<td>161</td>
</tr>
<tr>
<td>5.10</td>
<td>2D $v_{los}$ and $\sigma_{los}$ maps for $C22_{38}$, $C28_7$ and $GN1_{141}$.</td>
<td>162</td>
</tr>
<tr>
<td>5.11</td>
<td>Tidal signature Velocity histogram for $M2_{11}$: High resolution.</td>
<td>163</td>
</tr>
<tr>
<td>5.12</td>
<td>Tidal signature Velocity histogram for $M2_{11}$: Low resolution.</td>
<td>163</td>
</tr>
<tr>
<td>5.13</td>
<td>(Non-) Tidal signature velocity histogram for $C22_{38}$.</td>
<td>164</td>
</tr>
<tr>
<td>5.14</td>
<td>Tidal signature velocity histogram for $MC2_{C33-0}$.</td>
<td>165</td>
</tr>
</tbody>
</table>
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>CVn I data from GMOS-N.</td>
<td>42</td>
</tr>
<tr>
<td>2.2</td>
<td>Confidence limits of sub-population detections</td>
<td>54</td>
</tr>
<tr>
<td>2.3</td>
<td>Confidence limits of sub-population detections using metallicity</td>
<td>56</td>
</tr>
<tr>
<td>2.4</td>
<td>Confidence limits for two-dimensional detections</td>
<td>56</td>
</tr>
<tr>
<td>3.1</td>
<td>Milky way disk properties</td>
<td>70</td>
</tr>
<tr>
<td>3.2</td>
<td>Stellar component parameters for split power fit.</td>
<td>75</td>
</tr>
<tr>
<td>3.3</td>
<td>The first choice of parameter ranges for MCMC</td>
<td>76</td>
</tr>
<tr>
<td>3.4</td>
<td>Galactic coordinates of the centre of Carina</td>
<td>77</td>
</tr>
<tr>
<td>3.5</td>
<td>Summary of the high resolution numerical test inputs</td>
<td>83</td>
</tr>
<tr>
<td>3.6</td>
<td>List of lower resolution numerical tests</td>
<td>86</td>
</tr>
<tr>
<td>3.7</td>
<td>Extreme orbit test list</td>
<td>91</td>
</tr>
<tr>
<td>3.8</td>
<td>Medium resolution simulation list</td>
<td>94</td>
</tr>
<tr>
<td>4.1</td>
<td>Surface brightness profile for Carina</td>
<td>100</td>
</tr>
<tr>
<td>4.2</td>
<td>Projected velocity dispersion profile for Carina</td>
<td>101</td>
</tr>
<tr>
<td>4.3</td>
<td>MCMC set split power parameters</td>
<td>104</td>
</tr>
<tr>
<td>4.4</td>
<td>MCMC set parameter ranges</td>
<td>104</td>
</tr>
<tr>
<td>Table Number</td>
<td>Table Title</td>
<td>Page</td>
</tr>
<tr>
<td>-------------</td>
<td>------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>4.5</td>
<td>Density constraint for MC2</td>
<td>125</td>
</tr>
<tr>
<td>4.6</td>
<td>Constraints on parameters found for all MCMC sets</td>
<td>146</td>
</tr>
<tr>
<td>4.7</td>
<td>The models with the best $\chi^2$ for each MCMC set</td>
<td>147</td>
</tr>
<tr>
<td>5.1</td>
<td>Model parameters of the high resolution simulations</td>
<td>150</td>
</tr>
<tr>
<td>5.2</td>
<td>Tidal radii of the high resolution models</td>
<td>155</td>
</tr>
</tbody>
</table>
1

Introduction

Our location in the Universe can be described as on a planet with a moon in the Solar System surrounded by billions of stars in the Milky Way Galaxy, which is one of the largest spiral galaxies in the Local Group surrounded by more than 30 satellite galaxies (Figure 1.1) in the Virgo Super Cluster of Galaxies, in a 13.7 billion year old Universe.

This hierarchy offers two perspectives on the importance of galaxies. According to the first one, thousands to billions of stars are born, evolve and die in galaxies.
Introduction

The physical processes on these scales are dominated by the baryonic matter which is contained in the stars. From the second perspective, the structure of and interactions between individual galaxies govern the formation and evolution theories of structure in the Universe. For both of these, although our own galaxy’s structure can be studied in detail, some of the outstanding questions are easier to address by looking at other nearby galaxies that we can resolve in detail.

The dwarf spheroidal galaxies (dSphs) are the closest neighbours of the Milky Way that have been accompanying it since the early ages of the Universe. They are intriguing objects to the galactic archeologist trying to understand galaxy formation and more specifically the history of the Local Group. In this thesis we study two of these dSphs in detail and hence try to obtain information about dSphs in general.

First, we investigate the previously claimed multiple stellar populations in the Canes Venatici I (CVn I) dSph and calculate the confidence with which the substructures in it can be detected. Second, we use more than 10000 N-body simulations to model the dynamics of the Carina dSph. Apart from finding some models that fit Carina’s data, we also find the degeneracies between model parameters that need to be taken into account in the dynamical modelling and mass estimation of dSphs.

In this chapter, the motivation for studies of dSphs is first given in the context of the standard model of cosmology and its predictions. We then explain the motivation for our research shown through the complexity of dSphs’ structures and the difficulty of accurate mass modelling.
1.1 Cosmological framework: $\Lambda$ Cold Dark Matter model

In the standard cosmological model, the Universe started from a singularity called the Big Bang and has since been expanding. The baryonic matter which consists of atoms and molecules makes up only 4% of the total mass-energy content of the Universe. The most powerful component of the total content is the dark energy which consists of 72% of the total energy of the universe. The remaining 23% consists of so-called dark matter. Little is known about the nature of dark matter, apart from its lack of interactions other than through gravitational forces. However, the currently favored cosmological model is based on the dark matter being 'cold' so that it consists of non-relativistic particles that were able to form clumps of matter very early in the evolution of the Universe. In this hierarchical structure formation paradigm, the first dark halos that were formed by the gravita-
tional collapse of these first clumps are then thought to become the building blocks of ever larger structures from Galaxy Superclusters (Bahcall & Soneira, 1983) to Great Walls (Haynes & Giovarelli, 1986; Geller & Huchra, 1989).

As the Universe expanded, the baryonic matter has undergone two main phase changes. For the first 400000 years, the matter was in the form of a hot plasma of photons, protons and electrons. The photons interacted with free electrons through Thomson scattering where the energy and momentum of the incident photons were conserved. Due to the high cross section of Thomson scattering for the high frequency photons in the very dense early Universe, the mean free path of the photons was very short at this epoch making the Universe opaque. As the density kept decreasing, the Universe became too cold for continuous ionisation and the protons and electrons started to combine together to form neutral Hydrogen. In contrast to the frequent scattering of the photons in this hot plasma, the scattering of the photons occurred rarely in the neutral Universe after Hydrogen recombination, as it was only possible for photons that had the right energies to excite the energy levels of neutral Hydrogen. Thus, the Universe became transparent and the radiation emitted by the last scattering of the photons at the beginning of recombination has travelled through the Universe since then, reaching us today from every direction in the Universe as the Cosmic Background Radiation (CBR).

The Universe was at this point in thermodynamic equilibrium and hence the emitted radiation was that of a black body as described by Planck’s Law:
Introduction

1.1. Cosmological framework: Λ Cold Dark Matter model

The spectral radiance \( I_\lambda \) of the CMB as a function of wavelength \( \lambda \) is a black body radiation curve at temperature \( T = 2.726 \text{K} \).

\[
I_\lambda = \frac{2\pihc^2}{\lambda^5 \exp(hc/kT\lambda)} - 1
\]

where \( I_\lambda \) is the intensity of the radiation at wavelength \( \lambda \), \( h \) is Planck’s constant, \( k \) is the Boltzmann constant, \( c \) is the speed of light and \( T \) is the characteristic temperature of the black body. It is seen from Equation (1.1) that the wavelength at which the intensity is a maximum depends on the temperature. Although the initial characteristic temperature of the CBR was very high, as the density in the Universe decreased it has cooled down since it was first emitted and currently it peaks at \( \lambda = 1.9 \text{mm} \) corresponding to a temperature \( T=2.726 \text{K} \) (Figure 1.2). This wavelength is in the microwave window and therefore the radiation is called the Cosmic Microwave Background Radiation (CMBR).

While being a nearly perfect black body radiation spectrum, there are small temperature fluctuations in the CMBR due to quantum fluctuations in the hot plasma. These primordial fluctuations trace the small density inhomogeneities that were the
seeds of the large scale structures we see in the Universe today. The fluctuations grew as the denser areas attracted more matter and hence the density in the lower areas kept decreasing. In the absence of fundamental interactions other than gravity in this Early Universe, these dark matter clumps formed the first galaxy progenitors as they collapsed due to gravitational instability and were able to capture gas particles. Eventually as the first stars started to form, the radiation due to the nuclear reactions in their cores started to reionise the intergalactic medium. The reionisation epoch occurred between 150 million years and one billion years after the Big Bang. It is thought that during this epoch galaxy formation was suppressed as the first stars in these proto-galaxies and other bright objects such as QSOs were radiating energy, stopping the gas from collapsing to form more stars in the gravitational wells of the dark matter haloes. Despite the reionisation of the matter, the Universe remained transparent as the the density decreased in the constantly expanding, low density Universe. Understanding the exact time that reionisation began and the way it occurred is an important question that cosmological models try to answer, as these have implications for how rapidly the first galaxies would have to form in the early Universe and would also affect the star formation histories of old galaxies.

A good review of the development of CMBR research can be found in the 2007 Nobel Lecture in Physics by George Smoot (Smoot, 2007), while a more quantitative discussion can be found in Binney & Tremaine (2008).
1.1.1 Hot, warm or cold?

Finding dark matter and understanding its nature are major problems in cosmology and particle physics. There are several dark matter candidates that have been considered in the past years from Massive Compact Halo Objects (MACHOs), such as Black Holes and brown dwarfs, to Weakly Interacting Massive Particles (WIMPS) such as neutrinos, and exotic super-symmetric particles (Ellis & Olive, 2010). MACHOs have been largely ruled out as Dark Matter candidates as observations of microlensing and binary stars showed that their contribution to the total mass of our Galaxy is too small to account for the excess mass (Gates et al., 1996; Griest et al., 2001; Yoo et al., 2004; Quinn et al., 2009). Experimental particle physicists build sensitive underground observatories like Super Kamiokande in order to detect solar neutrinos, or particle accelerators like the Large Hadron Collider (LHC) to look for super symmetric particles that are predicted by the extensions to the Standard Model of Particle Physics.

On the other hand, astronomers rely on detection methods where the presence of dark matter is inferred indirectly, by observing its effect on the luminous structures in the Universe and comparing the predictions of cosmological models with the observed properties. One of the methods that are used to address these issues is studying the CMBR. Its highly isotropic black body spectrum at $T=2.725$ (Mather, 1994) is used to determine the temperature at the time of reionisation. It can also be used to map the density inhomogeneities in the last scattering surface through the small fluctuations in its temperature distribution. The amplitudes of these inhomogeneities determine the time scales at which structures can form in different dark matter scenarios. In addition, the power spectrum of these on different angu-
lar scales in the sky is used to constrain other cosmological parameters such as the
density at the time of recombination and the flatness of the Universe.

![Figure 1.3: CMB power spectrum from 7 years data of WMAP. Figure is from Larson et al. (2010).](image)

The amplitudes of these fluctuations are so small ($\Delta T/T = 10^{-5}$) that they already
rule out Hot Dark Matter (HDM), as in the cosmological simulations relativistic
HDM particles smooth out the initial perturbations and the galaxies form through
fragmentation at times later than seen in the observations. As a result, slower dark
matter particles in Warm Dark Matter (WDM) and Cold Dark Matter are invoked
in an attempt to allow the particles to group into clumps easily as they have lower
initial velocity dispersions. On the other hand, comparison of the fluctuations of
the Gaussian random field with the high accuracy CMBR data from Wilkinson Mi-
crowave Anisotropy Probe determined that the reionisation time had to be before
$z=20$, ruling out the Warm Dark Matter model as well, in which the first clumps
would not have formed in time (Spergel et al., 2003).

Another important method used in the Dark Matter search is gravitational lensing
Introduction

1.1. Cosmological framework: \( \Lambda \) Cold Dark Matter model

which results from bending of space time by massive objects. As the light from a distant source travels towards us, passing close to a massive object on its trajectory, it follows the curvature of space-time and becomes distorted by an amount that depends on the object that causes the lensing. In strong lensing, a single background object’s image is very distorted by a foreground mass distribution and the background object is detected as several objects forming a ring or just as distinct multiple images of the same object. In weak lensing, the effect is too small to be detected for a single source but systematic distortions in the shapes of a group of lensed objects is detected and the lens mass can be calculated using the statistical properties of all the distortions in the area. Finally, in microlensing only a small amount of change in the luminosity of the background source with time is detected. By careful analysis of the background object’s light curve, the mass of the lensing object can be estimated. Large gravitational lensing surveys can be very useful as they can be used to map the mass distribution on the sky and compare it to the prediction of the theories. Figure [1.4] shows the Bullet Cluster which is a very good demonstration of the strength of gravitational lensing while looking for the evidence of dark matter. The picture shows the resulting mass distribution in the Bullet Cluster that was formed through the collision of two galaxy clusters. The red areas in the centre of the Figure show the hot gas emitting X-Rays. The gas is heated through the shocks caused by the colliding gas of each galaxy. In such a collision, the dark matter components of the clusters are expected to have travelled further out than the gas, as the dark matter basically moves as a collisionless fluid, i.e., is able to travel further without being slowed down by interactions. This is confirmed by the gravitational lensing data which show that although the total stellar mass is smaller than the gas mass, most of the mass in the Bullet Cluster is out in the blue areas where the amount of gas is negligible (Clowe et al., 2006).
1.1. Cosmological framework: $\Lambda$ Cold Dark Matter model

Currently the CDM is the best dark matter candidate as it agrees with most of the observations about structure formation. However, despite its success over the WDM and HDM scenarios, $\Lambda$-CDM theory has its own problems which will be discussed in the next section.

1.1.2 Problems with $\Lambda$ Cold Dark Matter

One of the challenges that $\Lambda$-CDM needs to overcome is the so-called missing satellites problem, first highlighted by (Klypin et al., 1999; Moore et al., 1999).

The cosmological simulations that were made by these authors using $\Lambda$-CDM predicted that hundreds of sub-haloes should exist near large Milky Way-like haloes.
1.1. Cosmological framework: \(\Lambda\) Cold Dark Matter model

Although some of these would merge together to grow the halo of their host galaxy, the expected number of sub-haloes at the present is still an order of magnitude larger than the observed satellites of Milky Way, if we assume that the satellite haloes seen in the simulations correspond to the observed satellite galaxies.

However, the number of known dSph satellites of the Milky Way has increased significantly in the past few years, due to the discovery of new dSphs by Willman et al. (2005); Zucker et al. (2006a,b); Belokurov et al. (2006, 2007); Walsh et al. (2007) in the data from the Sloan Digital Sky Survey (SDSS; York et al., 2000). Since the SDSS covers only about one quarter of the sky, it is thus likely that the total number of satellites surrounding the Milky Way may be at least a factor of four larger than previously thought, although the extrapolation from the SDSS survey to the whole sky requires careful analysis (see e.g. Tollerud, 2008).

In addition, careful analysis of the data can decrease the discrepancy further as all the sub-haloes in the simulations might not be detectable if due to their formation epoch and the effect of reionisation they were not able to capture gas and form stars (e.g. Benson et al., 2002). Furthermore, the observed dSphs may be merely a particular subset of the sub-haloes that were able to survive subsequent tidal interactions with the Milky Way (e.g. Stoehr et al. (2002); Diemand et al. (2005); Moore et al. (2006); Simon & Geha (2007); Strigari et al. (2007); Bovill (2009)) for reasons of their masses and orbits etc.

The second challenge \(\Lambda\)-CDM is confronted with is the discrepancy between the observed and predicted density profiles in the inner parts of the Low Surface Brightness (LSB) galaxies. The rotation curves of the LSBs are similar to those of spi-
ral galaxies remaining flat out to large radii (red curve in Figure 1.5) suggesting that their mass profiles are dominated by extended dark matter distributions (see McGaugh et al., 2001). This means that the stellar distribution of LSB galaxies can be used to trace the dark matter distribution in the inner parts of these galaxies. Although $\Lambda$-CDM simulations predict that all dark matter halos have steep density profiles in their inner parts, these overdensities (cusps) are not found when the dark matter haloes of the observed LSBs are modelled using their rotation curves (de Blok & Bosma, 2002). Their rotation curves are instead consistent with constant density distributions (cores).

![Figure 1.5: Rotation curves of galaxies. The blue curve belongs to a Keplerian rotation where the velocity is inversely proportional to the radius, while the red dots show an example of data, that give the real spiral and low surface brightness galaxies’ flat rotation curves.](image)

Another point to consider regarding the core/cusp discrepancy, when interpreting the results of cosmological simulations is that, despite forming only 4% of the total matter in the Universe, the baryonic matter is not always negligible since on small scales it interacts in several ways and with stronger interactions. This is a
fundamental problem with the cosmological simulations that do not take into account the baryonic physics and simulate all the structures in the Universe only with dark matter and through gravitational interactions. Including the baryonic physics in these simulations might result in halting the formation of the cusps at early stages as energy is dissipated through the infall of gas, or matter is swept out by the winds from the first stars or later by supernovae. In addition, even if cusps are initially formed they might have been transformed to cores. The mechanisms that have been considered for transforming the density cusps to constant density cores include the existence of a central bar where angular momentum is transferred from the rotating bar (Valenzuela et al., 2007) to the dark matter halo, dynamical friction in the early Universe where merging gas clouds disrupt cusps (Tonini et al., 2006) and other methods taking into account baryon-dark matter interactions (see de Blok (2009) for an overview of this field). However, there currently exists no common consensus in the area and the core-cusp problem remains one that challenges the Λ-CDM model (de Blok, 2009). Being the closest Dark Matter dominated systems (Mateo, 1998) and potentially the most Dark Matter dominated ones, dSphs could help to explore the cusp-core discrepancy. Accurate kinematic data from dSphs could help to constrain the actual density profiles of dSph haloes or to improve the dynamical models and hence provide us with clues about the halo evolution.

As dark matter does not interact except gravitationally, aside from particle accelerators and underground neutrino detectors, we rely on accurate mass estimation of stellar systems for testing dark matter theories. Although gravitational lensing is very promising, the current surveys are not extensive enough to answer all the questions and not every galaxy is amenable to a lensing study. In the next section, we therefore introduce other methods used for mass estimation of galaxies.
1.2 Mass modelling and tides

In theory, an orbit of a star in a galaxy can be calculated using Newtonian dynamics to calculate the gravitational force exerted on the star by each of the remaining stars in the galaxy (Equation 1.2). Direct N Body codes use this approach to simulate the motions of millions of stars at the same time. In practice however, our observational data is not complete to account for all the stars in a galaxy and instead we approximate the gravitational field in which a particle \( j \) is in, by a smoothed gravitational potential, \( \phi \).

\[
F = \sum_{i=1}^{N} \frac{Gm_i m_j (r_i - r_j)}{|r_i - r_j|^3} = -m_j \nabla \phi \tag{1.2}
\]

The total gravitational force \( F \) that act on each star in the system determines the orbit of the star via \( a_j = F/m_j \), and the gravitational potential is proportional to the mass of the system that provides the potential. Thus the orbital velocities of the stars can be used directly to estimate the mass of the system.

**Definition: Virial theorem:** Classical Newtonian Mechanics shows that in a self gravitating system, the potential energy of the system (\( P \)) is equal to twice the total kinetic energy (\( K \)) of all the individual masses in it: \( 2K = -P \). For a total of \( N \) particles with masses \( m \), in a system of radius \( R \), the gravitational potential energy of the system is calculated using \( P \approx -\frac{Gm^2N^2}{2R} \), and the total kinetic energy is the sum of the individual kinetic energies, ie. \( K = \frac{Nmv^2}{2} \), where \( v^2 \) is the mean of the squares of all the velocities.

In 1937, Fritz Zwicky used the virial theorem to show that in the Coma Cluster, the high velocities of the galaxies in the orbits were pointing to a higher total mass.
than the mass seen in the luminous objects (Zwicky, 1937). In 1975 Vera Rubin also showed that the rotation curves of many spiral galaxies were not decreasing with the increasing radius as expected in Keplerian motion \( v = (GM/r)^{1/2} \) but rather stayed flat all the way through to the outer radii of the observed mass. As the velocity in these outer parts was larger than expected, for galaxies in virial equilibrium this would mean that the total mass was greater than the luminous component of the galaxies and was much more extended. Despite the success of the rotation curves analysis in these early days, failure to detect the dark matter particles to this day and the persisting discrepancy between the observations and the theory necessitate more accurate mass modelling to resolve the issue. In elliptical galaxies, which are thought to form by merging spiral galaxies, the systems usually do not have net rotation, nor are there large amounts of gas that can be traced out to large radii. In these galaxies, the stars do not rotate around the centre of the galaxy as in Keplerian motion and the pressure that supports the galaxy against gravitational collapse comes from the random motions of the stars. In this case the velocity dispersion, \( \sigma^2 \), of these random velocities is instead used for mass calculations.

When looking at galaxy clusters, the temperature of the hot gas stripped from the infalling galaxies can be calculated using the X-Ray spectra of the gas. This temperature is directly related to the pressure of the gas and is the origin of the force supporting the gas in the cluster against gravitational collapse. In this thesis, however, we are interested in much smaller objects which do not contain large amounts of hot gas and we need another method to estimate the masses.
1.2.1 Collisionless Boltzmann Equation and Jeans analysis

As a more sophisticated method to calculate the masses of galaxies, in cases where
the system does not lose or gain stars, we can model the phase space distribution
of the system as a fluid (Equation 1.3), and find a solution which conserves the
phase space density of the system. Equation 1.3, called the Collisionless Boltz-
mann Equation, is used for this purpose. The mass distribution is represented by a
density function, ν, and we define a phase space density f(x,v,t) where each star is
described by its position in the 6 dimensional phase space (x,v).

\[ \frac{\partial f}{\partial t} + v \frac{\partial f}{\partial r} + \nabla \phi \frac{\partial f}{\partial v} = 0 \]  

(1.3)

This partial differential equation describes the distribution function of a collision-
less system as an incompressible fluid where the total number of stars in the system
is conserved. Then, instead of using the Newtonian approach to calculate the inter-
actions between all individual pairs of particles it uses the probability distribution
function (DF) in the 6 dimensional phase space of positions (r) and velocities (v).

Although f represents a density function, in systems with low (number) density
such as galaxies, the density is not continuous. Therefore f is smoothed out to a DF
of the probabilities of any point in the phase space (r,v) being occupied at a given
time. This DF is conserved in the phase space and if its functional form is known,
the system’s state can be described at any time with its mass and potential described
accurately. However, the distribution function is often not trivial to calculate.

In this case, the mass of the galaxy can be found by taking moments of the CBE,
multiplying each term of the CBE by the velocity and taking the integral over all
velocities to obtain Jeans Equation that relates the number density $\nu$ to the mean square of the velocities $v^2$ (Equation 1.4). The time derivative in the continuity equation is neglected as we assume that the system is in equilibrium. Most dSphs are pressure supported systems: the kinematic data for their stars show no sign of net rotation of the system (Strigari, 2009) and it is their random motions, described by the velocity dispersion, that support them against gravitational collapse. For spherical systems, the Jeans Equation 1.4 relates the total mass to the dispersion $\sigma$ of these random motions.

$$\frac{\partial (\nu v^2)}{\partial r} = -\nu \frac{d\phi}{dr}$$  \hspace{1cm} (1.4)

where it is assumed that the velocities have no preferred direction (tangential($v_\theta$) or radial $v_r$) and $v^2 = \sigma^2$. Since the potential $\phi$ is given as $r \frac{d\phi}{dr} = GM(r)/r = \nu_c^2$ the total mass of the system within a radius $r$ can be found as $M(r) = -\frac{r^2}{G\nu} \frac{d(v^2)}{dr}$. The Jeans equation given in Equation 1.4 can be projected to relate the 2 dimensional surface brightness $\Sigma$ to the line of sight velocity dispersion $\sigma$ instead of the 3 dimensional $\nu$ and $v^2$, and therefore can be used to directly relate observational quantities of the stellar distributions to the total potential and hence the mass distribution of the system.

### 1.3 Estimating dSph masses

While their proximity, suggested origins and amount of dark matter make them ideal laboratories for studying dark matter, the star formation mechanisms, kine-
matic structures and dynamical evolution of dSphs are not trivial to understand. Due to their low surface brightnesses, the modelling of their masses is often made by a simple Jeans analysis using only the central velocity dispersions and assuming isotropy (although see Lokas et al. (2005)). Furthermore, the observations often do not cover large enough radii where the stellar velocities would be tracing the complete mass of the haloes. For these reasons, in cases where the spherical isotropy or equilibrium assumptions fail, the Jeans analysis as described might give very crude answers.

In addition, accurate modelling would require well constrained orbital data for the dSphs around the Milky way for calculating the effects of the tidal force exerted on them by the Milky Way. But most dSph proper motions are not measured or are very uncertain as in the example of Carina where the velocities are measured as $\mu_\alpha = 15 \pm 9\text{mas/century}$ and $\mu_\delta = 22 \pm 9\text{mas/century}$ (Piatek, 2003). The degeneracy between the orbital and structural parameters would normally mean that mass and shape of the dSphs would be constrained by the orbit the dSph is on. However, the uncertainty in the observational data permits several different dynamical models that could be used to describe the data in some way. For instance on a given orbit, a dSph that has a very massive and extended halo could look like a lower mass but more concentrated halo as both would be only slightly affected by the Milky Way’s tidal force in their outer parts. Both of these would be losing some stars to the Milky Way, which would be seen in the form of a tidal tail that is forming in the outer radii and an increase of the velocity dispersion. This mass-orbital anisotropy degeneracy becomes even more severe for the poorly constrained orbital models and the distinction between different models should be made by first breaking the degeneracies (see Lokas et al., 2006). It is therefore important to understand the
correlations between orbital and structural parameters and how they relate to the uncertainties in the data.

In this thesis, we choose Carina for our study as the extent of its data and the possible existence of tidal tails in its outer parts gives us the opportunity to study the degeneracies and the effect of the tides.

### 1.4 Dwarf spheroidal galaxies

DSphs are ideal individual galaxies to search for dark matter, as their proximity to us in the Local Universe makes it possible for us to observe the projected density profiles of their stellar components despite their very low surface brightnesses. We can also measure the radial velocities for the individual stars in these systems which facilitates detailed dynamical modelling. Measurement of the chemical abundances of individual stars are also possible.

The luminosity of the classical dwarf spheroidal galaxies are in the range of $10^5$ to $10^8 L_\odot$ (Mateo, 1998), and as low as $10^3 L_\odot$ in some of the recently discovered ultrafaint dwarf galaxies (Martin, 2008). However, despite these luminosities that are comparable to those of large globular clusters, the extent of the dSphs out to kpc scales and their large velocity dispersions of the order of 10km/s, need to be explained since for a galaxy in virial equilibrium this yield a mass of $M_{dSph} = \frac{v^2}{G} \approx 10^8 M_\odot$ within the observable radius. Taking into account their low surface brightnesses, the dSphs would seem to be the most dark matter dominated systems known in the Universe (Mateo, 1998).
The possibility of dSphs being formed by the tidal shocking and stripping of globular clusters in the outer parts of merging galaxies has also been considered (Kroupa, 1997). Although recent numerical simulations have shown that many of the dSphs may not be immune to tidal disturbance by the Milky Way (e.g. Muñoz et al., 2008; Łokas, 2008), their observed properties such as high central velocity dispersions and long survival times under the tidal effect of the Milky Way still require the presence of massive dark-matter haloes which protect them against a complete tidal disruption (Klessen et al., 2003; Maschenko et al., 2007).

Over the past two decades, a significant amount of observational work has focussed on quantifying both the amount of dark matter in these systems, and its spatial distribution (e.g. Gilmore et al., 2007; Simon & Geha, 2007; Walker et al., 2007). In addition to their proximity and dark matter content, these acquired observational data make the dSphs our best nearby laboratories in which to test dark-matter theories.

### 1.5 Stellar structures of the dSphs and Ultra-faint dwarfs

Given that dSphs occupy the low luminosity end of the galaxy luminosity function (Schechter, 1976; Flint et al., 2001), their star formation histories provide useful insights into the star formation process. Although all dSphs have old stellar populations more than 10 billion years old, part of the discrepancy between the number
of observed dSphs and the subhaloes predicted in the simulations can be due to
the present-day dSphs being a population distinct from the early Universe dwarfs
which contributed to the formation of Milky Way's halo (Robertson et al., 2005;
Font et al., 2006), although see (Frebel et al., 2010). This view is supported by the
stellar abundances, which are significantly different between the Milky Way's halo
stars and those of the present-day dwarf spheroidal galaxies (e.g. Shetrone et al.,
2001; Helmi et al., 2006). We note however that rather than solving the missing
satellites problem, this argument only suggests that it is not the earliest dwarf pop-
ulation that was formed closest to the Milky Way that we expect to observe in the
present day.

Evidence of metallicity gradients has been found using spectroscopic estimates of
[Fe/H] (e.g. Tolstoy et al., 2004; Battaglia et al., 2006; Koch et al., 2006). Although
the lack of observed HI gas in dSphs is in agreement with the simulation results (e.g.
Moore et al., 1999), who found that these small galaxies would have used or lost
their gas very early, the fact that these systems had extended star formation periods
is at odds with these predictions. It is not clear how they kept their gas despite the
reionization, supernovae and tidal interactions and how all of them are observed in
a gasless state today. For example in the case of Carina, the youngest of the three
stellar populations observed is 500 Million years old and yet there is no gas ob-
erved in this system.

Analyses of spatial variations in colour-magnitude diagrams indicated that popu-
lation gradients could exist in dSphs (e.g. Harbeck et al., 2001). In the case of
the Sculptor dSph, the metal-rich and metal-poor populations have significantly
different spatial distributions and kinematics (Tolstoy et al., 2004; Battaglia et al.
Kinematically cold sub-populations have been detected in the Ursa Minor and Sextans dSphs (Kleyna, 2003; Walker, 2006; Battaglia et al., 2010). These substructures provide information about the star formation histories of the dSphs if they were formed at different epochs, as well as constraining the dynamics. For instance, in Canes Venatici I, an extended metal poor population and a centrally focused metal rich population were detected (Ibata et al., 2006; de Jong et al., 2008).

The very different velocity dispersions of these two populations yielded two very different mass estimates as the centrally concentrated population had a velocity dispersion near zero. Although this could mean that CVnI is not in equilibrium, the very low surface brightness of this dSph affects the significance of this detection and, as will be shown in this thesis, the population gradient could not be confirmed (also see Simon & Geha (2007)).

In addition, the substructures have implications on the shapes of the halo profiles. If the dark matter haloes of dSphs are cusped, as predicted by cosmological simulations, such internal substructures would be expected to be destroyed in a few Gyrs (Kleyna, 2003). They can survive much longer if their haloes are cored. If localised substructures were found to be common in dSphs, this could be difficult to reconcile with a picture in which dSphs occupy cusped haloes (although see Penarrubia et al., 2009, for evidence that triaxial haloes may permit the survival of substructures even in the presence of a cusp). Note however, that if the populations in CVnI were real, the central population might have been recently accreted and might not have had enough time to relax. The potential power of substructures to discriminate between dark matter halo models adds further motivation to our goal of establishing the level of confidence with which sub-populations can be detected in small data sets.
The ultra-faint dwarfs discovered in SDSS are thought to be the most dark matter dominated systems known as shown by kinematic data (Simon & Geha, 2007). The finding of an extremely metal poor star in Sculptor by Frebel et al. (2010) re-opened the case of the satellite galaxies to be the building blocks of Milky Way. In addition, their mass-to-light ratios seem to increase with decreasing luminosity. It has also been claimed that there is a common mass scale of $10^7 M_\odot$ inside 300pc for all dwarf satellites of the Milky Way (Strigari et al., 2008; Walker et al., 2009). Salvadori, & Ferrara (2009) showed that if the Ultra-faint dSphs were pre-reionisation-formed star-forming mini-haloes, their mean metallicites would be lower than the classical dSphs, as at the epoch of formation the gas was less metal-enriched, and that their Metallicity Distribution Function would be broader as their star formation period was more extended. Including radiative feedback together with gravitational calculations in their simulations they were able to reproduce the low star formation efficiencies of dSphs which could explain why, despite having the same mass scales inside 300pc, the luminosity of dSphs span a large range of values.

As the brightest of the ultra-faint dSph population, CVn I can help answering the outstanding question of whether the ultra-faint dSphs represent the low-luminosity tail of the dSph population, or are instead the brightest members of a population of hitherto unknown faint stellar systems, bridging dSphs and star clusters (e.g. $\omega$ Centauri: Bekki & Norris, 2006). The presence of multiple, distinct kinematic populations in a low-luminosity dSph would set it apart from the majority of low-luminosity star clusters. In addition, the presence of a spread in the stellar abundances would suggest an association with the brighter dwarf galaxies and would
also be interesting in terms of its implications for star formation.

## 1.6 Outline

The outline of this thesis is as follows:

In Chapter 2, we try to answer the following questions: Does Canes Venatici I dSph have multiple kinematic stellar populations which yield different estimates for its dynamical masses? Can we use the small data sets of dSphs to draw meaningful results on the structure of sub-populations?

In Chapter 3 we present the method that will be used in the parameter space search for modelling the mass of the Carina. This chapter contains background information on N-body simulations and Markov Chains. We then present our algorithm that studies the dynamics through N-body simulations, and show the numerical tests performed to guide our choices of the parameters to be used in the simulation codes.

In Chapter 4, we try to answer the following questions: Can the mass of Carina dSph be estimated with the currently available orbital parameters and kinematic data? What are the degeneracies, and can we break them? Can we distinguish between different inner halo slopes, and hence test the Λ-CDM model?

In Chapter 5, we try to answer the following questions: Are the results of our parameter space search using mid-resolution simulations in Chapter 4 robust, i.e. do we obtain comparable results using high resolution simulations? How do these
compare to Carina’s data? Can we detect and predict tidal signatures using two dimensional maps and velocity distributions?

Finally in Chapter 6, we summarise our findings, and briefly explain the future work that would be useful, building on the results we obtained in this thesis.
2

Kinematic sub-populations in
dwarf spheroidal galaxies

2.1 Introduction

It is now widely accepted that dwarf spheroidal (dSph) galaxies are the most dark-matter dominated stellar systems known in the Universe (e.g. Mateo, 1998). Over the past two decades, a significant amount of observational work has focussed on quantifying both the amount of dark matter in these systems, and its spatial distri-
Kinematic sub-populations in dwarf spheroidal galaxies

2.1. Introduction

bution (e.g. Gilmore et al., 2007; Simon & Geha, 2007; Walker et al., 2007). Although recent numerical simulations have shown that many of the dSphs may not be immune to tidal disturbance by the Milky Way (e.g. Muñoz et al., 2008; Łokas, 2008), their observed properties still require the presence of massive dark-matter haloes which protect them against complete tidal disruption. The dSphs thus provide us with nearby laboratories in which to test dark-matter theories.

Given that dSphs occupy the low luminosity end of the galaxy luminosity function, their star formation histories provide useful insights into the star formation process. Analyses of spatial variations in colour-magnitude diagram morphology provided early evidence of population gradients in a number of dSphs (e.g. Harbeck et al., 2001). Such gradients are to be expected in deep potential wells. More recently, evidence of metallicity gradients has been found using spectroscopic estimates of [Fe/H] (e.g. Tolstoy et al., 2004; Battaglia et al., 2006; Koch et al., 2006). In at least one case, that of the Sculptor dSph, the metal-rich and metal-poor populations have significantly different spatial distributions and kinematics (Tolstoy et al., 2004; Battaglia et al., 2008). Although little evidence of similar features has been found in other dSphs (e.g. Koch et al., 2006, 2007a,b), the presence of dynamically distinct stellar populations within dSphs, as well as the complex interplay between the dynamical, spatial and chemical properties of their stars, is of great interest as it has implications for star formation and galaxy evolution.

It is, however, important to note that although the hierarchical build-up of structure in the standard Λ-Cold Dark Matter (ΛCDM) paradigm implies that satellite galaxies contribute significantly to the stellar haloes of their hosts, detailed abundance studies of stars in the more luminous dSphs have demonstrated that
Kinematic sub-populations in
dwarf spheroidal galaxies

2.1. Introduction

Their properties are significantly different from those of the Milky Way halo (e.g. Shetrone et al., 2001; Helmi et al., 2006). Among the significant differences between the halo and the dSphs, the more important chemical differences are in the alpha-elements (Unavane et al., 1996; Venn et al., 2004). Thus, it appears that the primordial dwarf satellites, which many models assume existed and were disrupted to form the Milky Way halo, must have had stellar populations distinct from those seen in the present-day dSphs (Robertson et al., 2005; Font et al., 2006).

Given their high estimated mass-to-light ratios, the observed dSphs are usually identified with the large population of sub-haloes which are observed to surround Milky Way-sized haloes in cosmological simulations assuming a standard ΛCDM universe. However, it was noted early on that the number of dSphs around the Milky Way was much lower than the expected number of satellite dark-matter haloes (e.g. Moore et al., 1999). A number of possible explanations for the apparent lack of Milky Way satellites have been presented in the literature (e.g. Stoehr et al., 2002; Diemand et al., 2005; Moore et al., 2006; Strigari et al., 2007; Simon & Geha, 2007; Bovill, 2009). All these models are based on the reasonable postulate that out of the full population of substructures around the Milky Way, the observed dSphs are merely the particular subset that (for reasons of mass, orbit, formation epoch, re-ionisation, etc.) were able to capture gas, form stars, and survive any subsequent tidal interactions with the Milky Way.

The number of known dSph satellites of the Milky Way has increased significantly in the past few years, due to the discovery of nine new dSphs (Willman et al., 2005; Zucker et al., 2006a,b; Belokurov et al., 2006, 2007; Walsh et al., 2007) in the data from the Sloan Digital Sky Survey (SDSS; York et al., 2000). Since the SDSS
covers only about one quarter of the sky, it is thus likely that the total number of satellites surrounding the Milky Way may be at least a factor of four larger than previously thought, although the extrapolation from the SDSS survey to the whole sky requires careful analysis (see e.g. Tollerud, 2008). To compare the properties of the newly discovered satellites with those of sub-haloes in cosmological simulations, as well as to confirm their nature as true satellite galaxies of the Milky Way, as opposed to star clusters or disrupted remnants, spectroscopic observations of their member stars are essential in order to estimate dynamical masses from the observed stellar kinematics. The extremely low luminosities of these objects (in some cases as low as $10^3 L_\odot$: Martin, 2008) present significant observational challenges, as the kinematic data sets are small, making it difficult to obtain statistically significant results.

The Canes Venatici I (CVn I) dSph is the brightest and the most extended of the newly discovered population of ultra-faint SDSS dSphs (Zucker et al., 2006a). Ibata et al. (2006) presented spectra for a sample of CVn I member stars obtained using the DEIMOS spectrograph mounted on the Keck telescope. They identified two kinematically distinct stellar populations in this data set: an spatially extended metal-poor population with high velocity dispersion and a centrally-concentrated metal-rich population with a dispersion of almost zero. Their analysis of the mass of CVn I suggested that the two populations might not be in equilibrium, as the mass profiles obtained based on the individual populations were inconsistent with each other. A signature of the presence of multiple photometric populations was also identified by de Jong et al. (2008). However, a subsequent study of CVn I by Simon & Geha (2007), using a larger sample of Keck spectra, did not reproduce this bimodality.
An important outstanding question is whether the ultra-faint dSphs represent the low-luminosity tail of the dSph population, or are instead the brightest members of a population of hitherto unknown faint stellar systems, bridging dSphs and star clusters (e.g. ω Centauri: Bekki & Norris, 2006). The presence of multiple, distinct kinematic populations in a low-luminosity dSph would set it apart from the majority of low-luminosity star clusters. In addition, the presence of a spread in the stellar abundances would suggest an association with the brighter dwarf galaxies and would also be interesting in terms of its implications for star formation. It is thus important to determine whether the sub-population identified by Ibata et al. (2006) in CVn I is real. One goal of our study was to shed some light on this issue by using spectra obtained with a different spectrograph to those in the previous two studies of CVn I. In addition, we wanted to investigate the extent to which sub-populations can be reliably detected in the very small kinematic data sets that are observable for the ultra-faint dSphs.

In addition to their potential importance for probing the star formation histories of dSphs, kinematic substructures can be used to test another key feature of the hierarchical structure formation paradigm. Kinematically cold sub-populations have been detected in the Ursa Minor and Sextans dSphs (Kleyna, 2003; Walker, 2006). If the dark matter haloes of dSphs are cusped, as predicted by cosmological simulations, such internal substructures would be expected to be destroyed on timescales of at most a few Gyrs (Kleyna, 2003). They can survive much longer if their haloes are cored. If localised substructures were found to be common in dSphs, this could be difficult to reconcile with a picture in which dSphs occupy cusped haloes (although see Penarrubia et al., 2009, for evidence that triaxial haloes may permit the survival of substructures even in the presence of a cusp). As the brightest of the
ultra-faint dSph population, CVn I is a good candidate in which to search for possible sub-populations. The potential power of substructures to discriminate between dark matter halo models adds further motivation to our goal of establishing the level of confidence with which sub-populations can be detected in small data sets.

The outline of this chapter is as follows. In Section 2.2, we present a new kinematic data set for stars in CVn I, based on spectra obtained with the Gemini North telescope, and calculate a mass estimate for the galaxy from these data. In Section 2.3, we look for kinematic sub-populations in our data, and compare our findings with those of Ibata et al. (2006). Section 2.3.2 discusses the general detectability of sub-populations in small kinematic data sets. Finally, in Section 2.4, we draw some general conclusions and suggest possible differences between the two data sets for CVn I that we have compared.

### 2.2 Canes Venatici I

#### 2.2.1 Data reduction

Twenty eight stars in the CVn I dSph were observed on 2007 March 26 and 2007 April 7 and 8 using the GMOS-N spectrograph mounted on the Gemini North telescope. Our targets were chosen by cross-matching of GMOS-N pre-images (taken in the i-band) with existing SDSS photometry. As Figure 2.1 shows, all the selected stars lie in the red giant branch (RGB) region of the CVn I colour-magnitude diagram (CMD). A total of three GMOS slit-masks were observed, with the spectra centred on the spectral region containing the Ca triplet region (around 860nm). Our
Kinematic sub-populations in
dwarf spheroidal galaxies

2.2. Canes Venatici I

Figure 2.1: SDSS \((g-i, i)\) color-magnitude diagram for stars in a field of radius 15 arcmin centred on CVn I. Our likely CVn I members are indicated as solid triangles. The stars which are excluded from our analysis (see Table 2.1) are shown as open circles. The coordinates of the centre of CVn I are \(RA=13\,28\,03.5\) and \(Dec=33\,33\,21\) (Zucker et al., 2006a).

Masks covered three distinct fields in CVn I. Figure 2.2 shows the locations of the fields relative to the spatial distribution of stars in CVn I. The masks were cut with slitlets of width 0.75 arcsec. Figure 2.3 shows the radial distribution of the 28 stars observed.

The GMOS detector consists of three adjacent CCDs. As the dispersion axis of the slits is perpendicular to the spaces between the CCDs, the spectra contain gaps corresponding to the inter-CCD gaps. In order to achieve continuous wavelength coverage throughout the spectral region of interest, each mask was observed in two configurations with different central wavelengths (855nm and 860nm). All observations were taken using the R831+ G5302 grating and CaT G0309 filter, with \(2\times4\) spectral and spatial binning, respectively. The spectra thus obtained have a nominal resolution of 3600. The three fields were observed for a total of 10,800s, 9,000s and 12,600s, respectively, with the observations divided into individual exposures.
2.2. Canes Venatici I

Figure 2.2: Distribution of our GMOS target fields in CVn I. The data points show the positions of stars satisfying our CMD selection cut. The slight excess of stars in the region $-10 < X < 10$ and $-5 < Y < 5$ indicates the location of the main body of CVn I. The ellipse shows the half-light radius of the system, with semi-major axis 8.9 arcmin. The open circles are our observed stars.

Figure 2.3: Cumulative radial distribution of the observed stars.
of 1800s to facilitate cosmic ray removal.

The raw data were reduced using the standard *gemini* reduction package which is run within the Image Reduction and Analysis Facility (IRAF) environment. All data were first bias subtracted and flat-field corrected. The individual spectral traces were identified from flat field images (obtained using quartz halogen continuum lamp exposures). The wavelength calibration of the spectra was performed using CuAr lamp exposures adjacent in time to the science exposures as calibration frames. The typical r.m.s. uncertainty in the wavelength calibration, obtained by fitting a polynomial to the line positions in the CuAr spectra, was 0.01Å, which corresponds to a velocity error of ~ 0.4 km s$^{-1}$ at a wavelength of 860nm. This wavelength solution was then applied to the reduced science spectra. Sky subtraction was performed by using the sky flux in the regions of the slit not dominated by light from the target to estimate the sky spectrum. Finally, the object spectra were extracted from the CCD images using a fifth order Chebyshev polynomial fit. Figure 2.4 shows examples of a good quality spectrum (top panel), a low quality spectrum (middle panel), and typical quality spectrum (bottom).

### 2.2.2 Velocities

The velocities of the stars were calculated using the *fxcor* task in IRAF to cross-correlate the stellar spectral lines with the lines in a template Ca triplet spectrum. The synthetic template consisted of three Gaussian lines at the wavelengths of the

---

1IRAF is distributed by the National Optical Astronomy Observatories, which are operated by the Association of Universities for Research in Astronomy Inc. (AURA), under cooperative agreement with the National Science Foundation.
Ca triplet lines, whose widths were chosen to match those typical of RGB stars. We first cross-correlated the individual science exposures as a preliminary diagnostic of whether any spectra were obviously anomalous and should be excluded. As none of the spectra seemed to have serious problems, all frames were used in the velocity calculations, and we combined all heliocentric-corrected exposures of the same mask together in order to increase the signal-to-noise ratio.

The \texttt{fxcor} task returns estimated velocity uncertainties which are based on the Tonry-Davis Ratio for the fitted cross-correlation peak (Tonry, 1979). These errors are often found not to be an accurate reflection of the true uncertainties (see e.g. Kleyna et al., 2002; Muñoz et al., 2005). In order to estimate the actual uncertainty in our velocity determinations, we measured separately the velocities $v_1$ and $v_2$ for the spectra with central wavelengths 855nm and 860nm, respectively. We combined these estimates to obtain the mean velocity for each star $\bar{v} = 0.5(v_1 + v_2)$ and defined
Kinematic sub-populations in dwarf spheroidal galaxies

2.2. Canes Venatici I

A χ² statistic via

\[ \chi^2 = \frac{(v_1 - \bar{v})^2}{(dv_1)^2} + \frac{(v_2 - \bar{v})^2}{(dv_2)^2}, \]  

(2.1)

where \(dv_1\) and \(dv_2\) are the formal errors returned by \texttt{fxcor}. We then rescaled the velocity errors in our sample by a factor \(f\) so that the sum of equation 2.1 over all stars was \(2N\), where \(N\) is the size of the velocity sample. Finally, using the rescaled errors, we calculated \(P(\chi^2)\) for each star, using the routine \texttt{gammq} from Numerical Recipes [Press et al., 1991]. The final velocities and errors are given in Table 2.1.

Following the error rescaling, only one star (13:28:24.02, +33:35:32.7) was found to have an extremely low value of \(P(\chi^2) (< 10^{-4})\). As Table 2.1 shows, this star also has the largest velocity in the sample and a relatively large estimated velocity error, possibly due to its low signal-to-noise ratio, and we therefore excluded it from our final sample. We also excluded one star which has very different radial velocity, \(v_R = -39.1 \text{ km s}^{-1}\), compared to the mean velocity of the rest of our target stars (25.8 ± 0.3 km s\(^{-1}\); see Section 2.2.4). Figure 2.5 shows the velocity histogram for our final sample, consisting of 26 stars. We note that our sample includes 8 stars from the high signal-to-noise ratio sample of Ibata et al. (2006) sample; the upper panel of Figure 2.6 is a histogram representing the difference between the velocities of these stars from both studies normalised by their \(1\sigma\) measurement uncertainties. Thus, we calculate \(\Delta v/\langle \sigma \rangle\) as \([v_{\text{Keck}} - v_{\text{GMOS}}]/\sqrt{dv_{\text{Keck}}^2 + dv_{\text{GMOS}}^2}\), after applying a velocity shift of -3.4 km s\(^{-1}\) to our estimates in order to bring the medians of the two data sets into agreement. The plot shows that apart from the two outliers, at 8.6 and 4.6σ, with very different velocities in the two sets, the differences are normally distributed. The outliers are possibly stars in binary systems that have changed their velocity between the two observations. The Ibata et al. (2006) data were taken in May 2006, i.e., around ten months earlier than our data. The observed velocity dif-

36
ferences of $8 - 10\text{ km s}^{-1}$ over this baseline are consistent with tight binary orbits.

### 2.2.3 Metallicities

It is well-established that the line strength of the near-infrared Ca triplet lines in the spectrum of an RGB star can be used to estimate the metallicity $[\text{Fe/H}]$ of the star (e.g., Armandroff & Zinn, 1988; Armandroff, 1991; Carrera et al., 2007; Bosler et al., 2007). We note that the accuracy of this method may be less reliable when extrapolating below metallicities of $\sim -2.2$ where globular cluster calibrators are missing (Koch et al., 2008), although comparisons of high-vs-low resolution data by (Battaglia et al., 2008) have shown that CaT-based estimates may be correct down to $[\text{Fe/H}] \sim -3$. In practice, we normalized the spectra using a seventh order Legendre polynomial, fitted each of the triplet lines using a Penny function (see Cole et al., 2004), and integrated the profile over the standard band passes of (Armandroff & Zinn, 1988). The final $[\text{Fe/H}]$ metallicities, on the scale of Carretta & Gratton (1997), were calculated using the calibration of Rutledge et al. (1997a,b), namely

$$[\text{Fe/H}] = -2.66 + 0.42[\Sigma W + 0.64(V - V_{\text{HB}})],$$  \hspace{1cm} (2.2)

where we parameterised the line strength of the Ca triplet as

$$\Sigma W = 0.5w_1 + w_2 + 0.6w_3,$$  \hspace{1cm} (2.3)

where $w_1$, $w_2$ and $w_3$ are the widths of the individual lines. In Eq. 2.2, $V$ is the $V$-band magnitude of the star, and $V_{\text{HB}}$ is the magnitude of the horizontal branch.
of the system. For the latter, we used a value of $V_{HB} = 22.4$, obtained by visual inspection of the (V-I,V) colour-magnitude diagram from transformed SDSS data of CVnI. We note that this is very similar to the value of $V_{HB} = 22.5$ used by Martin et al. (2008a). The uncertainty of $\sim 0.1$ magnitudes in $V_{HB}$ gives rise to a negligible additional uncertainty in our [Fe/H] estimates. The random errors on the [Fe/H] metallicities were calculated using the formalism of Cayrel (1988) for the errors on the single line widths, and are based on the spectral signal-to-noise ratio. These were then propagated through the calibration equations, accounting for photometric errors. The final metallicity estimates are given in Table 2.1. The lower panel in Figure 2.6 plots the differences between our estimated metallicities and those found by Ibata et al. (2006) for those stars we have in common. As for the velocities in the top panel, the [Fe/H] differences are shown normalised by their combined 1σ error, following the application of a shift of [Fe/H]=$-0.105$ to bring the median values into agreement. As in the case of the comparison between the velocity data (top panel of the figure), the [Fe/H] differences are normally distributed indicating very good agreement between the two sets of [Fe/H] measurements.

Figure 2.7 shows the distribution of velocity versus [Fe/H] for our CVn I sample. The error-weighted mean [Fe/H] is $-1.9 \pm 0.02$. This value agrees well with previous estimates of the mean [Fe/H] for CVn I, which lie in the range [Fe/H]=$-2$ to [Fe/H]=$-2.1$ (Ibata et al., 2006; Simon & Geha, 2007; Kirby, 2008). We note that all previous studies of CVn I have found a significant spread in [Fe/H], of order $0.25 - 0.5$ dex (Ibata et al., 2006; Simon & Geha, 2007; Kirby, 2008). Figure 2.7 suggests that our data favour intrinsic dispersions towards the lower end of this range. As the figure shows, there appear to be no obvious correlations between velocity and [Fe/H] in our sample.
2.2. Canes Venatici I

Figure 2.5: Velocity histogram for our data set of 26 likely members of CVn I. The two stars which are excluded from our analysis (see Table 2.1) have not been included in this figure.
2.2. Canes Venatici I

**Figure 2.6:** Histograms of the normalised differences in estimated velocity (top panel) and metallicity (bottom panel) for the ten stars observed both by us and Ibata et al. (2006). Differences have been normalised by the combined error from the two estimates (see text for details). Apart from the two significant velocity outliers, both distributions are close to the overplotted unit Gaussians.

**Figure 2.7:** Line of sight velocity and [Fe/H] for our sample of likely CVn I members. The two stars which are excluded from our analysis (see Table 2.1) have not been included in this figure.
## Kinematic sub-populations in dwarf spheroidal galaxies

### 2.2. Canes Venatici I

<table>
<thead>
<tr>
<th>α (J2000)</th>
<th>δ (J2000)</th>
<th>g</th>
<th>i</th>
<th>(v_t)</th>
<th>(d_{W})</th>
<th>(d\Sigma W)</th>
<th>[Fe/H]</th>
<th>d[Fe/H]</th>
</tr>
</thead>
<tbody>
<tr>
<td>13 28 10.07</td>
<td>+33 33 41.6</td>
<td>20.7</td>
<td>19.7</td>
<td>29.2</td>
<td>1.7</td>
<td>3.12</td>
<td>0.07</td>
<td>−1.94</td>
</tr>
<tr>
<td>13 28 10.31</td>
<td>+33 33 06.0</td>
<td>20.3</td>
<td>19.1</td>
<td>39.9</td>
<td>1.6</td>
<td>3.02</td>
<td>0.05</td>
<td>−2.10</td>
</tr>
<tr>
<td>13 28 16.78</td>
<td>+33 32 54.0</td>
<td>21.1</td>
<td>20.0</td>
<td>21.9</td>
<td>2.4</td>
<td>2.83</td>
<td>0.08</td>
<td>−1.96</td>
</tr>
<tr>
<td>13 28 15.00</td>
<td>+33 32 01.6</td>
<td>21.3</td>
<td>20.3</td>
<td>30.4</td>
<td>2.1</td>
<td>2.72</td>
<td>0.14</td>
<td>−1.94</td>
</tr>
<tr>
<td>13 28 18.31</td>
<td>+33 33 17.3</td>
<td>21.2</td>
<td>20.4</td>
<td>14.5</td>
<td>1.9</td>
<td>2.96</td>
<td>0.11</td>
<td>−1.85</td>
</tr>
<tr>
<td>13 28 08.49</td>
<td>+33 33 29.4</td>
<td>21.5</td>
<td>20.6</td>
<td>35.2</td>
<td>2.4</td>
<td>2.94</td>
<td>0.19</td>
<td>−1.77</td>
</tr>
<tr>
<td>13 28 08.82</td>
<td>+33 34 41.2</td>
<td>20.4</td>
<td>19.2</td>
<td>25.9</td>
<td>1.5</td>
<td>3.22</td>
<td>0.05</td>
<td>−1.99</td>
</tr>
<tr>
<td>13 28 10.35</td>
<td>+33 34 27.7</td>
<td>20.7</td>
<td>19.5</td>
<td>24.4</td>
<td>1.5</td>
<td>3.55</td>
<td>0.06</td>
<td>−1.77</td>
</tr>
<tr>
<td>13 28 16.23</td>
<td>+33 34 09.9</td>
<td>20.7</td>
<td>19.5</td>
<td>24.8</td>
<td>1.6</td>
<td>3.61</td>
<td>0.10</td>
<td>−1.68</td>
</tr>
<tr>
<td>13 28 11.31</td>
<td>+33 35 06.1</td>
<td>21.3</td>
<td>20.2</td>
<td>21.9</td>
<td>1.8</td>
<td>3.77</td>
<td>0.15</td>
<td>−1.52</td>
</tr>
<tr>
<td>13 28 07.19</td>
<td>+33 36 05.4</td>
<td>20.7</td>
<td>19.5</td>
<td>23.8</td>
<td>1.0</td>
<td>2.51</td>
<td>0.08</td>
<td>−2.21</td>
</tr>
<tr>
<td>13 28 13.74</td>
<td>+33 35 55.6</td>
<td>20.2</td>
<td>18.9</td>
<td>21.3</td>
<td>1.3</td>
<td>3.19</td>
<td>0.06</td>
<td>−2.07</td>
</tr>
<tr>
<td>13 28 24.02</td>
<td>+33 35 32.7</td>
<td>21.4</td>
<td>20.5</td>
<td>44.0</td>
<td>3.2</td>
<td>2.23</td>
<td>0.18</td>
<td>−2.10</td>
</tr>
<tr>
<td>13 27 38.76</td>
<td>+33 32 55.3</td>
<td>20.5</td>
<td>19.3</td>
<td>12.3</td>
<td>1.4</td>
<td>2.65</td>
<td>0.05</td>
<td>−2.20</td>
</tr>
<tr>
<td>13 27 33.80</td>
<td>+33 32 57.3</td>
<td>21.2</td>
<td>20.1</td>
<td>31.9</td>
<td>1.4</td>
<td>3.20</td>
<td>0.09</td>
<td>−1.77</td>
</tr>
<tr>
<td>13 27 28.41</td>
<td>+33 33 26.3</td>
<td>21.3</td>
<td>20.2</td>
<td>−39.1</td>
<td>2.0</td>
<td>4.26</td>
<td>0.13</td>
<td>−1.31</td>
</tr>
<tr>
<td>13 27 28.29</td>
<td>+33 32 10.6</td>
<td>20.2</td>
<td>18.8</td>
<td>26.9</td>
<td>1.6</td>
<td>4.06</td>
<td>0.13</td>
<td>−1.71</td>
</tr>
<tr>
<td>13 27 30.35</td>
<td>+33 32 04.3</td>
<td>21.5</td>
<td>20.6</td>
<td>15.6</td>
<td>2.1</td>
<td>3.01</td>
<td>0.13</td>
<td>−1.75</td>
</tr>
<tr>
<td>13 27 30.78</td>
<td>+33 29 38.9</td>
<td>21.1</td>
<td>20.1</td>
<td>19.8</td>
<td>1.8</td>
<td>4.12</td>
<td>0.14</td>
<td>−1.42</td>
</tr>
</tbody>
</table>

*Summary of the properties of our CVn I data. Columns give (1) Right ascension; (2) Declination; (3,4) SDSS g, i magnitudes; (5,6) radial velocity and error, in km s\(^{-1}\); (7,8) combined equivalent width of Ca triplet lines, with error, obtained using equation \[\text{eq.2.3}\] (9,10) estimated metallicity [Fe/H], with error, obtained using equation \[\text{eq.2.2}\].*
Kinematic sub-populations in
dwarf spheroidal galaxies

2.2. Canes Venatici I

<table>
<thead>
<tr>
<th>RA</th>
<th>Dec</th>
<th>V</th>
<th>V disp</th>
<th>V errately</th>
<th>V errandar</th>
<th>Sigma</th>
</tr>
</thead>
<tbody>
<tr>
<td>13 27 31.21</td>
<td>+33 29 59.2</td>
<td>19.9</td>
<td>18.4</td>
<td>14.8</td>
<td>1.1</td>
<td>4.01</td>
</tr>
<tr>
<td>13 27 16.19</td>
<td>+33 32 22.2</td>
<td>20.6</td>
<td>19.4</td>
<td>21.0</td>
<td>1.2</td>
<td>2.98</td>
</tr>
<tr>
<td>13 28 52.91</td>
<td>+33 29 26.0</td>
<td>20.3</td>
<td>19.0</td>
<td>31.3</td>
<td>1.7</td>
<td>3.11</td>
</tr>
<tr>
<td>13 28 46.21</td>
<td>+33 30 49.7</td>
<td>21.4</td>
<td>20.5</td>
<td>35.9</td>
<td>1.9</td>
<td>3.29</td>
</tr>
<tr>
<td>13 28 50.64</td>
<td>+33 31 20.9</td>
<td>20.5</td>
<td>19.3</td>
<td>42.1</td>
<td>1.6</td>
<td>2.93</td>
</tr>
<tr>
<td>13 28 50.29</td>
<td>+33 32 31.6</td>
<td>20.3</td>
<td>18.9</td>
<td>37.0</td>
<td>1.3</td>
<td>4.24</td>
</tr>
<tr>
<td>13 28 45.29</td>
<td>+33 31 35.1</td>
<td>21.4</td>
<td>20.4</td>
<td>37.1</td>
<td>2.1</td>
<td>3.05</td>
</tr>
<tr>
<td>13 28 53.47</td>
<td>+33 33 22.5</td>
<td>21.1</td>
<td>20.1</td>
<td>22.1</td>
<td>1.9</td>
<td>2.12</td>
</tr>
<tr>
<td>13 28 44.26</td>
<td>+33 34 11.8</td>
<td>20.0</td>
<td>18.7</td>
<td>31.0</td>
<td>1.7</td>
<td>2.47</td>
</tr>
</tbody>
</table>

Table 2.1: (*1) This velocity is calculated as $v = 57.3 \pm 4.37 \text{ km s}^{-1}$ in the first observation at $\lambda = 855\text{nm}$ and $v = 29.1 \pm 4.61 \text{ km s}^{-1}$ in the second observation at $\lambda = 860\text{nm}$. Although an average value $v = 44 \text{ km s}^{-1}$ with $dv = 3.2 \text{ km s}^{-1}$ can be found, the significant difference results in a very small likelihood of this is $P(\chi^2) = 9 \times 10^{-6}$, (compared to the typical value of $P(\chi^2) = 0.33$). We therefore excluded this star from our analysis. (*2) This star ($v = -39.1 \text{ km s}^{-1}$) is a clear outlier from the mean velocity of $v = 25.8 \pm 0.3 \text{ km s}^{-1}$ and is also excluded from our analysis. V and I magnitudes are calculated from SDSS photometry using the transformations of Lupton (2005: http://www.sdss.org/dr4/algorithms/sdssUBVRITransform.html#Lupton). Lupton derived these equations by matching photometry from SDSS Data Release 4 to Stetson’s published photometry.

2.2.4 Mass calculation

In order to estimate the mass of CVn I, we calculate the velocity dispersion of the system using the new velocity set discussed in the previous section. We use a maximum likelihood method (e.g. Kley et al., 2004) to calculate the velocity dispersion and the mean velocity of our data. We apply an iterative $3\sigma$ cut in velocity - however, we note that this did not remove any stars (ie. it converged after a single
iteration). Based on the CMD in Figure 2.1, we do not expect a significant foreground contamination in our velocity sample, and we therefore use all 26 of our stars when estimating the dispersion. We find a dispersion of $\sigma = 7.9^{+1.3}_{-1.1}$ km s$^{-1}$ and a mean velocity of $v = 25.8 \pm 0.3$ km s$^{-1}$. The latter is somewhat smaller than the value of $v = 30.9 \pm 0.6$ km s$^{-1}$ found by Simon & Geha (2007). Ibata et al. (2006) found dispersions of 13.9 km s$^{-1}$ and 0.5 km s$^{-1}$ for the two populations they identified. As we discuss below, we do not find evidence of multiple populations in our data and we therefore quote only a single value for the dispersion.

We now use our dispersion measurement to constrain the mass of CVn I. In order to proceed we need to parameterise the spatial distribution of our data. We assume that our tracer population is drawn from a Plummer distribution, and we find the scale-length for which the likelihood of the positions of our tracer data set is maximised. Based on the positions of our tracer stars only, we find a Plummer scalelength of $a = 4.6^{+8.2}_{-2.2}$ arcmin. While this is small compared to the value of $8.5 \pm 0.5$ arcmin found by Zucker et al. (2006a) and $8.9 \pm 0.4$ arcmin found by Martin (2008) for the full stellar distribution, it is within the limits allowed by the uncertainties. The mass is then calculated using the isotropic Jeans equation (Binney & Tremaine, 1998, eq. 4.56), under the additional assumptions of spherical symmetry and a constant velocity dispersion.

We find a mass of $4.4^{+1.6}_{-1.1} \times 10^7 M_\odot$ within the volume probed by our data (i.e. out to a radius of 720 parsecs, 11 arcmin). The mass-to-light ratio is calculated assuming a luminosity of $L = (2.3 \pm 0.3) \times 10^5 L_\odot$ (Martin, 2008). Assuming symmetric errors on $M$ we find $M/L = 192 \pm 76 M_\odot/L_\odot$. If we take the value of the scale-length reported by Zucker et al. (2006a) to be that of our tracers, we obtain a mass...
of $3.1^{+0.9}_{-0.8} \times 10^7 M_\odot$. We note that both these estimates are in reasonable agreement with the mass of $M = (2.7 \pm 0.4) \times 10^7 M_\odot$ reported by Simon & Geha (2007) using their larger data set. Ibata et al. (2006) obtained two very different mass estimates using the distinct populations which they identified in their data. An important point to keep in mind while dealing with small data sets is that in spherically symmetric systems, the Jeans equations remain valid for density-weighted averages of the spatial distributions, velocity dispersion profiles, and velocity anisotropy profiles of multiple tracer populations (Strigari et al., 2007). Thus, it is legitimate to use a data set which may contain multiple sub-populations when estimating the mass of the system. As long as all sub-populations are in dynamical equilibrium, this estimate will be more reliable than the noisier estimates based on the smaller, individual populations.

2.3 Sub-populations

2.3.1 Canes Venatici I

As we noted above, Ibata et al. (2006) identified two kinematically distinct populations in CVn I. Given the potential importance of sub-populations in dSphs discussed in the introduction, we now investigate whether our data exhibit any evidence of multiple populations. Although the wide scatter in the abundances in Figure 2.7 might be due to an extended star formation period, making CVn I similar to the classical, non-ultrafaint dSphs, no clear signature of distinct sub-populations is seen. In order to confirm this visual impression more quantitatively, we fitted our velocity distribution with multiple Gaussians and tested the significance of the
fits using Monte Carlo realisations of our data. Our approach, which is essentially a likelihood ratio test, is similar to the KMM test (Ashman et al., 1994), which is designed to detect multiple Gaussian populations with different means and dispersions within a single data set, although unlike the KMM test we do not include a determination of which sub-population the individual stars belong to.

The first step of this process was to fit a single Gaussian to our velocity data. We then repeated the fit for a two-Gaussian model in which a fraction of the data belonged to a population with mean and dispersion , and the remaining data had mean and dispersion . As expected, the two-Gaussian model yielded higher likelihoods. In order to determine whether this was only due to the increased number of fitting parameters or was a real detection, we tested the significance of the results with artificial data. To do this, we generated 1000 data sets of 26 stars drawn from a single Gaussian and calculated the improvement of the fit with a two-Gaussian model. The improvement between fits is quantified as a parameter which is the logarithm of the ratios of likelihoods and of the two fits. The distribution of probability ratios are shown in Figure 2.8 and in Figure 2.9. We note that the precise parameter values in the fits are dependent on the grid chosen for the parameter search but the results converge for finer grids - for our finest grids the velocity parameters vary by $\lesssim 0.5 \text{ km s}^{-1}$.

The values we obtained for the GEMINI and KECK data (for the same grid spacings) are shown as the single dot in each panel of Figure 2.8 and Figure 2.9. In the left panel of Figure 2.8, this point coincides with the peak obtained by fitting two Gaussians to artificial data consisting of a single population. We conclude that we do not see evidence for a second population in our data. Figure 2.9 shows the
equivalent test for the Ibata et al. (2006) data (where we have taken the data for their 26 stars with S/N > 15, as listed in Table 2 of Martin et al., 2007). In this case the improvement is larger than would be expected to arise by chance in a single-Gaussian data set. We note that a similar result is obtained when the two data sets are combined. Therefore we conclude that there is evidence of a second population containing 40 per cent of the total number of stars, in the Ibata et al. (2006) data set, at almost the 3σ confidence level. We find that the dispersions of these populations are around 1.0 km s\(^{-1}\) and 13. km s\(^{-1}\), respectively. Although these values are similar to those found by Ibata et al. (2006), we note that the populations we have identified may be different to those in that paper, as in that case the separation of the populations included an explicit velocity cut.

The middle panels of Figure 2.8 and Figure 2.9 show that performing the same two-Gaussian test for the [Fe/H] distributions does not suggest any evidence of significant sub-populations in either our data or the Ibata et al. (2006) data. The best-fitting populations to the Ibata et al. (2006) data are found to have mean metallicities of [Fe/H]=−2.1 and [Fe/H]=−1.8, compared to [Fe/H]=−2.0 and [Fe/H]=−1.5 found by Martin (2008). In the right-hand panels of the Figure s we show the results of two-dimensional tests, i.e. instead of looking for subpopulations in the kinematics or [Fe/H] independently, we instead describe each population as a combination of two Gaussians, one each in velocity and [Fe/H]. As the figure shows, in the case of CVn I, this approach yields the same results regarding the presence of multiple populations as we obtained using the velocity data alone.
Kinematic sub-populations in dwarf spheroidal galaxies

2.3. Sub-populations

Figure 2.8: Distribution of likelihood ratios $\Delta P = \log P_2 - \log P_1$ between a single-Gaussian fit ($P_1$) and a two-Gaussian fit ($P_2$) to 1000 Monte Carlo realisations of 26 stars drawn from a single Gaussian distribution. The single dots indicate the values obtained for our GMOS-N data. The left-hand panel shows results for the velocity data, the middle panel shows results for the $[\text{Fe}/H]$ data. In the right-hand panel we fit the velocity and $[\text{Fe}/H]$ data simultaneously - each population is assumed to be a Gaussian in both velocity and $[\text{Fe}/H]$. There is no evidence of multiple populations in our data. See text for a detailed discussion.

Figure 2.9: Distribution of likelihood ratios $\Delta P = \log P_2 - \log P_1$ between a single-Gaussian fit ($P_1$) and a two-Gaussian fit ($P_2$) to 1000 Monte Carlo realisations of 26 stars drawn from a single Gaussian distribution. The single dots indicate the values obtained for the Keck data of Ibata et al. (2006). The left-hand panel shows results for the velocity data, the middle panel shows results for the $[\text{Fe}/H]$ data. In the right-hand panel we fit the velocity and $[\text{Fe}/H]$ data simultaneously - each population is assumed to be a Gaussian in both velocity and $[\text{Fe}/H]$. In contrast to our GMOS-N data, a sub-population containing a fraction of 40 per cent of the stars is detected in the Ibata et al. (2006) velocity data. See text for a detailed discussion.
2.3. Sub-populations

The detectability of a sub-population depends on i) the total number of stars in the data set; ii) the fraction of stars in the sub-population; iii) the difference in velocity dispersion between the populations; iv) the observational errors on the velocities. We investigate the importance of each of these in turn.

The total number of the stars is crucial for the detection of multiple populations. Figure 2.10 shows three tests done with data sets of $N = 120, 60$ and $30$ stars. In
each panel a comparison is made between data sets that have either a single population or two populations containing equal numbers ($N/2$) of stars. Motivated by the case of CVn I, we consider data sets having two sub-populations with $\sigma = 13$ km s$^{-1}$ and $\sigma = 1$ km s$^{-1}$. The two populations are thus clearly distinct, and we are thus isolating the effect of the sample size in the result. In Figure 2.10 we plot the improvement in probability obtained using a two-Gaussian fit to the single and double populations as dashed and solid histograms, respectively. We determine the $1\sigma$ and $3\sigma$ range of the distribution of values obtained from the single population (control) sample, and define a $1\sigma$ ($3\sigma$) detection of a sub-population to be one in which $\Delta P$ is larger than the $1\sigma$ ($3\sigma$) limits of the control sample. Although the difference between the dispersions is large, as we reduce the sample size, the significance of the detection of multiple populations decreases, as would be expected. Nevertheless, even for $N = 30$ stars, in 34.8 per cent (97.5 per cent) of cases, the subpopulation is detected at the $3\sigma$ ($1\sigma$) confidence level.

So far the tests have been made with sub-populations of equal sizes, i.e., the fraction of the sizes is set at $f=0.5$. We now look at the detectability of sub-populations with different values of $f$. This is important for the CVn I populations, since it is possible that a cold population in the centre could have been missed in our sample if it contained a smaller number of stars. Our preliminary tests showed that a cold population could not be detected even in a large sample if it only made up $\sim 0.1$ of the total number of stars. Figure 2.11 shows results for cold populations with fractional sizes $f = 0.3$, $f = 0.5$ and $f = 0.7$. In this test, the dispersions of the sub-populations are $13$ km s$^{-1}$ and $1$ km s$^{-1}$ and the velocity error is $2$ km s$^{-1}$. For 120 stars, a $3\sigma$ detection was made for all the samples with a cold population of fractional size $f = 0.5$ and $f = 0.7$. We found (see Figure 2.11) that when the cold
Kinematic sub-populations in
dwarf spheroidal galaxies

population has a smaller fractional size in the sample i.e. $f = 0.3$, it was detected in
75.1 per cent (99.8 per cent) of cases at the $3\sigma$ ($1\sigma$) level. It is thus easier to detect
a sub-population if its dispersion is larger than that of the main population, rather
than a cold sub-population.

We next consider the impact of velocity errors on our ability to detect multiple pop-
ulations with similar velocity dispersions. The sub-populations in this case have
$\sigma_1 = 7 \text{ km s}^{-1}$ and $\sigma_2 = 4 \text{ km s}^{-1}$. As Figure 2.12 shows, decreasing the errors from
$dv = 2 \text{ km s}^{-1}$ to $dv = 1 \text{ km s}^{-1}$ gives rise to a small change in the distribution of
$\Delta P$ values. However, this does not lead to a significant increase in the probability
of detecting the multiple populations. We therefore conclude that, in this situation,
reducing the velocity errors to $1 \text{ km s}^{-1}$ (similar to the CVn I data) does not improve
our ability to identify sub-populations.

Next, to see the effect of the difference between the velocity dispersions of the
populations we investigate samples in which the main population has a disper-
sion of $15 \text{ km s}^{-1}$ while the cold sub-populations have dispersions ranging from
$\sigma = 1 \text{ km s}^{-1}$ to $14 \text{ km s}^{-1}$. We consider two sample sizes, with a total num-
ber of either 120 or 60 stars. We find that even for a relatively large sample
of stars ($N = 120$), a $3\sigma$ detection is possible for all the samples only when
$\sigma_1/\left(\sigma_1 - \sigma_2\right) \leq 1.1$, i.e. when the velocity dispersions of the individual populations
are $\sigma_1 = 15 \text{ km s}^{-1}$ and $\sigma_2 = 1 \text{ km s}^{-1}$. A $1\sigma$ level detection is possible for all 1000
samples for $\sigma_1/\left(\sigma_1 - \sigma_2\right) \leq 1.3$, in which case the sub-populations’ dispersions are
$\sigma_1 = 15 \text{ km s}^{-1}$, $\sigma_2 = 3 \text{ km s}^{-1}$. However, populations with $\sigma_1/\left(\sigma_1 - \sigma_2\right) \leq 1.4$
and $\sigma_1/\left(\sigma_1 - \sigma_2\right) \leq 1.9$ can be detected at $3\sigma$ and $1\sigma$ levels for 68 per cent of the
1000 samples. For a smaller sample ($N = 60$), a $3\sigma$ detection for all samples is not
possible for even the largest ratios of of $\sigma_1/\sigma_2$. In this case $3\sigma$ and $1\sigma$ detections for 68 per cent of the samples require $\sigma_1/(\sigma_1 - \sigma_2) \leq 1.3$ and $\sigma_1/(\sigma_1 - \sigma_2) \leq 1.7$ respectively. We note that the claimed CVn I populations in [Ibata et al. (2006)] have an even more extreme dispersion difference $\sigma_1/(\sigma_1 - \sigma_2) = 1.04$. A Monte Carlo experiment with 26 stars and this dispersion ratio shows that in this case populations can be identified with $3\sigma$ confidence in 90.5 per cent of the samples. Table 2.2 summarises our results for the full range of dispersions ratios we have considered.
Kinematic sub-populations in dwarf spheroidal galaxies

2.3. Sub-populations

Figure 2.12: Histograms illustrating the effect of velocity errors on the detection of sub-populations with similar velocity dispersions. The total number of stars is 120 and the sub-populations contain equal numbers of stars. In each case, the double-populated sample with $\sigma_1 = 7 \text{ km s}^{-1}$ and $\sigma_2 = 4 \text{ km s}^{-1}$ is shown by the solid-line histogram and compared to a single population system with $\sigma = 7 \text{ km s}^{-1}$, shown as a dotted histogram. In the left panel, the velocity errors for both histograms are $dv = 2 \text{ km s}^{-1}$. As this error is relatively large compared to the difference in the dispersion, in the right panel we repeat the same experiment with $dv = 1 \text{ km s}^{-1}$.
Kinematic sub-populations in
dwarf spheroidal galaxies

2.3. Sub-populations

<table>
<thead>
<tr>
<th>$\sigma_{12}$</th>
<th>N</th>
<th>$N &gt; \sigma$</th>
<th>$N &gt; 3\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>120 stars</td>
<td>1000</td>
<td>1000</td>
</tr>
<tr>
<td></td>
<td>60 stars</td>
<td>1000</td>
<td>984</td>
</tr>
<tr>
<td>1.2</td>
<td>120 stars</td>
<td>1000</td>
<td>999</td>
</tr>
<tr>
<td></td>
<td>60 stars</td>
<td>999</td>
<td>895</td>
</tr>
<tr>
<td>1.3</td>
<td>120 stars</td>
<td>1000</td>
<td>963</td>
</tr>
<tr>
<td></td>
<td>60 stars</td>
<td>989</td>
<td>685</td>
</tr>
<tr>
<td>1.4</td>
<td>120 stars</td>
<td>997</td>
<td>782</td>
</tr>
<tr>
<td></td>
<td>60 stars</td>
<td>946</td>
<td>397</td>
</tr>
<tr>
<td>1.5</td>
<td>120 stars</td>
<td>966</td>
<td>480</td>
</tr>
<tr>
<td></td>
<td>60 stars</td>
<td>851</td>
<td>189</td>
</tr>
<tr>
<td>1.7</td>
<td>120 stars</td>
<td>874</td>
<td>203</td>
</tr>
<tr>
<td></td>
<td>60 stars</td>
<td>711</td>
<td>82</td>
</tr>
<tr>
<td>1.9</td>
<td>120 stars</td>
<td>738</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>60 stars</td>
<td>572</td>
<td>31</td>
</tr>
<tr>
<td>2.1</td>
<td>120 stars</td>
<td>562</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>60 stars</td>
<td>451</td>
<td>12</td>
</tr>
<tr>
<td>2.5</td>
<td>120 stars</td>
<td>424</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>60 stars</td>
<td>347</td>
<td>4</td>
</tr>
<tr>
<td>3.3</td>
<td>120 stars</td>
<td>313</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>60 stars</td>
<td>276</td>
<td>2</td>
</tr>
<tr>
<td>3.8</td>
<td>120 stars</td>
<td>227</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>60 stars</td>
<td>228</td>
<td>2</td>
</tr>
</tbody>
</table>

*Continued on the next page*
Kinematic sub-populations in dwarf spheroidal galaxies

2.3. Sub-populations

<table>
<thead>
<tr>
<th>Ratio</th>
<th>Total Number of Stars</th>
<th>Two-population Samples</th>
<th>Three-population Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>120 stars 193 0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>60 stars 195 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.5</td>
<td>120 stars 160 0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>60 stars 174 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>120 stars 145 0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>60 stars 154 1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.2: Confidence limits for the detection of sub-populations with different kinematics.

Columns are: (1) Ratio of main velocity dispersion $\sigma_1$ to the difference between the populations $\sigma_1 - \sigma_2$; (2) total number of stars in the data set; (3) Number of two-population samples for which $\Delta P$ is greater than the 1σ upper limit of $\Delta P$ obtained from single-population samples; (4) Number of two-population samples for which $\Delta P$ is greater than the 3σ upper limit of $\Delta P$ obtained from single-population samples. We compare populations of 60 and 120 stars containing two sub-populations. In each case $\sigma_1 = 15 \text{ km s}^{-1}$ while $\sigma_2$ lies in the range 1 km s$^{-1}$ to 14 km s$^{-1}$. Each dispersion ratio has been tested for 1000 data sets.

Finally, in order to test the effect of using the metallicity information of the data sets in our experiments we repeat some of the tests in two dimensions, as we did for CVn I (see Figure 2.8). This time each individual population is described by a combination of two Gaussians, one for metallicity and the other for velocity. As before, we generate data from either one or two populations, and then plot the distributions of $\Delta P$ obtained by fitting either one pair, or two pairs of Gaussians to the data. The results are shown in Table 2.3 and Table 2.4. In Table 2.3 the results of the one-dimensional velocity fits are different from Table 2.2. This arises solely due to the stochastic nature of the data sets and thus gives an indication of the level of noise in the experiments. We note, however, that our conclusions are not strongly sensitive to this noise.
Table 2.3 also shows that there is a qualitative difference between the metallicity and velocity data analysis, in that the metallicity data yield more significant results than the velocity data. This is due to differences in the properties of the data. In particular, the mean velocities of the input kinematic populations are similar to each other and the overlapping populations are distinguished differences in the velocity dispersion. On the other hand the input metallicity subpopulations have the same dispersions but differ in their mean values. As the separation in mean [Fe/H] is large compared to their dispersions in our test data, we are able to recover the two populations even with a small data set. We note that this table is not meant to be complete, but aims to give a subset of one-dimensional fits to be compared with the two-dimensional fits in Table 2.4.

As two-dimensional tests are much more computationally expensive, we consider only the case from Table 2.3 in which the metallicities of the populations are clearly separable. Table 2.4 and Figure 2.13 show that in this case, the inclusion of metallicity information in the analysis makes it possible to separate kinematic populations which were not separable using the kinematic data alone.

Finally, we note that, as in the one dimensional case, the results are dependent on the parameters of the underlying sub-populations. For example, in the case of our CVn I data, inclusion of metallicity information does not change the results obtained. If there are multiple kinematic populations in the CVn I data, their metallicities are likely quite similar; we find no evidence of multiple metallicity populations in either CVn I data set.
2.3. Sub-populations

<table>
<thead>
<tr>
<th>Populations</th>
<th>( N )</th>
<th>( N_{1D} &gt; \sigma )</th>
<th>( N_{1D} &gt; 3\sigma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_{12} = 5 )</td>
<td>120</td>
<td>228</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>198</td>
<td>1</td>
</tr>
<tr>
<td>( \sigma_{12} = 2.1 )</td>
<td>120</td>
<td>699</td>
<td>120</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>510</td>
<td>14</td>
</tr>
<tr>
<td>( \text{feh}_{12} = 0.5 )</td>
<td>120</td>
<td>985</td>
<td>349</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>887</td>
<td>38</td>
</tr>
</tbody>
</table>

Table 2.3: Confidence limits for detection of sub-populations as in Table 2.2. The table shows the results of one-dimensional fits to data sets using either velocities (top two populations) or metallicities (lower populations). The two velocity dispersion ratios that are used are \( \sigma_{12} = 2.1 \) and \( \sigma_{12} = 5 \). The mean metallicities are \([\text{Fe}/\text{H}]_1 = -2.0\) and \([\text{Fe}/\text{H}]_2 = -1.5\), yielding a value of \( \text{feh}_{12} = [\text{Fe}/\text{H}]_2 - [\text{Fe}/\text{H}]_1 \) of 0.5. The dispersion in the metallicities is the same for both populations and is equal to \( \sigma_{[\text{Fe}/\text{H}]} = 0.1 \). See text for discussion.

<table>
<thead>
<tr>
<th>([\text{Fe}/\text{H}]_{12} )</th>
<th>( \sigma_{12} )</th>
<th>( N )</th>
<th>( N_{2D} &gt; \sigma )</th>
<th>( N_{2D} &gt; 3\sigma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>5</td>
<td>120</td>
<td>1000</td>
<td>941</td>
</tr>
<tr>
<td>0.5</td>
<td>5</td>
<td>60</td>
<td>993</td>
<td>507</td>
</tr>
<tr>
<td>0.5</td>
<td>2.1</td>
<td>120</td>
<td>1000</td>
<td>967</td>
</tr>
<tr>
<td>0.5</td>
<td>2.1</td>
<td>60</td>
<td>997</td>
<td>617</td>
</tr>
</tbody>
</table>

Table 2.4: Confidence limits for two-dimensional detections. Comparison of the numbers with those in Table 2.3 show that in cases where the sub-populations differ both in kinematics and metallicities, the simultaneous use of both the parameters improves the results significantly (see also Figure 2.13). Note that this does not change the results for our CVn I data. See text for further discussion.
Figure 2.13: The figure shows the improvements of the fits for a two-dimensional fit (right), relative to separate velocity (middle) and metallicity (left) fits given in Table 2.3 and Table 2.4. The solid histograms show the results obtained for input data containing a single population, while the dashed histograms show results for input data containing two populations. The first and the third lines have histograms for data sets having 120 stars in total. The upper one is for $\sigma_{12} = 5$ and the lower one is for $\sigma_{12} = 2.1$. The second and fourth lines have again $\sigma_{12} = 5$ and $\sigma_{12} = 2.1$, respectively, for data sets consisting of 60 stars. [Note that in the [FeH] column both dashed histograms for N=120 (i.e. the leftmost panels in rows 1 and 3) are the same as the only difference between the two cases shown is in the velocity dispersions of the subpopulation which do not affect the results in the first column. The same is true for the N=60 data (rows 2 and 4).]
2.4 Conclusions

In this section we have presented a new data set of velocities and metallicities for the Canes Venatici I (CVn I) dSph, based on spectra taken with the GMOS-N spectrograph on the Gemini North telescope. A maximum likelihood fit to the velocity distribution yields a mean velocity of $v = 25.8 \pm 0.3 \text{ km s}^{-1}$ and a dispersion of $\sigma = 7.9^{+1.3}_{-1.1} \text{ km s}^{-1}$. Assuming a constant, isotropic velocity dispersion and a Plummer profile for the mass distribution, we find a mass of $4.4^{+1.6}_{-1.1} \times 10^7 M_\odot$ in the volume where our tracer stars are located. Although this value is larger than the value $2.7 \pm 0.4 \times 10^7 M_\odot$ calculated by Simon & Geha (2007), this is most likely due to the assumptions made for our models and the distribution of our particular subsets of stars.

One of the original aims of our study was to investigate the claimed multiple stellar populations in CVn I. As we discussed above, the two previous studies by Ibata et al. (2006) and Simon & Geha (2007) did not agree on the existence of a cold sub-population in CVn I. The two populations found in the former study were puzzling as they led to two different mass estimates. The authors suggested that this might indicate that the system had recently accreted a younger population and was not yet in equilibrium.

In this paper we looked for evidence of multiple populations in our data, under the assumption that each population was Gaussian. Based on this analysis, we concluded that there was no reason to suspect the presence of a second population in our data. We also applied our analysis to the Ibata et al. (2006) data, where we found evidence of a statistically significant sub-population with a dispersion of
\[ \sigma = 0.6 \text{ km s}^{-1} \] (compared to \( \sigma = 13.6 \text{ km s}^{-1} \) for the main population).

Although, statistically there is a possibility of having chosen two very different samples leading to different results because of the small sample sizes, our analysis suggests that there is a qualitative difference between our data and those of Ibata et al. (2006). Although further data would be necessary to resolve this issue, we note that the spatial distributions of these two data sets are different, which could potentially account for the differences in the detected populations. However, our central field is centred close to the blue/young star population which Martin et al. (2008a) find in their photometry from the Large Binocular Telescope, and which they identify with the cold population of Ibata et al. (2006). The exact fraction of stars in each population found by Martin et al. (2008a) is currently unclear, however, and so it is possible that we have not picked up any stars associated with the cold population.

We have also carried out a study of the detectability of sub-populations in small kinematic data sets. Under the assumption of Gaussian populations, we studied the effects of four parameters. We obtained confidence limits for the detection of sub-populations in samples with different numbers of stars, different population ratios and velocity dispersions. We found that for a case with distinct populations having similar velocity dispersions, reducing the observed velocity errors did not improve our ability to detect the sub-population. For a given sample size, our ability to detect two populations increased as the ratio of their dispersions, \( \sigma_1/\sigma_2 \), increased. However, even for large \( \sigma_1/\sigma_2 \) and equal population size, a sample of 30 stars yielded a 3\( \sigma \) detection in only about 35 per cent of cases. As expected, for larger sample sizes, this detection rate was significantly higher. We also showed that a cold population needs to constitute a larger fraction of the total sample than is required to
detect a hot sub-population. Finally, we showed that the simultaneous use of both kinematic and metallicity information can be a useful way to increase the significance of a detected sub-population in the case where there is a correlation between metallicity and kinematics and the populations are more distinct in metallicity than in kinematics.

Our results suggest that the robust detection of the sub-populations associated with any surviving sub-haloes within a dSph would require samples of many hundreds of stellar velocities. In this case, localised substructures could be detected by windowing the data, provided that a window whose spatial size coincided with plausible sub-halo scales would contain a sample of at least 100 stars. As such data sets are now becoming available for many of the larger dSphs, this test may soon be feasible. We note that the claim of multiple global populations in Sculptor (Tolstoy et al., 2004; Battaglia et al., 2008) was based on a large data set.

Finally, we note that all our significance tests were based on the assumption of Gaussian populations, which was the case for all our Monte Carlo samples. However, for real data, the true distributions will not be known, and are not necessarily well-approximated by Gaussians. It is therefore difficult in a real case to assign a robust statistical significance to a particular detection of a sub-population.

As we have shown, for small data sets, many Monte Carlo realisations do not yield significant detections of the sub-populations. In the absence of a robust estimate of the confidence level of a particular detection, or additional, independent evidence of the presence of multiple populations, we conclude that one should exercise great caution in decomposing data sets of fewer than 100 stars into multiple populations.
Acknowledgments

UU acknowledges funding from the European Commission under the Marie Curie Host Fellowship for Early Stage Research Training SPARTAN, Contract No MEST-CT-2004-007512, University of Leicester, UK. MIW acknowledges support from a Royal Society University Research Fellowship. AK acknowledges support by an STFC postdoctoral Fellowship. T.C.B. acknowledges partial funding of this work from grant AST 07-07776, as well as from PHY 02-16783 and PHY 08-22648: Physics Frontiers Center / Joint Institute for Nuclear Astrophysics (JINA), awarded by the U.S. National Science Foundation.

Based on observations obtained at the Gemini Observatory, which is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (United States), the Science and Technology Facilities Council (United Kingdom), the National Research Council (Canada), CONICYT (Chile), the Australian Research Council (Australia), CNPq (Brazil) and CONICET (Argentina). Program ID: GN-2007A-Q-66.
Carina is one of the most studied dSphs to date (e.g. Monelli, 2003; Koch et al., 2006; Muñoz et al., 2008). However, motivated by the complexity of its structure (Koch et al., 2006), the uncertainties in the observed orbital parameters (Piatek, 2003) and the possible degeneracies between the structural parameters, we choose to study it more extensively with a robust statistical analysis such as using Monte Carlo methods. In this chapter, we first introduce the concepts of N-body simulations and the Metropolis Markov Chain Monte Carlo (MCMC) method that was used to span the relevant parameter space of dark matter, orbital and stellar prop-
Sampling and Simulating 3.1. N-body simulations in astrophysics

properties for over 10000 simulations. A large suite of N-body tests was performed in order to establish the reliable limits for numerical and physical parameters. These will be presented in Chapter 4 together with the results of the science simulations and the analysis.

3.1 N-body simulations in astrophysics

Due to the impossibility of designing laboratory experiments that can duplicate astrophysical processes exactly, and often not even being able to observe a whole astrophysical process as it can take millions of years, computer simulations are our best experimental tools to study astrophysics. Simulations provide an environment in which we can apply the models built using our understanding of the laws of physics to simulate astrophysical processes. The comparison of the simulation results to the observational data is then the key to test our knowledge. The goal of this chapter is to understand the evolution of the Carina dSph satellite of the Milky Way. As in dSphs where the present-day amount of gas is very small, the long-term evolution of the system can be studied excluding the effect of the gas. In this chapter we focus on N-body simulations where the only physical interaction that is calculated is the gravitational interaction between stars and dark matter in a model galaxy represented by a large number of particles.

An N-body code calculates the Newtonian gravitational forces between particles $i$ and $j$ given in eq.3.1
3.1. N-body simulations in astrophysics

\[
F_j = \sum_{i=1}^{i=N} \frac{Gm_i m_j (r_i - r_j)}{|r_i - r_j|^3} - m_j \nabla \psi_{\text{ext}} = m_j \frac{\partial^2 r_j}{\partial t^2}
\]  (3.1)

where \( G \) is the gravitational constant, \( r \) is the position, \( m \) is the mass of each individual body and \( \psi \) is any external potential that might affect the system. Equation 3.1 is used to integrate the orbits of the individual particles as it is solved to calculate the acceleration \( \mathbf{a} \) of each particle.

Although ideally every star would be represented with a single point mass in an N-body simulation, the number of these interactions in a system that has \( N \) stars would be \( N \times (N - 1) \). When the number of particles \( N \) is large (as is the number of stars in a galaxy), solving the differential equation 3.1 is very computationally expensive at each time step. In addition, considering that the calculations need to be done for many time steps in order to follow the time evolution of the system, the total number of required calculations is very large for simulating a galaxy, since even the smallest galaxies have millions of stars. In addition, in this thesis a high number of simulations need to be done for a comprehensive search in the parameter space. In order to decrease the computational cost, often tens to thousands of stars will be represented by a single particle. Moreover, the N-body code that will be used for the simulation can be optimised, by using efficient distribution of the data to the computing nodes. In this chapter we explain one of these optimisation methods: the hierarchical tree structure.
3.1.1 Tree codes

A tree algorithm as described in *Barnes & Hut* (1986) can significantly reduce the number of calculations by tessellating the root node into daughter nodes, ie. the whole simulation into sub-cells containing only a small fraction of the total number of particles. The common octree structure consists of eight daughter nodes which themselves will be tessellated into more nodes and finally to “leaves”, forming a hierarchical tree structure. The denser areas are divided more than the less dense areas in order not to have too few or too many particles in the cells assuring that the load of computations is at similar levels everywhere in the simulation. As the gravitational attraction between particles decreases with the square of the distance between particles (Eq. 3.1), the interactions between distant areas of the simulation can be approximated by calculating the forces between close groups of particles.

In this case, a leaf that is far away is treated as a single particle which has a mass equal to the sum of all the particles’ masses inside that leaf and located at its centre of mass. The numerical parameter used to determine the minimum distance between leaves before they can be treated as distant point masses is the opening angle, \( \theta \). Figure 3.1 illustrates the effect of changing the value of \( \theta \). A larger opening angle means that the approximation will be applied to the leaves at smaller distances. It is important to use an opening angle which keeps the accuracy of the force at an acceptable level while diminishing the number of forces calculated to save computational time.

Another issue with the computational affordability comes from Eq. 3.1. In reality, galaxies should be treated as collisionless sytems. Because of their low stellar
Sampling and Simulating

3.1. N-body simulations in astrophysics

Densities, stellar close encounters are rare and are of little importance for the general dynamics of the system, in contrast to star clusters whose evolution can be dominated by these interactions. In a galaxy, each star will experience a smooth acceleration under the gravitational forces acting on it by the large scale structure of the galaxy, i.e., the combined forces of all the individual stars and dark matter particles. However, in galaxy simulations where stars are represented point masses, Eq. [3.1] gives an infinitely large force when the inter-particle distance approaches to zero. As the acceleration will be very large for those very small inter-particle distances, in order to calculate the accelerations caused by close encounters accurately, the calculations would have to be made with very small time steps, increasing the simulation time. Also, these interactions are unphysical since particles are more massive than stars.
The solution is to enforce the quasi-collisionlessness of the system by modifying the point mass assumption slightly. The point masses are converted to smoothed particles with radius $\epsilon$, the softening parameter. Although the masses are still presented by near point masses, the softening parameter avoids having infinitely large accelerations calculated in infinitesimal time steps. In addition, by limiting the minimum distance between particle pairs, the softening parameter prevents stars from being captured by each other’s gravitational potentials and thus solves another numerical problem that can arise by the formation of binary stellar systems from the point masses in the simulations. Otherwise, these could lead to a high number of unphysical binary systems being formed because of the point mass approximation.

In the presence of softening, the Newtonian gravitational potential is modified for particles closer than $2h$ where $h$ is the softening length. The spline smoothed gravitational kernel used by PKDGRAV is given in Equations 3.2 to 3.5 (see Stadel, 2001) where $\rho$, $\phi$ and $a$, are density, potential and acceleration respectively.

\[
\rho = \frac{m}{\pi h^3} \begin{cases} 
1 - \frac{3}{2}(2-u)u^2 & \text{if } u < 1 \\
\frac{1}{4}(2-u)^3 & \text{if } 1 \leq u < 2
\end{cases}
\]  

\[
\phi = \frac{m}{h} \begin{cases} 
\frac{7}{5} - u^2 \left(\frac{2}{3} - \frac{1}{10}u^2(3-u)\right) & \text{if } u < 1 \\
-\frac{1}{15}u^{-1} + \frac{8}{5} - u^2 \left(\frac{4}{3} - u \left(1 - \frac{1}{10}u(3 - \frac{1}{3}u)\right)\right) & \text{if } 1 \leq u < 2
\end{cases}
\]  

\[
u = \frac{|r|}{h}
\]
Sampling and Simulating

3.1. N-body simulations in astrophysics

\[ a = \frac{r m}{h^3} \begin{cases} \frac{4}{3} - \frac{1}{2} \left( \frac{12}{5} - u \right) & \text{if } u < 1 \\ -\frac{1}{15} u^3 + \frac{8}{3} - u \left( 3 - u \left( \frac{6}{5} - \frac{1}{6} u \right) \right) & \text{if } 1 \leq u < 2 \end{cases} \] (3.5)

It is important to remember that as in the case of the opening angle \( \theta \), \( \epsilon \) should also be chosen with care in order to ensure the reliability of the force calculation while shortening the simulation time. Later in this chapter the time tests take into account these trade offs, and determine optimum values of these parameters to be used in our main Carina simulations.

3.1.2 PKDGRAV

**Definition: K-d tree**: The N-body code PKDGRAV (Stadel, 2001) used for the simulations in this thesis has a different tree structure than the octree described above. The k-d tree (k-dimensional tree) (Bentley, 1975) used in PKDGRAV is a binary tree where each node is divided along its long axis which separates equally loaded daughter sub-trees (Stadel, 2001). Eventually the tree is composed of local trees and leaf cells called buckets which have a maximum of eight particles.

Each bucket has an opening radius about its centre of mass which is related to the user specified opening angle \( \theta \) by

\[ r_o = \frac{2B_{\text{max}}}{\sqrt{3} \theta} \] (3.6)

\( B_{\text{max}} \) is the maximum distance of any particle in the bucket to the centre of mass of
the bucket. If the opening radius is smaller than the distance between two buckets, these are far enough to be treated as single particles at their centre of mass. On the other hand if the opening radius is large enough to reach the bucket under consideration, this needs to be “opened”, ie. the interactions will be calculated between individual particles.

PKDGRAV uses three parameters to set the time scales for the N-body integrations. The first one is $\delta$ which sets the maximum time step allowed for the particles on the maximum rung (the main node of the simulation), and the time scale of each consecutive rung (daughter node) is half of the previous one. For $\delta = 10^8$ years, the tree is constructed with the root node set to particles with time steps of $10^8$ years, the first set of daughter nodes allow $5 \times 10^7$ years, the next set are $2.5 \times 10^7$ and so on.

The second time parameter is $d\eta$ which sets the smallest time steps of the simulation. It is specified by the user as a fraction of local dynamical time, which is set by $\delta$ for each rung. The value of $d\eta$ should be chosen taking into account the trade off between the simulation time and accuracy in the force calculations. Although $\delta$ has a smaller effect on the simulation time it needs to be observed as well to make sure that the particles are distributed into different rungs in a sensible way, ie. not spanning too large or too small range of time steps in the simulation.

The final time parameter is the length of the simulation. Although, Carina is an old system which has stars of the order of 10 Gyrs of age, we choose to do our simulations for 5 and 6 Gyrs, as the Milky Way potential might have been considerably different further back in time and we do not have a prescription for tracing back orbits reliably to earlier times (Lux, 2010).
As Carina is not an isolated galaxy but a satellite which has been and still is interacting with the Milky Way, a fixed external potential was implemented into PKDGRAV by adding another code: GalPot (Dehnen & Binney, 1998). The chosen model for the Milky Way is Model 4a from Dehnen & Binney (1998) where the Sun’s distance to the Galactic Centre is given as $R_0 = 7.5\text{kpc}$. In the simulations we hold all the parameters of the Milky Way model fixed. The Milky Way model consists of a three component flattened disk, each of the form given in (Eq. 3.7) composed of interstellar medium, a thin and a thick stellar disk, and two spheroids of the form given (Eq.3.8), one for the dark matter halo and one for the bulge:

$$\rho_d = \frac{\Sigma_d}{2z_d} \exp\left(-\frac{R_m}{R_d} - \frac{R}{R_d} - \frac{|z|}{z_d}\right)$$  \hspace{1cm} (3.7)

<table>
<thead>
<tr>
<th>Component</th>
<th>Contribution to $\Sigma(R_0)$</th>
<th>$R_d/R_d, \ast$</th>
<th>$R_M[kpc]$</th>
<th>$z_d[kpc]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISM</td>
<td>0.25</td>
<td>2</td>
<td>4</td>
<td>0.04</td>
</tr>
<tr>
<td>Thin disk</td>
<td>0.7</td>
<td>1</td>
<td>0</td>
<td>0.18</td>
</tr>
<tr>
<td>Thick disk</td>
<td>0.05</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3.1: Milky Way disk properties as given in Dehnen & Binney (1998). $R - d$ is the scalelength of the particular disk that is being taken into account (namely, the scalelength of the ISM, thin and thick stellar disks), while $R_d, \ast$ is the scalelength of the stellar disk.

$$\rho_s = \rho_0 \left(\frac{m}{r_0}\right)^{-\gamma} \left(1 + \frac{m}{r_0}\right)^{\gamma-\beta} e^{-m^2/r_i^2}$$  \hspace{1cm} (3.8)

where $m = (R^2 + q^2z^2)^{1/2}$. The bulge parameters apart from the central density $\rho_{0,b}$ are constrained by observational data directly: $\beta_b = \gamma_b = 1.8, q_b = 0.6, r_{0,b} = 1\text{kpc}$ and $r_{t,b} = 1.9\text{kpc}$. The central surface density is found by a least squared fit depending on the model and is equal to $49.6\text{M}_\odot\text{pc}^{-2}$ in the model that we use. The halo
flattening $q_h = 0.8$ is chosen arbitrarily, while the rest of the halo parameters are calculated using least squared fits to the velocities in the inner parts and the rotation curves in the outer parts. These are $\gamma_h = 1.76$, $\beta_h = 2,108$, $r_{0,h} = 1\text{kpc}$, $r_{r,b} = \infty$ and $\rho_h = 0.76\text{M}_\odot\text{pc}^{-3}$.

Before going into the details of the tests done with PKDGRAV, we will have a look at the description of the statistical methods used together with the N-body simulations in order to search the space of Carina’s parameters efficiently and obtain the distribution of parameters.

### 3.2 Markov Chain Monte Carlo method: The Metropolis-Hastings algorithm

Monte Carlo methods are computational algorithms based on repeating experiments a large number of times using randomly chosen parameters. In other words, we use computers to make thousands of experiments on an object we want to understand, instead of doing a single experiment where we try to make the right “guess” for the initial conditions of the system. In our case the object is Carina and the experiment will be N-body simulations made for a number of progenitors that might have existed $5\text{Gyrs}$ ago. The randomness of the input and high number of simulations are crucial to make certain that the sampled models are representative of all different states of the physical system, and to ensure that in the end we can find a distribution of the probabilities for all the likely models for Carina as well as the most likely progenitor models.
3.2. Markov Chain Monte Carlo Method: Metropolis-Hastings algorithm

Being completely deterministic, computers can only generate pseudo-random numbers. These will be generated by a function that uses an initial seed as the starting point, therefore in reality the numbers are completely predictable in advance. However, the predictability is not a problem as long as the sample is 'random enough' for the current problem which is ensured by the use of sophisticated algorithms that combine several functions together. Python uses the Mersenne Twister (Matsumoto, 1998) Pseudo-Random Number Generator (PRNG), which has a repetition period of $2^{19937} - 1$ and performs well in many randomness tests. The effect of any non-randomness is then neglected as it is not correlated to the physical parameters. The biases that might be introduced will cancel out in general and will not bias any physical parameter in a given direction.

**Definition: Markov Chain**: A Markov chain is a discrete random process with the Markov Property, i.e. the next state of the system only depends on the current state, and the past states are forgotten.

The Markov Chain Monte Carlo (MCMC) method uses a random walk algorithm, i.e., it takes randomly sized and directed steps to choose samples in the parameter space. At each step, the probability of the current physical state in the simulation is compared to the probability of the last state. The more “favorable” point which gives a “closer” result to the real data is then chosen as the starting point for the next step of the random walk. The walk eventually converges towards a stationary state where the best point of the parameter space is found and the experiment is not trapped in a local minimum.
The Metropolis-Hastings algorithm (e.g. Metropolis, 1953, 1987) is an MCMC algorithm that samples the parameter space efficiently using a proposal distribution, in our case a normal distribution. It starts by selecting a "reasonable" starting value for the parameter to be explored. The random steps are then picked from a Gaussian distribution generated by the PRNG. In the statistical system where random steps are taken the algorithm looks for a solution with smaller $\chi^2$. This is done by calculating the likelihoods of the models $L_i = (e^{-\chi^2_i/2})/ \sqrt{2\pi}$ at each step and comparing the ratio of the likelihoods of the consecutive steps given in eq. \[3.9\]

$$\frac{L_2}{L_1} = e^{-\frac{\chi^2_1 - \chi^2_2}{2}}$$

where

$$\chi^2_i = \frac{1}{n} \sum \frac{(p_{model} - p_{observed})^2}{dp_{model}^2 + dp_{observed}^2}.$$  \[3.10\]

In eq. \[3.10\] n is the number of data points used in the calculations, $p$ is the value of the parameter used and $dp$ is the error in the values. The ratio of the likelihoods is used to decide if a model is accepted and therefore becomes the current state of the chain, before choosing the next random direction in the parameter space. If the most recent model is a better fit to the data ie. $(\chi^2_i < \chi^2_{i-1})$, the ratio $L_i/L_{i-1}$ in eq. \[3.9\] is larger than one and this model passes the test likelihood. If the $\chi^2$ of the most recent model is larger than the previous one, $L_i/L_{i-1}$ is instead compared to a random number $0 < x < 1$. This way a model that has larger difference from the observations (larger $\chi^2$) still has the chance of being accepted if $L_i/L_{i-1} > x$. This is the strength of the algorithm as it prevents the search from being trapped around a local minimum for in $\chi^2$ states. It instead keeps looking for a global minimumum where the parameters are closest to their real values. However, it still weighs the better likelihoods more, leading to an overall convergence. If the most recent model
is accepted, it becomes the new starting point for the next random step. If it fails the likelihood test, the previous model is re-accepted, and the next step is taken from there in a different random direction. The chain obeys the Markov Property as the only likelihoods of two consecutive states are compared.

The chain needs to take enough steps in order to reach the stationary state. The exact number depends on the dimensions of the parameter space, the range of values allowed for each parameter, the maximum size of the steps that is allowed and the burn-in period. The burn-in period is the number of simulations that the chain needs to complete before converging. These first steps remember the starting point of the chain, i.e., they are biased towards the initial values of the parameters and therefore need to be excluded from the final analysis. At the point where the likelihood of consecutive simulations do not have any systematic changes, hence it is only the noise in the likelihoods that can be seen, the chain has converged. From this point all the steps can be used in the statistical analysis. If multiple chains are being run at the same time, the chains are converged and can be added together once the variance between the means of the chains is smaller than the variances in the parameters within individual chains.

### 3.3 Implementing MCMC for Carina

So far in this chapter, the general methods that are used in the Carina project have been explained. Now that both the simulating and the sampling methods have been introduced, we look at the implementation of the MCMC for Carina, and we then give the details of the tests that are made in order to establish the limitations of our
3.3. Implementing MCMC for Carina

3.3.1 Defining the parameter space

We chose four structural and two orbital parameters of Carina as the free parameters we want to study: Stellar mass $M_*$, dark matter mass $M_{DM}$, scale length of the dark matter halo $r_{DM}$, scalelength of the stellar component $r_s$, and the two components of the proper motion in the direction of right ascension and the declination, $\mu_\alpha$ and $\mu_\delta$ respectively. Although the final state of the MCMC does not depend on the starting point in the parameter space, initially we choose to use good fits to the data as this can help to minimize the burn in period.

The dark matter halo of Carina is modelled using a Hernquist density profile where $\alpha = 1$, $\gamma = 1$ and $\beta = 4$ in eq (3.11). The structural parameters for the stellar component ($\alpha$, $\beta$, $\gamma$, $r_s$) are fixed as well and are given in Table 3.2 together with the observed values of the free MCMC parameters ($M_*$, $M_{DM}$, $r_{DM}$, $\mu_\alpha$, $\mu_\delta$).

$$\rho(r) = \frac{\rho_0}{(r/a)^\gamma \left(1 + (r/a)^{\beta-\gamma}\right)}$$  \hspace{1cm} \text{(3.11)}

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\gamma$</th>
<th>$r_s$</th>
<th>$M_*$</th>
<th>$M_{DM}$</th>
<th>$r_{DM}$</th>
<th>$\mu_\alpha$</th>
<th>$\mu_\delta$</th>
<th>$V_{Helio}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>7</td>
<td>0.37</td>
<td>0.37</td>
<td>$4.3 \times 10^3$</td>
<td>$10^8$</td>
<td>1</td>
<td>15</td>
<td>22</td>
<td>223</td>
</tr>
</tbody>
</table>

Table 3.2: Parameter values for the stellar component of the models. The units are $[M_\odot]$ for masses, [kpc] for distances and [mas/century] for the proper motions and [km/s] for the radial velocity.

The range of values allowed for each parameter is given in Table 3.3. The choice of these ranges will be explained in more detail in the next chapter.
3.3. Implementing MCMC for Carina

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MIN</td>
<td>4.3×10^5</td>
<td>10^7</td>
<td>0.5</td>
<td>-12</td>
<td>-5</td>
</tr>
<tr>
<td>MAX</td>
<td>2.15×10^6</td>
<td>5×10^8</td>
<td>1</td>
<td>42</td>
<td>49</td>
</tr>
</tbody>
</table>

Table 3.3: The ranges of the MCMC parameter values for Carina. The units are as given in Table 3.2

3.3.2 Algorithm

The galaxies in the chain are generated using falcON (Dehnen, 2000, 2002) which solves the Poisson equation \((\nabla^2 \phi = 4\pi G \rho)\) for self-consistent multiple component potentials corresponding to the density of the form given in Eq. 3.11. It then generates each component with the desired number of particles according to the distribution functions calculated. Although there is the possibility of adding a disk in the models, we only generate two-component dSph models consisting of a spherical dark matter halo and stellar component. In order to test the stability of our models and take into account the numerical noise, each dwarf galaxy is given enough time to relax initially. This is done by putting the dwarf on a circular orbit at 300kpc where the tidal force exerted on it by the Miky Way is negligible and running a short N-body simulation of 50Myrs. The resulting galaxy is stable in isolation as will be shown at the end of this chapter. This stable galaxy is then put on Carina’s orbit and a longer N-body simulation is made to study the interaction of the satellite with Milky Way over the past five or six billion years.

To determine the initial position and velocity of Carina for our simulation, we integrate a point mass backwards in time from the present day. The orbit of the satellite around the Milky Way at the present is initially determined by the observed values: \(\alpha = 06^h 41^m 36.7^s, \delta = 50^\circ 57' 58''\) (J2000), \(\mu_\alpha = 22 \pm 9\) mas/century, \(\mu_\delta = 15\pm9\) mas/century at a distance of 102 kpc (Piatek, 2003) and \(v_{\text{helio}} = 223\) km/s as calculated by Mateo (1993). We calculate the radial velocity with respect to the
galactic centre by calculating the galactic coordinates and velocities as given in Table 3.4 and find $V_R = 3\text{kms}$. 

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>$V_X$</th>
<th>$V_Y$</th>
<th>$V_Z$</th>
</tr>
</thead>
<tbody>
<tr>
<td>24.54</td>
<td>-92.01</td>
<td>-38.156</td>
<td>73.89</td>
<td>10.17</td>
<td>35.89</td>
</tr>
</tbody>
</table>

Table 3.4: The galactic coordinates of the centre of Carina calculated using the right ascension, declination, the heliocentric radial velocity and proper motion. The distances are given in [kpc] while the velocities are in units of [km/s].

The position and velocity coordinates that the satellite will have at the beginning of the N-body simulation are then found by integrating the current orbit back for 5 or 6 Gyrs using a Leapfrog time integrator with adaptive timesteps written in C++ by Hanni Lux (see Lux, 2010). As in PKDGRAV the orbit integration is made in the presence of the Dehnen potential 4a (Dehnen & Binney, 1998). The orbit integrator provides the position and velocity coordinates of the centre of mass 5 or 6 Gyrs ago and every particle in the relaxed satellite is then shifted together, thus shifting the centre of mass to these new coordinates. The sign of the bulk velocity is then reversed and the orbit is integrated forwards again as a full N-body simulation of 6 Gyrs. There are two reasons for choosing the integration time to be 5 or 6 Gyrs. The first one, that is mentioned earlier in this chapter, is the probable change in Milky Way’s own structure in the earlier periods of its lifetime, that would need to be simulated with an evolving external potential. The second reason is the chaotic nature of the Milky Way’s satellites’ orbits at high redshifts where mergers and inter-galactic accretion are dominant making the predictions of the orbits unreliable (see Lux, 2010). Therefore we choose the integration to be 6 Gyrs as it is a safe choice for modelling the Milky Way’s potential (Penarrubia et al., 2006), and it is a long enough time to see the effects of tides on Carina.
As soon as the first simulation is finished, the result for this first model is compared to the observational data. The simulation data are binned the same way as the observational data (see Table 4.1 and Table 4.2 in Chapter 4) and the reduced $\chi^2$ taking into account both the surface brightness $\Sigma(R)$ and the projected velocity dispersion $\sigma(R)$ profiles is calculated using

$$
\chi^2 = \frac{1}{n_{\Sigma,\text{bin}}} \sum \frac{(\Sigma_{\text{sim}} - \Sigma_{\text{obs}})^2}{\sqrt{d\Sigma_{\text{sim}}^2 + d\Sigma_{\text{obs}}^2}} + \frac{1}{n_{\sigma,\text{bin}}} \sum \frac{(\sigma_{\text{sim}} - \sigma_{\text{obs}})^2}{\sqrt{d\sigma_{\text{sim}}^2 + d\sigma_{\text{obs}}^2}}
$$

(3.12)

where $n_{\Sigma,\text{bin}}$ and $n_{\sigma,\text{bin}}$ are the number of bins in the surface brightness and the velocity dispersion, respectively. As this is the first model in the chain it is accepted regardless of the value of $\chi^2$. However, this value is stored for comparison with the next model.

The exploration of degeneracies between the orbital and structural parameters with MCMC starts at this point where the algorithm takes the first random step to choose the next values of the five parameters given in Table 3.2. The basic step size for each parameter is fixed to be $s_p = (P_{\text{max}} - P_{\text{min}})/10$, where $P_{\text{max}}$ and $P_{\text{min}}$ are the maximum and minimum values for each parameter given in Table 3.3. The new value $P_{\text{new}}$ is then computed for each parameter as in Eq. 3.13 for which the random numbers $r_N$ are generated from Gaussian distributions with mean at zero and $\sigma = 1$.

$$
P_{\text{new}} = P_{\text{old}} + r_N \times s_p
$$

(3.13)
The choice of the base step is made to ensure that the chain takes steps large enough for each parameter to span the space completely and quickly while making sure that it is small enough to distinguish between closer models. This is enforced by monitoring the acceptance rate every 50 steps in the chain and the step size is changed accordingly. If the acceptance rate is too high, the algorithm is adjusted to take larger steps \( (s_{\text{new}}) \) to reach different areas of the parameter space by:

\[
s_{\text{new}} = e^{\ln(s_{p})/\ln(10)+d\,s}/2\times\ln(10)
\]

where \( d\,s = s_{p} \). If the acceptance rate is too low and smaller steps are necessary then \( d\,s = -s_{p} \).

Once the new parameter values are determined, the above steps in the MCMC are repeated by generating a new galaxy that is allowed to relax and put on a new orbit. The new N-body simulation is performed and finally the new \( \chi^2_{\text{new}} \) is calculated between this new model and Carina’s data. As explained in Section 3.2, if the new model gives a better fit to the observational data therefore \( \chi^2_{\text{new}} < \chi^2_{\text{old}} \), it is accepted and the next random step will be taken from this models’ position. On the other hand, if \( \chi^2_{\text{new}} > \chi^2_{\text{old}} \), the choice between the old and new models is made by comparing the ratio of the likelihoods (Eq. 3.9) to a random number between zero and one. If the new model is rejected the old one is counted as being accepted once more and it will be from its position that a random step will be taken again. Ideally, the condition of the likelihoods comparison allows about 40% of the new models to be accepted even if some of them have larger \( \chi^2 \) than the models that they are compared to. This is the property of the MCMC that reduces the risk of the search.
being caught at a local minimum while still providing more weight to better models by accepting models with better likelihoods more often.

At the end of the MCMC, the distribution of likelihoods in the parameter space is obtained and the distribution of each parameter is plotted in a 2D plot showing the degeneracies between each pair of parameters. High resolution simulations are made for the few models that are very good fits to the data in order to check that our conclusions are qualitatively similar to lower resolution models. These will be presented in Chapter 5.

### 3.4 Tests

In this section, the results of the numerical tests done before running the MCMC for a real dSph will be explained. These were done for three reasons: 1) To establish the sensitivity of the results to the choice of numerical parameters; 2) To understand the constraints that were imposed on the numerical parameters by the computational resources and time; 3) To show that the MCMC can work for this project. The last one of these points was investigated by applying our algorithm to match the output of a simulated galaxy. This showed that our analysis methods were able to distinguish between different models and reveal degeneracies between the parameters. The results of these will be presented in the next chapter, while here we will focus on the first two that are related to the numerical parameters.

The trade off between the efficient use of resources and the accuracy that can be acquired is established in this section by investigating the effects of the four nu-
merical parameters: the softening length $\epsilon$, the minimum time step $d\eta$ for the force calculation, the opening angle $\theta$ and the number of particles used in the simulations $N$. In all cases the particles were shared equally between the halo and the stellar distribution. This leads to a higher resolution in the less massive and less extended stellar distribution which was preferred as it offers the only means of direct comparison to the observational data. On the other hand $N$ also needed to be large enough for the halo profile to be traced accurately.

### 3.4.1 High resolution tests

Initially two simulations with a total of $2 \times 10^6$ particles, $N_{\text{tot}}$, are carried out using the physical parameters given in Table 3.2 and the numerical parameters given in Table 3.5. The simulations follow the evolution of the model dSph for 5 Gyrs on its orbit around the Milky Way. The parameter $\delta$ which gives the maximum time step of the highest rung in PKDGRAV is only monitored at the beginning of these simulations and will not be changed again. We find that $\delta = 10^8$ years gives the following rung distribution for the particles:

$$(0, 0, 0, 1, 62, 3678, 420149, 1576090, 20, 0, 0, 0),$$

using which time scale is found by calculating $t_{\text{rung}} = \delta/n_{\text{rung}}$.

Starting from left the first four rungs corresponding to $10^8$, $5 \times 10^7$, $2.5 \times 10^7$, $1.25 \times 10^7$, years have no particles and the range of time steps covering between a maximum of $4 \times 10^5$ to $6 \times 10^6$ years, gives a balanced distribution of particles in the rungs. The other time related parameter $d\eta$ is chosen so that $d\eta = 0.1$ and therefore each particle is integrated using approximately different points on a single orbit.
We choose a test value of $\epsilon = 30\text{pc}$ for the softening parameter. The acceleration gained by a particle $i$ through its interaction with all the particles in the field (left hand side of the Equation 3.15) needs to be larger than the acceleration it can gain through any two body interactions (right hand side of the Equation 3.15).

\[
\frac{Gm_i m_j N_{r<30}}{30^2} \gg \frac{Gm_i m_j}{\epsilon^2}
\]  \hspace{1cm} (3.15)

giving

\[
\epsilon \gg r/ \sqrt{N_{r<30}}
\]  \hspace{1cm} (3.16)

Equation 3.16 puts a lower limit on $\epsilon$. We test this value for the innermost 30pc of the galaxy. Out of the $2 \times 10^6$ particles, 1003 in the halo, and 2117 in the bulge were in the innermost 30pc providing the conditions $\epsilon \gg 1.06\text{pc}$ and $\epsilon \gg 0.65\text{pc}$ respectively, both consistent with the $\epsilon = 30\text{pc}$ choice.

The range of values suggested for the opening angle to use in PKDGRAV is $0.5 < \theta < 0.7$ for cosmological simulations. However, as this is an arbitrary value for the simulations we plan to do, in these two high resolution runs we investigate whether a larger value of $\theta$ could safely be used or not. Table 3.5 and Figure 3.2 show that both runs provide good fits to the observational data with no significant change depending on the opening angle.

As seen in the table, these runs which used 32 processors for several days are both
very slow and computationally expensive. They will be used from now on as the reference points of the accuracy tests that will be done for a number of numerical parameters with the more feasible number of $2 \times 10^5$ particles.

<table>
<thead>
<tr>
<th>Name</th>
<th>$N_s$</th>
<th>$N_{DM}$</th>
<th>$N_p$</th>
<th>$\epsilon$</th>
<th>$d\eta$</th>
<th>$t_{step}$ (Myrs)</th>
<th>$d\theta$</th>
<th>PKDGRAV</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>$10^6$</td>
<td>$10^6$</td>
<td>32</td>
<td>30</td>
<td>0.1</td>
<td>100</td>
<td>1.0</td>
<td>53h45m</td>
<td>1.01</td>
</tr>
<tr>
<td>B7</td>
<td>$10^6$</td>
<td>$10^6$</td>
<td>32</td>
<td>30</td>
<td>0.1</td>
<td>100</td>
<td>0.7</td>
<td>91hrs</td>
<td>1.01</td>
</tr>
</tbody>
</table>

Table 3.5: Two high-resolution simulations with different opening angles $\theta$. Although the result of the simulation is not affected by the choice of $\theta$, the simulation time decreases by 40% when $\theta = 1$. The $\chi^2$ is calculated separately between each simulation and the observational data.

### 3.4.2 Lower resolution tests

In this section, we present the lower resolution tests that have been done in order to find the optimal values for the N-body simulations of our science MCMC runs. The physical model of the dSph galaxy is the same as in the previous tests. Table 3.6 shows all the tests that were made with $2 \times 10^5$ particles.
Figure 3.2: The high resolution test simulation results for B1 (upper panel) and B7 (lower panel). Although the velocity dispersions show that this will not be the best fit to the data, the surface brightnesses are well reproduced. The dotted red line on the panels on the right is our resolution limit set to $\epsilon = 30\text{pc}$.
### 3.4. Tests

<table>
<thead>
<tr>
<th>Name</th>
<th>$N_*$</th>
<th>$N_{DM}$</th>
<th>$N_p$</th>
<th>$\epsilon$ [pc]</th>
<th>$d\eta$</th>
<th>$t_{\text{step}}$ (Myrs)</th>
<th>$d\theta$</th>
<th>PKDGRAV (min)</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>8</td>
<td>30</td>
<td>0.1</td>
<td>100</td>
<td>0.7</td>
<td>17h49m*</td>
<td>1.68</td>
</tr>
<tr>
<td>M2</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>8</td>
<td>30</td>
<td>0.1</td>
<td>100</td>
<td>1.0</td>
<td>9h</td>
<td>2.10</td>
</tr>
<tr>
<td>M3</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>8</td>
<td>30</td>
<td>0.2</td>
<td>100</td>
<td>0.7</td>
<td>9h*</td>
<td>1.84</td>
</tr>
<tr>
<td>M4</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>8</td>
<td>30</td>
<td>0.2</td>
<td>100</td>
<td>1.0</td>
<td>4h15m</td>
<td>2.12</td>
</tr>
<tr>
<td>M5</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>8</td>
<td>30</td>
<td>0.3</td>
<td>100</td>
<td>0.7</td>
<td>5h</td>
<td>2.13</td>
</tr>
<tr>
<td>M6</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>8</td>
<td>30</td>
<td>0.3</td>
<td>100</td>
<td>1.0</td>
<td>3h39m</td>
<td>2.42</td>
</tr>
<tr>
<td>M7</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>8</td>
<td>30</td>
<td>0.4</td>
<td>100</td>
<td>0.7</td>
<td>3h20m</td>
<td>3.27</td>
</tr>
<tr>
<td>M8</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>8</td>
<td>30</td>
<td>0.4</td>
<td>100</td>
<td>1.0</td>
<td>2h50m</td>
<td>3.68</td>
</tr>
<tr>
<td>M9</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>8</td>
<td>30</td>
<td>0.25</td>
<td>100</td>
<td>0.7</td>
<td>5h49m</td>
<td>2.50</td>
</tr>
<tr>
<td>M10</td>
<td>$10^5$</td>
<td>$10^6$</td>
<td>8</td>
<td>30</td>
<td>0.25</td>
<td>100</td>
<td>0.8</td>
<td>5h11m</td>
<td>2.52</td>
</tr>
<tr>
<td>M11</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>8</td>
<td>30</td>
<td>0.25</td>
<td>100</td>
<td>0.9</td>
<td>4h25m</td>
<td>2.57</td>
</tr>
<tr>
<td>M12</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>8</td>
<td>30</td>
<td>0.25</td>
<td>100</td>
<td>1.0</td>
<td>5h10m</td>
<td>2.57</td>
</tr>
<tr>
<td>M13</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>8</td>
<td>100</td>
<td>0.25</td>
<td>100</td>
<td>0.7</td>
<td>3h12m</td>
<td>3.83</td>
</tr>
<tr>
<td>M14</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>8</td>
<td>100</td>
<td>0.25</td>
<td>100</td>
<td>1.0</td>
<td>2h13m</td>
<td>3.53</td>
</tr>
<tr>
<td>M15</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>8</td>
<td>50</td>
<td>0.25</td>
<td>100</td>
<td>0.7</td>
<td>6h45m</td>
<td>2.24</td>
</tr>
<tr>
<td>M16</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>8</td>
<td>50</td>
<td>0.25</td>
<td>100</td>
<td>1.0</td>
<td>3h43m</td>
<td>2.00</td>
</tr>
<tr>
<td>M17</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>16</td>
<td>30</td>
<td>0.25</td>
<td>100</td>
<td>0.9</td>
<td>3h30m*</td>
<td>2.26</td>
</tr>
<tr>
<td>M18</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>16</td>
<td>30</td>
<td>0.25</td>
<td>100</td>
<td>1.0</td>
<td>3h30m</td>
<td>2.31</td>
</tr>
<tr>
<td>M19</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>4</td>
<td>30</td>
<td>0.25</td>
<td>100</td>
<td>0.9</td>
<td>7h30m*</td>
<td>2.30</td>
</tr>
<tr>
<td>M20</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>4</td>
<td>30</td>
<td>0.25</td>
<td>100</td>
<td>1.0</td>
<td>6h28m*</td>
<td>2.63</td>
</tr>
<tr>
<td>M21</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>16</td>
<td>50</td>
<td>0.25</td>
<td>100</td>
<td>1.0</td>
<td>2h8m</td>
<td>1.71</td>
</tr>
<tr>
<td>M22</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>16</td>
<td>50</td>
<td>0.4</td>
<td>100</td>
<td>1.0</td>
<td>1h25m</td>
<td>4.38</td>
</tr>
<tr>
<td>M23</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>16</td>
<td>30</td>
<td>0.3</td>
<td>100</td>
<td>0.9</td>
<td>3h13m</td>
<td>2.09</td>
</tr>
</tbody>
</table>

Continued in the next page
3.4. Tests

<table>
<thead>
<tr>
<th>Name</th>
<th>$N_*$</th>
<th>$N_{DM}$</th>
<th>$N_p$</th>
<th>$\epsilon_{[pc]}$</th>
<th>$d\eta$</th>
<th>$t_{step}$ (Myrs)</th>
<th>$d\theta$</th>
<th>PKDGRAV (min)</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>M24</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>16</td>
<td>30</td>
<td>0.3</td>
<td>100</td>
<td>1.0</td>
<td>2h27m</td>
<td>2.07</td>
</tr>
<tr>
<td>M25</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>16</td>
<td>30</td>
<td>0.4</td>
<td>100</td>
<td>0.9</td>
<td>2h15m</td>
<td>3.57</td>
</tr>
<tr>
<td>M26</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>16</td>
<td>30</td>
<td>0.4</td>
<td>100</td>
<td>1.0</td>
<td>1h39m</td>
<td>3.40</td>
</tr>
<tr>
<td>M27</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>8</td>
<td>50</td>
<td>0.4</td>
<td>100</td>
<td>1.0</td>
<td>2h4m</td>
<td>4.80</td>
</tr>
<tr>
<td>M28</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>16</td>
<td>50</td>
<td>0.3</td>
<td>100</td>
<td>1.0</td>
<td>2h5m</td>
<td>3.31</td>
</tr>
<tr>
<td>M29</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>16</td>
<td>50</td>
<td>0.3</td>
<td>100</td>
<td>0.9</td>
<td>2h30m</td>
<td>3.12</td>
</tr>
</tbody>
</table>

Table 3.6: $N_p$ is the number of processors the simulation was performed with. * These simulations have all been suspended. M1 was suspended 4 times for a duration of 9 minutes. M3 was suspended for a total of 7 times and a period of 10 minutes. M17 was suspended for a total of 4 times for a period of 15 minutes. M19 was suspended for 2 minutes and M20 was suspended for 1 minute. Note that the simulation times do not always strictly follow expected trends as the times depend on the nodes that each simulation is using (ie. load due to other running programs).

In the comparisons between the test runs from M1 to M29 we use the values of the $\chi^2(obs, sim)$ that is the $\chi^2$ value calculated between the observations and the simulated models. At this point the $\chi^2$ is not aimed to test the fits to the observational data but is used as an indicator of the change in the results depending on numerical parameters. It is seen from Table 3.6 that, as in the high resolution runs, the effect of changing $\theta$ is to reduce the simulation time up to 50% without causing a large difference in the result (eg. M1 vs M2 or M15 vs. M16). However, it is clear that the choice of the right value for $\epsilon$ is important. For instance, models M9, M13 and M15, amongst which the only difference is $\epsilon$, show that although the results for $\epsilon = 30 pc$ and $\epsilon = 50 pc$ are very similar, once the softening is increased to
Sampling and Simulating

3.4. Tests

\[ \epsilon = 100\text{pc}, \] the loss of accuracy of the force calculations affect the simulation result considerably.

On the other hand, the comparison between models M1, M3, M5 and M7 illustrates the effect of changing \( d\eta \). Although the increasing \( d\eta \) values from M1 to M7 gives a desirable reduction in the time that the simulations take, by the time \( d\eta = 0.4 \) ie. in model M7, the \( \chi^2(\text{obs, sim}) \) between has increased by a factor of two. This level of difference is hard to monitor in the chains as we would not know if the \( \chi^2 \) differences were due to the physical properties of the models or were random due the the reduced accuracy of the simulations are not reliable with the numerical parameters that are being used. Therefore, we rule out \( d\eta = 0.4 \). Model M21, also tries an intermediate value of \( d\eta = 0.25 \). Comparing the result of M21 to M22, we find that a change from \( d\eta = 0.25 \) to \( d\eta = 0.4 \), on its own leads to a factor of two change in \( \chi^2 \). The model M21(or M16 on 8 processors) itself is on a good level of accuracy like M1 to M4, while it is much faster. We therefore choose this model for our MCMC runs. Figures [3.3] to [3.5] compare some of these models with the results of the high resolution simulations. In all of these, the low resolution models show an evolution in the inner part of the cusped halo profiles. However, it is seen that, in M21, this evolution is inside 50 pc limit that is the range our simulations cannot resolve. Therefore, at the end of the time tests, the numerical parameters of model M21 is chosen to be used in the MCMC as it has the fastest set of numerical parameters that do not compromise the simulation results.

Before using the above results to start the Markov Chain, we do some more tests. The first of these are the stability tests done in order to show that in the absence of an external potential the galaxies are in equilibrium when using these parameters.
3.4. Tests

For the stability tests we use the model M21 that will be used in the MCMC chain. After the generation of the initial galaxy, it is put on a circular orbit at large radius \( r = 300 \text{ kpc} \) and evolved for \( 4.8 \times 10^9 \) years to check that any initial evolution caused by numerical imperfections halts at the early stages (see Figure 3.7) and the galaxy stays in equilibrium after that point (see Figure 3.8). Figure 3.6 shows the evolution of the \( \chi^2 \) values between the models and the observed Carina data at steps of 300Myrs. The evolution of the \( \chi^2 \) is very small compared to the differences that
Tests for satellites in extreme orbits

Although the stability tests show that the evolution of the satellite seen in the models with gentle tides will not be dominated by numerical problems, a final test is
made in order to test the same issue for models on more extreme orbits. This time, a different model where the eccentricity of the orbit is very high, and the satellite is completely disrupted by the Milky Way, is simulated (see Table 3.7). Figure 3.9 compares the velocity and density profiles of this model found by using the numerical parameters of B1, B7 and M21. Although the profiles are not exactly the same for the high and low resolution models the differences are much smaller than the observational uncertainties of the Carina’s data. Therefore using these tests we confirm that even in the extreme cases, where the satellite is being stripped apart by the tides, the numerical parameters used in M21 are not modifying the results in a way that can affect the distinction between different models (Figure 3.9). As can be seen in Figure 3.11 the orientation of the galaxy looks different between the $2 \times 10^5$ and $2 \times 10^6$ particles simulations at the end of 4.8 Gyrs. However, the rem-
3.4. Tests

Top: Evolution of the galaxy between 600Myrs (red) and 4.8Gyrs (blue) in the orbit. The relaxation differences are very small. After 600Myrs the rate of internal evolution is greatly reduced. Bottom: Residual plots of the test galaxy’s 600Myrs and 4.8Gyrs profiles.

nant shapes are still very similar and the models are approximately the same phase of their evolution. Since we only use the projected stellar density and velocity dispersion profiles in our MCMC runs, the model can be distinguished from others. The change in the orientation comes from the exaggeration of the mass loss caused by tides in the $N = 2 \times 10^5$ simulation where each particle represents a larger mass, which leads to a change in the satellite orbit as its mass decreases.

<table>
<thead>
<tr>
<th>Name</th>
<th>$N_s$</th>
<th>$N_{DM}$</th>
<th>$N_p$</th>
<th>$\epsilon$</th>
<th>$d\eta$</th>
<th>$t_{step}$</th>
<th>$T_{sim}$</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>M21</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>16</td>
<td>50</td>
<td>0.25</td>
<td>100</td>
<td>1.0</td>
<td>5.1573</td>
</tr>
<tr>
<td>B1</td>
<td>$10^6$</td>
<td>$10^6$</td>
<td>16</td>
<td>50</td>
<td>0.1</td>
<td>100</td>
<td>1.0</td>
<td>2.4089</td>
</tr>
<tr>
<td>B2</td>
<td>$10^6$</td>
<td>$10^6$</td>
<td>16</td>
<td>50</td>
<td>0.1</td>
<td>100</td>
<td>0.7</td>
<td>3.7237</td>
</tr>
</tbody>
</table>

Table 3.7: The extreme tests. The changes in $\chi^2$ are affected less than a factor of two for different opening angles, while the effect of the number of particles is larger.
3.4.5 Middle resolution tests

We have done tests in order to see if a higher resolution could be achievable in the MCMC than $2 \times 10^5$ particles. However, as can be seen from Table 3.8, even for the most extreme values of the numerical parameters, the simulation times are very long for $5 \times 10^5$ particle runs. Therefore, no further analysis of these tests were performed.

3.5 Summary

In this chapter we gave a general overview of the techniques that will be used in the rest of this thesis. In the first part of this chapter, N-body methods and tree
structure for data handling were explained. Specifically the details of the k-d tree used in PKDGRAV were given. Later an introduction to Monte Carlo methods and specifically Markov chains was given, followed by the details of the application of the MCMC for doing N-body simulations of Carina.

In the second part, numerical test results were presented which were used to choose the numerical parameters for the simulations taking into account Carina’s properties. The analysis of the test results showed that the maximum number of particles
3.5. Summary

**Figure 3.12:** Velocity distributions in the M21 (top) and B2 (bottom) galaxies of the extreme orbit tests. From left to right: xy plane, yz plane and xz plane. Blue particles are the bulge particles while red represents the halo.

<table>
<thead>
<tr>
<th>Name</th>
<th>N_s</th>
<th>N_DM</th>
<th>N_p</th>
<th>ε</th>
<th>dθ</th>
<th>t_step</th>
<th>dθ</th>
<th>T_sim</th>
<th>χ²</th>
</tr>
</thead>
<tbody>
<tr>
<td>MF2</td>
<td>5 × 10⁵</td>
<td>5 × 10⁵</td>
<td>16</td>
<td>30</td>
<td>0.1</td>
<td>100</td>
<td>1.0</td>
<td>28h45m</td>
<td>0.9</td>
</tr>
<tr>
<td>MF3</td>
<td>5 × 10⁵</td>
<td>5 × 10⁵</td>
<td>16</td>
<td>30</td>
<td>0.25</td>
<td>100</td>
<td>1.0</td>
<td>15h45m</td>
<td>1.1</td>
</tr>
<tr>
<td>MF4</td>
<td>5 × 10⁵</td>
<td>5 × 10⁵</td>
<td>16</td>
<td>30</td>
<td>0.4</td>
<td>100</td>
<td>1.0</td>
<td>10h40m</td>
<td>1.8</td>
</tr>
<tr>
<td>MF5</td>
<td>5 × 10⁵</td>
<td>5 × 10⁵</td>
<td>16</td>
<td>50</td>
<td>0.1</td>
<td>100</td>
<td>1.0</td>
<td>35h40m</td>
<td>0.9</td>
</tr>
<tr>
<td>MF6</td>
<td>5 × 10⁵</td>
<td>5 × 10⁵</td>
<td>16</td>
<td>50</td>
<td>0.25</td>
<td>100</td>
<td>1.0</td>
<td>11h38m</td>
<td>1.1</td>
</tr>
<tr>
<td>MF7</td>
<td>5 × 10⁵</td>
<td>5 × 10⁵</td>
<td>16</td>
<td>50</td>
<td>0.4</td>
<td>100</td>
<td>1.0</td>
<td>7h40m</td>
<td>2.6</td>
</tr>
</tbody>
</table>

Table 3.8: Middle resolution simulations. They are only added for reference. MF2 was suspended once for one minute, MF4 was suspended for 37 minutes and MF5 was suspended for 18 minutes.

To be used in the simulations was 2 × 10⁵ particles due to CPU time constraints, and the set of parameters belonging to test M21 were established to be the most sensible choice for use in the MCMC as the force calculations were still accurate and the individual simulation time for each galaxy could be decreased to a couple of hours. Using these results, 8 chains were submitted to run on 80 processors for 2 months and a total of 3300 simulations were obtained at the end for our artificial data fitting chain.

In the next chapter we present the results of this artificial run and the Carina chains.
run using our algorithm. This chain was run for artificial data of which we know
the properties will be used to quantify the accuracy of our algorithm.
In the previous chapter we introduced an MCMC algorithm to do an efficient parameter space search that can be used to investigate the effect of the Milky Way’s tidal force on a nearby dSph galaxy. In this chapter, we apply the method to data in the Carina dSph. The first part of the current chapter explains our motivation for choosing Carina as our showcase and how we chose the free parameters that will be used in the MCMC analysis. The second part consists of the results of several
Markov chains, including a first set of chains that were run for an artificial data set in order to test the reliability of our results. While we show that our method works, due to the uncertainties in the observational data we cannot provide strong constraints on the parameters for the moment.

### 4.1 Carina dSph

As mentioned in Chapter 3, Carina is one of the dSphs with the most extensive data sets: The currently published data sets consists of the photometry of 68000 stars from Monelli (2003), spectroscopy for \( \approx 900 \) of red giant stars in its central parts from Walker et al. (2007) and Muñoz et al. (2006), as well as bulk proper motions from Piatek (2003).

Based on our current knowledge, Carina seems to have had an exceptional star formation history. Although most dSphs have extended star formation periods, none other than Carina exhibit confirmed episodic star bursts. It has three distinct stellar populations of ages 11, 5, and 0.6 Gyrs (Monelli 2003) and shows a slight population gradient where the more metal rich population is more centrally concentrated in the galaxy (Koch et al. 2006).

The current data for Carina show tidal features with a slightly irregular surface brightness profile in the outer parts, as can be seen in Fig. 4.1 and Table 4.1 as well as projected velocity dispersion profile that increases at large radii seen in Fig. 4.2 and Table 4.2. It is however, important to note that this is the region where the contamination by the foreground stars would be most significant. As explained in the introduction, there are different formation scenarios for dwarf spheroidal galax-
ies. While in the CDM paradigm they are the systems that formed very early on in the evolution of the Universe, MOND models suggest that they were formed at later times on the tidal arms of merging galaxies as overdensities due to shocks were formed. This latter model, that is called a tidal dwarf, is modelled free of dark matter and because of its small mass is hard to reconcile with old dSphs that have survived tidal interactions with the Milky Way for long time. Investigating the effects of tides on Carina would help to distinguish between these models and to calculate the amount of dark matter that is expected to be contained in the dSph to explain its dynamics. Modelling its dark matter content could also be used in comparisons with the CDM simulations’ predictions of cuspy dark matter haloes. In haloes that have similar masses inside similar scale lengths, a cuspy density profile would make the galaxy more stable against tides. In order to investigate whether we can distinguish between different halo profiles, we model the halo of Carina using dark matter haloes with different inner density slopes.

In the past, several studies looked at modelling Carina’s mass distribution taking into account disruptive tidal interaction models. The mass loss rates found for these models range from $7.5\%$ Gyr$^{-1}$ (Muñoz et al., 2006) to $10\%$ Gyr$^{-1}$ (Muñoz et al., 2008) and $30\%$ Gyr$^{-1}$ (Johnston et al., 1999; Monelli, 2003) in the literature, and clearly depend on the observed proper motions $\mu_{\alpha} = 22 \pm 9$[mas/century] and $\mu_{\delta} = 15 \pm 9$[mas/century] that are highly uncertain. Therefore it is crucial to understand the degeneracies between different structural parameters that are used in the models.

Carina’s complex dynamical structure and star formation history raise several questions. The first of these is related to the early loss of gas predicted by the current
cosmological models (Moore et al., 1999). How did Carina hold on to its gas for so long? Furthermore since we currently do not observe gas in Carina (and in fact in any of the Local Group dSphs), where is the interstellar gas that can be associated with the recent star formation activity? Is it the shocks during perigalactic or Galactic disk passages or some other mechanism that have triggered the episodic star burst periods in Carina?

Muñoz et al (2008) have published simulation results of two component dSph models where the stellar and dark matter scale lengths were the same, which they compared to Carina’s observed density and velocity profiles. The stellar component in these Mass-Follows-Light (MFL) models would be protected against the tidal forces exerted on the galaxy by the Milky Way, because of the additional mass that would be added to the very central part of the galaxy. They have found models that could fit Carina’s data well without using additional constraints. The aim of this chapter is to see whether the MFL models that they found are the only models that can explain Carina’s surface brightness and velocity dispersion profiles or whether they are a subset of the probable models in a large parameter space. We investigate different halo shapes and compare tidally disrupting models with those that are not affected by tides in order to determine whether a definitive answer can be obtained with the current data.
An MCMC based search for dynamical models of the Carina dSph

4.1 Carina dSph

Figure 4.1: Surface brightness profile of Carina from photometric data of Walker et al. (2007)

<table>
<thead>
<tr>
<th>$R_{\text{max}}$ [kpc]</th>
<th>$\Sigma [M_\odot/kpc^2]$</th>
<th>$d\Sigma [M_\odot/kpc^2]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.148</td>
<td>$1.7 \times 10^6$</td>
<td>$8.0 \times 10^4$</td>
</tr>
<tr>
<td>0.297</td>
<td>$7.7 \times 10^5$</td>
<td>$3.4 \times 10^4$</td>
</tr>
<tr>
<td>0.445</td>
<td>$2.4 \times 10^5$</td>
<td>$1.6 \times 10^4$</td>
</tr>
<tr>
<td>0.593</td>
<td>$6.6 \times 10^4$</td>
<td>$1.0 \times 10^4$</td>
</tr>
<tr>
<td>0.742</td>
<td>$2.5 \times 10^4$</td>
<td>$8.3 \times 10^3$</td>
</tr>
<tr>
<td>0.890</td>
<td>$8.3 \times 10^3$</td>
<td>$6.8 \times 10^3$</td>
</tr>
<tr>
<td>1.038</td>
<td>$9.0 \times 10^3$</td>
<td>$6.2 \times 10^3$</td>
</tr>
<tr>
<td>1.187</td>
<td>$1.1 \times 10^3$</td>
<td>$5.7 \times 10^3$</td>
</tr>
<tr>
<td>1.335</td>
<td>$5.2 \times 10^3$</td>
<td>$4.2 \times 10^3$</td>
</tr>
<tr>
<td>1.483</td>
<td>$1.2 \times 10^3$</td>
<td>$2.5 \times 10^3$</td>
</tr>
<tr>
<td>1.632</td>
<td>$3.1 \times 10^2$</td>
<td>$1.3 \times 10^4$</td>
</tr>
</tbody>
</table>

Table 4.1: Surface brightness data for Carina (Walker et al. 2007)
An MCMC based search for dynamical models of the Carina dSph

4.1. Carina dSph

Figure 4.2: Projected velocity dispersion profile from spectroscopic data of Walker et al. (2007)

<table>
<thead>
<tr>
<th>R[kpc]</th>
<th>σ [km/s]</th>
<th>dσ [km/s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.056</td>
<td>5.9</td>
<td>1.9</td>
</tr>
<tr>
<td>0.105</td>
<td>6.1</td>
<td>2.0</td>
</tr>
<tr>
<td>0.139</td>
<td>5.6</td>
<td>1.9</td>
</tr>
<tr>
<td>0.17</td>
<td>6.4</td>
<td>1.9</td>
</tr>
<tr>
<td>0.2</td>
<td>6.5</td>
<td>2.1</td>
</tr>
<tr>
<td>0.229</td>
<td>7.6</td>
<td>2.2</td>
</tr>
<tr>
<td>0.262</td>
<td>6.8</td>
<td>2.0</td>
</tr>
<tr>
<td>0.321</td>
<td>5.5</td>
<td>1.7</td>
</tr>
<tr>
<td>0.426</td>
<td>6.0</td>
<td>1.9</td>
</tr>
<tr>
<td>0.829</td>
<td>8.4</td>
<td>2.4</td>
</tr>
</tbody>
</table>

Table 4.2: Dispersion profile for Carina from Walker et al. (2007)
4.2 MCMC

This chapter contains five different sets of Markov Chains (MC1-MC5) used to investigate the degeneracy between the mass and the orbital parameters of Carina. The free parameters and the ranges used for each set of chains are given in Table 4.4 while the fixed structural parameters used in each chain are given in Table 4.3. It can be seen from the table that the set of chains MC1, spanned a lower dimensional parameter space and smaller ranges for the parameters than the other sets. MC1 also consisted of N Body simulations of 5Gyrs for each model. The simulation time was increased to 6Gyrs in the rest of the sets to provide a greater time for galactic tides to have an effect. This difference between sets was mostly due to the fact that the set of 8 chains in MC1 was our initial exploration used to determine the requirements to look for more comprehensive answers within the observational constraints.

The scale length \( r_s \) and values of \( \gamma, \beta \) and \( \alpha \) power law indexes in the stellar component with a split power density profile (see Equation 3.11) are found by initially finding a minimum \( \chi^2 \) fit to the observational surface brightness. However, this does not necessarily give a best fit to the initial profile of the galaxy and was later adjusted by eye to obtain a profile in which the outermost points are slightly different than the observed value, in order to provide an opportunity for the stellar component to evolve with tides and fit the observational data at the end of the simulations. These: \( r_s = 0.37 \)kpc and values of \( \gamma = 0.37, \beta = 7 \) and \( \alpha = 0.4 \) are all used as fixed structural parameters in MC1. As will be seen later, the surface brightness profiles obtained with these initial structural parameters provide very good fits to the observational data. However, as some of our models might evolve due to tides, the stellar scale length might change over the course of a simulation. As the amount of this
An MCMC based search for dynamical models of the Carina dSph

4.2. MCMC

change would depend on the individual models, it is not clear that our choice of the scale length is appropriate for finding all the models that can be gently affected by tides in their outer parts and hence fit Carina’s present day data. In MC1, the models that were affected by tides usually were completely disrupted, and because Carina shows tidal signature only in the outer parts as would be expected by gentle tidal stirring these models did not fit Carina’s data well. therefore this set of chains only found models that were completely unaffected by tides and could fit Carina’s data well with rather flat velocity dispersion profiles that were due to the existence of extended dark matter haloes. Motivated by the lack of models that could match Carina’s data in the presence of gentle tides, we the initial $r_1$ of the dSph is added as a free parameter in the subsequent sets allowing the algorithm to choose the better scale lengths that would evolve tidally to match the present dat Carina scale length. This way, we think that all the parameters that have uncertainties were covered and the results of the chains became more complete. As Table 4.3 shows, these three sets of chains also had different inner halo slopes ranging from a cusped to a cored dark matter density profile ($\gamma = 1; 0.5; 0$).

In order to broaden our results further, one more set of chains was run: MC5 which consists of MFL models that include a two component Plummer model galaxy as in the Muñoz et al. (2008) study.

4.2.1 Choice of parameter ranges

The ranges for the free parameters in the chains are listed in Table 4.4. The motivation for our choices of these ranges are:

---

103
An MCMC based search for dynamical models of the Carina dSph

4.2. MCMC

<table>
<thead>
<tr>
<th>SET</th>
<th>$\gamma_h$</th>
<th>$\beta_h$</th>
<th>$\alpha_h$</th>
<th>$r_h$</th>
<th>$\gamma_b$</th>
<th>$\beta_b$</th>
<th>$\alpha_b$</th>
<th>$r_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC1</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>-</td>
<td>0.37</td>
<td>7</td>
<td>0.4</td>
<td>0.372</td>
</tr>
<tr>
<td>MC2</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>-</td>
<td>0.37</td>
<td>7</td>
<td>0.4</td>
<td>-</td>
</tr>
<tr>
<td>MC3</td>
<td>0.5</td>
<td>4</td>
<td>1</td>
<td>-</td>
<td>0.37</td>
<td>7</td>
<td>0.4</td>
<td>-</td>
</tr>
<tr>
<td>MC4</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>-</td>
<td>0.37</td>
<td>7</td>
<td>0.4</td>
<td>-</td>
</tr>
<tr>
<td>MC5</td>
<td>0.37</td>
<td>7</td>
<td>0.4</td>
<td>$r_s$</td>
<td>0.37</td>
<td>5</td>
<td>0.4</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4.3: Fixed structural parameters of the sets (see Equation 3.11 for definitions). Although the best fit value of $r_s$ was used for MC1, it has been left as a free parameter in the rest of the sets. The scale length are given in kpc.

<table>
<thead>
<tr>
<th>SET</th>
<th>$\mu_\alpha \cos(\delta)$</th>
<th>$\mu_\delta$</th>
<th>$M_*$</th>
<th>$M_h$</th>
<th>$r_h$</th>
<th>$r_s$</th>
<th>$T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC1</td>
<td>22 $\pm$ 27</td>
<td>15 $\pm$ 27</td>
<td>$M_5, 5M_\odot$</td>
<td>$10^5.5 \times 10^8$</td>
<td>0.5</td>
<td>0.37$^*$</td>
<td>5</td>
</tr>
<tr>
<td>MC2,3,4</td>
<td>22 $\pm$ 27</td>
<td>15 $\pm$ 27</td>
<td>$M_5, 5M_\odot$</td>
<td>$10^6.10^9$</td>
<td>0.5</td>
<td>0.07</td>
<td>6</td>
</tr>
<tr>
<td>MC5</td>
<td>22 $\pm$ 27</td>
<td>15 $\pm$ 327</td>
<td>$M_5, 5M_\odot$</td>
<td>$10^5 \times 10^8$</td>
<td>$r_s$</td>
<td>0.07</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 4.4: Ranges of free parameters for the MCMC sets. The proper motions: $\mu_\alpha$[mas/century], $\mu_\delta$[mas/century]; the stellar mass: $M_*$, given in units of Carina’s present day stellar mass $M_S = 4.3 \times 10^5 M_\odot$; the halo mass: $M_h[M_\odot]$; the halo and scale length: $r_h[kpc]$; the stellar scale length: $r_s[kpc]$; and the integration time $T[10^9\text{yrs}]$. ($^*$), fixed to the best fit value of the observations.

1. The proper motion ranges given in Table 4.4 come from the 3$\sigma$ errors of the observed proper motions given in Piatek (2003).

2. The minimum value of the stellar mass $M_*$ is set to the current luminous mass of Carina which is well constrained (Mateo, 1998). However, taking into account the large range of mass loss rates given in the literature, a maximum of $5M_{min}$ is allowed for the initial stellar mass of the models.

3. The dark matter mass is initially chosen between $10^7 M_\odot$ and $5 \times 10^8 M_\odot$ (Muñoz et al., 2008; Irwin & Hatzidimitriou, 1995) for MC1. However, motivated by the results of this set, masses as low as $10^6 M_\odot$ are allowed in the remaining sets as this allows for models with dark matter masses that are
An MCMC based search for dynamical models of the Carina dSph 4.2. MCMC comparable to the stellar mass of Carina. We also increase the maximum halo mass to $10^9 M_\odot$ for these further sets. A larger mass for a nearby dSph is not plausible as it would start to be large enough to affect the Milky Way significantly.

4. The scale length of dark matter, $r_h$ is initially chosen to be smaller than 3kpc. It seems reasonable to assume that $r_h$ would not be larger than $\approx 3$ kpc for a dSph (Mateo, 1998; Martin, 2008). In addition, the kinematic data only trace the potential out to 1kpc. In sets MC2, MC3 and MC4 a wider range of halo scale length values is allowed ($r_h < 5$ kpc) to account for more extreme cases. As the baryonic matter in the galaxy can collapse further than the dark matter by dissipating energy, the lower limit of $r_h$ is set to 0.5kpc and is chosen to always be larger than the stellar scale length.

The MFL case MC5 uses a Plummer model for both the stellar and the dark matter components of the dSphs following Muñoz et.al (2008). The limits are slightly varied for this case as the dark matter scale length cannot be as large as in the other sets, since it is limited by the limits of the stellar scale length. We therefore fix $r_h$ to $0.074 < r_s = r_h < 0.74$. In addition, the maximum halo mass allowed for this chain is decreased as it is not viable to have larger mass models with such small scale lengths which would result in very high velocity dispersion values that would not match Carina’s central velocity dispersion. While the velocity dispersion of the models on very extreme orbits might be reduced by significant loss of mass, as we will see later in this chapter, these models cannot match Carina’s data well.
4.3 Artificial data test

In the previous chapter we have given an outline of our algorithm that is designed to model the dynamical interactions of dSphs with the Milky Way. Before looking at the MCMC results for the Carina dSph, we show a full test MCMC made for an artificially generated galaxy model. Initially, a model that is being tidally disrupted was found. The end result of the simulation is then taken as our “observational” data for this test (Figure 4.3) and we run seven Markov chains in an attempt to recover this model (ART1 from now on) which we know is in our parameter ranges.

The initial physical properties of ART1 are \( M_s = 6.48 \times 10^5 M_\odot \), \( r_s = 0.37 \text{kpc} \), \( M_{DM} = 1.19 \times 10^8 M_\odot \), \( r_{DM} = 1.159 \text{kpc} \) and the orbital parameters are \( \mu_\alpha = 29[\text{mas/cent}] \), \( \mu_\delta = 8[\text{mas/cent}] \). This gives a very eccentric orbit with pericentre and apocentre distances of 15.8kpc and 106kpc, respectively. After 6Gyrs of N Body integration, the tidal rise in the outer part of the velocity dispersion can be seen (Figure 4.3), making this a good model in order to test the capacity of the algorithm to match tidal features. The error bars plotted in the figure are larger than the errors found for the simulated ART1 model. We introduce these, in order to make the “observational” errors of the ART1 model comparable to the observational errors of Carina. It is seen in the figure that the increase in the velocity dispersion that we try to match with our models, becomes statistically insignificant after we increase the errors.

Once the observational properties were calculated for ART1, seven random points were chosen in the parameter space as the starting points of the individual chains. At the end of each N-body simulation, the resulting surface brightness and velocity
An MCMC based search for dynamical models of the Carina dSph

4.3. Artificial data test

Figure 4.3: The “observed” model of the artificial chain test. The time evolution of the orbit show over 6Gyrs (top left), dark matter density (top right), surface brightness profile (bottom left) and the velocity dispersion (bottom right). The error bars are set to be the same size as that of Carina’s observed errors, while the data has additional random noise (obtained using Carina’s actual errors) added in order to take into account of uncertainties.

dispersion profiles were compared to the ART1 profiles given in Figure 4.3. Before calculating the $\chi^2$ between the models in the chain and ART1, we add noise to the “observed” data. We test whether, despite the fact that the right model is contained in our parameter space, all these uncertainties might prevent us from constraining the parameters. In addition, we investigate the correlations between model parameters.

First, we check that the chains find models that fit the data well by looking at the reduced $\chi^2$ values calculated between accepted models and the ART1 data. Figure 4.4
shows the evolution of the $\chi^2$'s for each chain in the MCART set. The chains evolve to models with smaller $\chi^2$ for all of the chains and give several very good models with $\chi^2 < 3.5$. We only include these models in our analysis of the MCART set of seven chains studied in this section.

![Figure 4.4: The $\chi^2$ evolution of the chains in the MCART set. The different colors belong to twelve individual chains run for MCART. The figure excludes the initial burnin period of 10 stars for each chain. In addition, most of the chains explore relatively worse regions in the parameter space as well where models with $\chi^2 > 5$ have been accepted (pink and black lines).](image)

Figure 4.5 shows the distribution of the initial stellar masses ($M_\odot$) that the models have (right panel), as well as the pericentre distances $r_p$ and eccentricities $e$ of their orbits calculated by following the orbit of a point mass representing the satellite in the full Milky Way potential (on the left). The stellar masses of most of the good models cover a range between $4.3 \times 10^5$ to $6.5 \times 10^5 M_\odot$, which is the right area for the value of this parameter in ART1. The correlation between the $r_p$ and $e$ is a common one that we find for all orbits we study as it will be seen later in this
chapter. The figure shows that the orbits with large perigalactic distances are often circular and that models that come closer are in general more eccentric. This is due to the fact that the closest circular orbits that can be used are at an orbital radius of 100 kpc as this is the current observed distance of Carina, and the orbits with closer perigalactic distances need to be more eccentric if they are to pass from Carina’s current position. The peak seen at small $r_p$ in this panel shows the right value for ART1, however, it is seen clearly from the figure that there are a large number of models with large perigalactic distances that could fit the data as well.

![Figure 4.5: On the left: The perigalactic distances and eccentricities found using the MCART set. The contours constrain models at 1σ, 2σ and 90% levels. On the right: The distribution of the initial stellar masses.](image)

As for the stellar mass, the models explore the whole parameter space but favor the correct region. There are several models at $10^8 M_\odot$ which is the actual value for the ART1 model, as well as several other models that fit the data well as seen in the $\chi^2$ distribution for different halo masses (on the panel on the right). It is seen however, that the lowest $\chi^2$ models are those with $1.5 \times 10^7 M_\odot < M_h < 3.2 \times 10^8 M_\odot$. 

Figure 4.6 shows the distribution of the initial halo masses that the models have.
An MCMC based search for dynamical models of the Carina dSph

4.3. Artificial data test

In this set we also check whether we can obtain consistent final masses from our chains, to match the model ART1’s final mass. The mass inside the last data point of the velocity dispersion profiles ($r = 0.82$ kpc) is $6.2 \times 10^6 M_\odot$ for the ART1 model.

The distribution of the final masses at the end of N-Body integrations are shown in Figure 4.7. The mean of the Gaussian is $6.8 \times 10^6 M_\odot$ where the 2σ limit covers the range of values between $6.2 \times 10^6 M_\odot$ to $7.4 \times 10^7 M_\odot$. The results are consistent with ART1.

Figure 4.6: On the left, the distribution of the initial halo masses that the MCART set’s good models. On the right, the distribution of the masses with $\chi^2$ values.

Figure 4.7: Distribution of the final model masses for the best models in MCART.
Figure 4.8 and Figure 4.9 both investigate the correlations that can be found between pairs of parameters. The first column of each of these figures show the stellar scale length values on the x-axes. It is seen that this parameter is constrained well, with a clear peak around $r_s = 0.4$ kpc, which is very close to the actual value of $r_s = 0.37$ kpc in the ART1 model. The second and third columns in Figure 4.8 have the proper motion parameters ($\mu_\alpha, \mu_\delta$) on their x-axes respectively. Given the large error bars in these parameters they are not well constrained: While the chains fail to find the right $\mu_\delta$ value, the $\mu_\alpha$ has a large range of accepted values. However, as seen in Figures 4.5 and 4.9 the right eccentricity and the perigalactic distance that belong to ART1, are sampled and found as peaks in the $r_p$ and $e$ histograms. Thus the chains are correctly recovering the physical properties of the orbit although the proper motions are incorrect.

The additional parameter $\sigma_{10}/\sigma_9$ plotted in Figure 4.9 is the ratio of the velocity dispersions in for the outermost bin in the data to the one before that. This parameter is used to detect when a model looks tidally heated in the outer parts, ie. with an increasing dispersion profile that would match ART1. The histogram shows that despite finding highly tidally disrupting models with $\sigma_{10}/\sigma_9 > 2$, the chains also find several models that have no or very small tidal heating with $\sigma_{10}/\sigma_9 <= 1$.

Finally, although the halo scale length $r_h$ has a wide range of values within the chains, it has a peak at about 1.6 kpc, ie., very close to the correct value for model ART1.

In this section we showed the results of the MCMC run in order to match an artificial galaxy model. The key point of this experiment with the artificial data was to
prove that the algorithm can be used to track the effects of tides and find the models that fit the data well, as well as showing the correlations between parameters. The chains were able to explore the whole parameter space and find the models with the right properties. In addition, we were able to show degeneracies between the orbital parameters. In the rest of this chapter, we will do several new experiments attempting to find good models for the Carina dSph.
4.4 MC1: Test set with Hernquist halo profile

As explained at the beginning of this chapter, before running our main sets of simulations MC2, MC3, MC4 and MC5 with different dark matter halo properties we made another test set where we tried to find models that match Carina’s data, assuming that the stellar scale length does not evolve in time, i.e., it is fixed to its current observed value from the beginning of the N-body simulations. This test set, MC1, considered models with Hernquist halo profiles and a conservative range of values for the 5 free parameters used in the MCMC, namely $\mu_0, M_s, r_h$ and $M_h$. Another difference of MC1 from the other sets was the relatively shorter N-body
integration time of 5Gyrs. Together with the exclusion of one of the free parameters this would mean decreasing the required number of simulations that would be needed for the chains to burn in and converge towards the right values of the parameters. MC1 consisted of eight chains which simulated 1860 models in total. As in the MCART, the chains started at randomly chosen areas of the parameter space and started to converge towards better models. They all have passed the burn in period after 100 simulations as can be seen from the evolutions of the $\chi^2$ of the models in Figure 4.10.

Out of the 1860 simulations that were run, we analyse the 1060 that are the models left after excluding the burn-in time of the chains. While studying the correlations between parameters, however, in order to investigate the models with the best results, we make an additional cut for small $\chi^2$s and exclude models with $\chi^2 > 3.5$ from our analysis. Figure 4.11 shows the stellar masses of the best models, which is the best constrained property that we calculate. We see that for this set the stellar mass of Carina was initially very near its current value, meaning that its stellar component will not appear to be highly tidally disrupted. This is further supported by the lack of highly tidally disrupted models found in this set. Figure 4.12 shows the best model found in MC1. Although there might be a small disturbance in the outer part of the light profile at $r > 1.5\text{kpc}$, this is too small to have a considerable effect in the total stellar mass.

The correlations found between the MCMC parameters are shown in Figure 4.13 and in Figure 4.14 which excludes the stellar mass and replaces the proper motions with derived orbital parameters instead, namely the pericentre distance and eccentricity. A full orbit integration is made for 5Gyrs with steps of 100Myrs.
An MCMC based search for dynamical models of the Carina dSph 4.4. MC1: Test set with Hernquist halo profile

**Figure 4.10:** The evolution of the chains towards smaller $\chi^2$. The upper panel excludes the first 10 simulations of each chain in order to make the figure easier to read. In the lower panel 100 simulations from each chain are excluded and the variation of the $\chi^2$ during the parameter space search is seen. We therefore consider the chains to have converged.
in order to determine the pericentre and apocentre passages using the closest and furthest distances from the centre of the Milky Way. The eccentricity is then calculated using \( e = (r_a - r_p)/(r_a + r_p) \). The trend of decreasing eccentricity with increasing perigalactic distance seen in MCART is also seen in the MC1 sets (see Figure 4.14). On the other hand, both \( r_p \) and \( e \) have clear peaks at \( 90 < r_p < 100 \text{kpc} \) and \( 0.2 < e < 0.3 \). This means that MC1 favoured models where Carina is on a nearly circular orbit at its current distance of 100kpc.

\( \sigma_{10}/\sigma_9 \) shown in Figure 4.14 is again used to determine rising dispersion profiles and quantifying the resemblance of the velocity dispersion profile to the observed value of 1.4 (see Table 4.2). The value of this parameter peaks just below 1 for MC1, therefore most of our models have nearly flat or slightly decreasing velocity dispersion profiles in the outer parts that would be expected in the presence of extended dark matter haloes and lack of strong tidal heating. This is also consistent
Figure 4.12: The dark matter density, stellar surface brightness and stellar velocity dispersion profiles for the best model of the MC1 set. The green curves are the observed profiles while the blue are from the simulation data. The details of the model are given at the end of this chapter in Table 4.7.
An MCMC based search for dynamical models of the Carina dSph

4.4. MC1: Test set with Hernquist halo profile

![Correlation diagram](image)

**Figure 4.13**: The correlations between the free parameters of the 5 dimensional MCMC space, obtained from 1060 simulations in the MC1 set. Each column belongs to the parameter of which the histogram is shown on the bottom row. The first column shows the probability density contours of $\mu_\alpha$ [mas/century], with four other parameters: $\mu_\delta$ [mas/century], $M_*$ [$M_\odot$], $M_h$ [$M_\odot$] and $r_h$ [kpc] from top to bottom. Columns 2, 3, 4 and 5 have $\mu_\delta$, $M_*$, $M_h$, and $r_h$ respectively. The contour levels are set to include 68%, 90%, 95% of the models respectively like before.

with the large pericentre orbits we find in this set.

The halo mass and scale length are both constrained to upper limits seen in their histograms: $M_h < 4 \times 10^8 M_\odot$, $r_h \leq 2$ kpc. In addition both have peak values towards the small end of the mass and scale length values. They are also correlated with each other (see the fourth column of the fourth row in Figure 4.14). The correlation between the mass and the scale length of the halo suggests that there might be a density constraint for the models as there is a proportional increase of the scale length with the mass for the likely models for Carina. If this is a density constraint, then we might be able to place a lower limit on the density of the satellite in order for it to survive tidal disruption; this is investigated further using the MC2 set.
An MCMC based search for dynamical models of the Carina dSph

4.4. MC1: Test set with Hernquist halo profile

In the absence of tidal disruption, the nearly flat dispersion profile in the outer parts would be due to the extended dark matter haloes as mentioned above and consistent with the low $\sigma_{10}/\sigma_9$ values we find in this set. The smallest $\chi^2$ in MC1 is obtained from the simulation of such a model (see Figure 4.12). The values for the MCMC parameters are given at the end of this chapter in Table 4.7. Although we might not have explored the whole parameter space yet, the best $\chi^2$ model of this first set was consistently found in the area of the parameter space where the probability density is highest.

It was explained in this section that although a perfect fit to the observed data was not found yet, MC1 set has already shown that both tidal and non-tidal models can be found to represent Carina data within the current orbital constraints. One thing
to keep in mind is that because of the nature of the data, the $\chi^2$ will naturally favor non tidal models as in this case only the outer data bins need to be neglected. To compensate for this and look for the effect of tides closely in the future we could also add the value of $\sigma_{10}/\sigma_9$ as a prior into our chains. However, in this thesis we do not include this prior in our chains and we let the sets explore the whole parameter space with as few prior assumptions as possible.

The several peaks seen in the contours in Figure 4.13 and Figure 4.14 show that four of the five free parameters might be degenerate to some extent. The stellar mass is the only well constrained parameter and all eight chains run in this set favor of models of Carina where the stellar mass has not changed much due to tides. On the other hand both the mass and the scale length of the dark matter halo are degenerate with each other and so are the eccentricity and perigalactic distance seen in Figure 4.14.

At the end of the analysis, the dispersion profiles of the MC1 models were examined by eye to look for a qualitative common property of the models. What we saw was that the models with the best $\chi^2$ were those that were not being affected much by the tidal effects of the Milky Way, either because they were protected against tides by their massive and extended dark matter haloes or because they were in distant orbits. On the other extreme we found models that were almost completely destroyed by tides and could not fit Carina’s data well. Furthermore, the histograms in Figure 4.13 indicate that $\mu_0$, and $M_{\text{halo}}$, for the models are very close to the allowed limits for these parameters. Motivated by these results, we changed some of our assumptions in the subsequent sets.
First we increased the N-body simulation time from 5 to 6 Gyrs. While still being a rather safe time scale considering the possible evolution of the Milky Way’s potential, this increase in the integration time provides more time for the heating of the outer parts of the satellites by gentle tides. In addition, we add the stellar scale length as a free parameter to the models. We expect that if tides have an effect on the observed properties of Carina, the value of this parameter might have been different in the past and later evolved to its current value. Finally, we also allow a larger range of values to be explored for the dark matter mass and scale length.

4.5 MC2: Hernquist halo profile set

In this section we show the results of the 3300 simulations made in the MC2 set where the dark matter haloes have Hernquist density profiles ($\gamma = 1$) as in MC1. In addition, we include the stellar scale length as a new free parameter with lower and upper limits set to $0.07 < r_s < 0.7kpc$. Increasing the number of parameters also means that the chains will take longer time to converge. However, as it can be seen from Figure 4.15 the initial burn in period is shorter because the chains are able to find good models quickly on including a new dimension. The figure shows the evolution of the $\chi^2$’s for the seven chains after a short burn in of 30 models. There is only one chain with a longer burn in period and another 90 models are excluded from this chain in the analysis.

Figures 4.16 and 4.17 show similar properties as in the MC1 set. First, the stellar mass is the best constrained parameter of the set. Second, the perigalactic distance has several peaks and is inversely proportional to the eccentricity. The peaks in Figure 4.17 show that the more distant models are still favoured in this set. However,
An MCMC based search for dynamical models of the Carina dSph

4.5. MC2: Hernquist halo profile set

Figure 4.15: The evolution of the chains towards smaller $\chi^2$ for MC2. The burn in period of 30 simulations are excluded from the chains. Only one chain (magenta) is still at higher $\chi^2$ values that go up to 60. In the following analysis 90 more models are excluded from this chain.

there are hundreds of other models distributed in two peaks with smaller perigalactic distances as well. As we will see later in this section, this means that in this set, we find models that fit Carina’s data well even though they are being affected by tides.

Figures 4.19, 4.20 and 4.21 show the MCMC parameter correlations. These latter two have an extra $\chi^2$ cut at $\chi^2 < 3.5$ in the remaining models after the initial burn in. The main difference can be seen comparing the histograms of the halo masses between Figure 4.19 and Figure 4.20. Figure 4.19 was only presented here to illustrate the effect of our $\chi^2$ cut. When the explicit $\chi^2$ cut is included the peak of the histogram moves to higher masses at around $m_h = 3 \times 10^8 M_\odot$. This also results in the loss of a whole set of models with small masses and large radii in Figure 4.19.

In Figures 4.20, the halo mass and scale length relation looks rather as these two
An MCMC based search for dynamical models of the Carina dSph

4.5. MC2: Hernquist halo profile set

**Figure 4.16:** The initial stellar mass is again the best constrained parameter in our studies and it does not correlate with any other parameters. The minimum value which is the limit of the histogram at $4.3 \times 10^5 M_\odot$ ([Mateo, 1998]) is set to the observed value of this property for the present day Carina.

**Figure 4.17:** Orbital constraints from the MC2 chain. The pericentre distance $r_p$ is inversely proportional to the eccentricity as in MCART and MC1 and can be omitted as an individual parameter in Figure 4.21. It is however, interesting to see that despite having models at several ranges, this set is not completely dominated by models with large perigalactic distances as in MC2.

**Figure 4.18:** The histogram shows the final masses of the best models inside $r = 0.8$ kpc which is the radius of the outermost kinematic data available. The gaussian that is overplotted has mean of $4.4 \times 10^7 M_\odot$ and $2\sigma$ limits at $2.3 \times 10^7 M_\odot$ and $8.4 \times 10^7 M_\odot$. 

123
parameters being proportional to each other which could be explained with a density constraint. To test this, we include a constant density curve passing through the peaks of the $r_h$ and $M_h$ contours in Figure 4.21.

The curve is plotted using $r = (M/(4\pi \rho_0/3))^{1/3}$ for a constant density $\rho_0 = 6.4 \times 10^6 M_\odot/kpc^3$. For comparison, one model from each peak in the $M_h - r_h$ plane was chosen and the densities calculated within the characteristic scale lengths $r_s$, $2r_s$, $3r_s$, $4r_s$, $r_h$, $3r_h$ of these individual models are compared to the value of $\rho_0$ given above. We find that $\rho_0$ corresponds approximately to the density constrained...
within the halo scale length $r_h$ for each model, which is similar for each model. The model parameters and densities are given in Table 4.5.

For our detailed analysis, we focus on the best models of the MC2 (those with $\chi^2 < 3.5$) presented in Figure 4.20 and Figure 4.21. The constraint on the stellar scale length is seen in the histograms where there is a clear upper limit for this parameter at $r_s = 0.4\text{kpc}$. In addition there is a peak value around 0.33 kpc. This is very close to the fixed $r_s$ value that we used in the MC1 set. However, allowing it to slightly vary helps us to find better models in general in this set. In particular, although the peaks at $\sigma_{10}/\sigma_9 > 1$ are not high, we find models that can fit Carina better in the outer parts with flat velocity dispersion profiles. In the next chapter we study two of these models in more detail with high resolution simulations.

<table>
<thead>
<tr>
<th>$r_s$</th>
<th>$r_h$</th>
<th>$M_h$</th>
<th>$\bar{\rho}(r_h)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.28</td>
<td>1.55</td>
<td>$1.4 \times 10^8$</td>
<td>$(3 \times 10^6 M_\odot/\text{kpc}^3)$</td>
</tr>
<tr>
<td>0.29</td>
<td>2.49</td>
<td>$4.25 \times 10^8$</td>
<td>$(2.4 \times 10^6 M_\odot/\text{kpc}^3)$</td>
</tr>
<tr>
<td>0.09</td>
<td>3.43</td>
<td>$8.5 \times 10^8$</td>
<td>$(2.1 \times 10^6 M_\odot/\text{kpc}^3)$</td>
</tr>
</tbody>
</table>

Table 4.5: The densities within the halo scale length calculated for three models chosen near the contour peaks at low, middle and large halo masses. The density used in the fitting curve of Figure 4.21 is $\rho_0 = 6.4 \times 10^6 M_\odot/\text{kpc}^3$.

Similarly to the multiple peaks seen in the $r_p$ and $e$ histograms, the proper motions also have multiple peaks in the parameter space. There are however, constraints on these, especially a lower limit at $\mu_s = 0\text{mas}/\text{century}$. We study the orbital parameters further in Figure 4.22 which shows that the large $m_h$, $r_h$ models that do not show tidal signatures are the models with the lowest $\chi^2$ models of this set. Nevertheless, there are models with $r_p \approx 10\text{kpc}$ which have very small $\chi^2$s and large $\sigma_{10}/\sigma_9$ on the bottom panel. From the top rows it is seen that these are again the most
4.5. MC2: Hernquist halo profile set

Figure 4.20: The correlations between the five free parameters of the 6 dimensional MCMC space, obtained with the MC2. The figure is organised the same way with Figure 4.8 and the $\chi^2 < 3.5$ condition is used for all the plotted parameters ensuring that only the best models are presented here. The remaining parameter is the stellar mass ($M_\ast$) that is found to be independent of the others and is given in Figure 4.16.
An MCMC based search for dynamical models of the Carina dSph

4.5. MC2: Hernquist halo profile set

The correlations between the orbital, mass and tidal parameters obtained with the MC2. The figure is organised the same way with Figure 4.9 and contains the $\chi^2 < 3.5$ cut. The perigalactic distance that is directly related to the eccentricity is shown in Figure 4.17. The red curve on the panel in the fourth column of the fourth row, is a constant density curve at $\rho_0 = 6.4 \times 10^6 M_\odot/\text{kpc}^3$.

Figure 4.21: The correlations between the orbital, mass and tidal parameters obtained with the MC2. The figure is organised the same way with Figure 4.9 and contains the $\chi^2 < 3.5$ cut. The perigalactic distance that is directly related to the eccentricity is shown in Figure 4.17. The red curve on the panel in the fourth column of the fourth row, is a constant density curve at $\rho_0 = 6.4 \times 10^6 M_\odot/\text{kpc}^3$.
extended, most massive models where the stellar component is shielded against the tides by a massive and extended dark matter halo.

In this section, we presented our first complete set of MCMC chains run for Carina dSph, and explored models with Hernquist halo profiles. We have shown that our chains find constraints on the stellar component’s structure, while the halo and orbit parameters might be degenerate with each other (small halo mass models that are found in less eccentric orbits in Figure 4.21). We also showed that there is a density constraint for the accepted models, which might be a lower limit for the density within the halo scalelength of a Hernquist halo satellite in close interaction with the Milky Way. In the following sections we look at the effects of different halo profiles on our results.

4.6 MC3: Cored halo profile set

The set MC3 consists of 6 chains with a total of 3600 simulations of models with a cored halo profile where the inner slope of the halo density is set to $\gamma = 0$. It was explained at the beginning of this chapter that cored and cusped density haloes respond differently to tides. In this section we investigate this using our chains.

We omit the stellar mass distribution plot for this chain as it is again the best constrained parameter and does not give us any new information about Carina or parameter correlations. The stellar scale length is well constrained as well (see Figure 4.23) with the same upper limit $r_s < 0.4\text{kpc}$ found as in MC2. It is however, degenerate with the halo parameters $M_h$, $r_h$ and the eccentricity $e$. The $M_h$ vs. $r_s$
An MCMC based search for dynamical models of the Carina dSph

4.6 MC3: Cored halo profile set

Figure 4.22: Comparison between the orbital and dark matter parameters of the models in the MC2 set. All three lines have two panels with the x-axes belonging to the perigalactic distance on the left, to the eccentricity on the right. From top to bottom rows, the panels are color coded for the halo mass in units of $10^8 M_\odot$, the scale length of the halo in kpc and the ratio of the velocity dispersions of the last two observational bins.
A simple explanation is obtained by using the Jeans equation: The mass of a system is given as $M \propto \sigma^2 r_s$, where $\sigma$ is the velocity dispersion of the tracer population and $r_s$ is corresponding characteristic scalelength. It is seen that for a given velocity dispersion, $r_s$ is directly proportional to the mass. Let’s assume we have two models with the same total mass, i.e. $M_{H1} = M_{H2} = M_T$ and the same central velocity dispersion $\sigma_1 = \sigma_2 = \sigma$ measured at a radius fixed by the observations: $r_{obs}$. The masses within $r_{obs}$, $M_{H1}(< r_{obs})$ and $M_{H2}(< r_{obs})$ are determined by the scalelengths $r_{H1}$ and $r_{H2}$ over which the total mass is distributed. For $r_{H2} = f r_{H1}$, where $f$ is a constant, these masses are related to each other by $M_{H2}(< r_{obs})/M_{H1}(< r_{obs}) = (r_{H1} + r_{obs})^2/(r_{H2} + r_{obs})^2$, (see derivation in the next paragraph ), hence are inversely proportional to the ratio square of the halo scalelengths. Since we have already seen above that $M_{H2}(< r_{obs})/M_{H1}(< r_{obs}) \propto r_{s1}/r_{s2}$, it is clear that for increasing halo scalelength, the stellar scalelength will decrease. This is due to the stars being confined in a deeper potential well and therefore being more concentrated if they are to match the same velocity dispersion profiles with models that have smaller masses inside the tracer radii. Therefore the apparent relation between the two scalelengths is in fact just a secondary demonstration of the $M_{\text{halo}}$ vs. $r_s$ and $M_{\text{halo}}$ vs. $r_h$ relations.

The masses used to calculate $M_{H2}(< r_{obs})/M_{H1}(< r_{obs})$ are calculated using the mass of a Hernquist profile where $\alpha = 1$, $\gamma = 1$, and $\beta = 4$, the mass for simplicity:

$$M_{H1}(< r_{obs}) = 4\pi \rho_0 r_{H1}^3 \frac{(r_{obs}/r_{H1})^2}{\left(1 + (r_{obs}/r_{H1})\right)^2}. \quad (4.1)$$
An MCMC based search for dynamical models of the Carina dSph

(See equation 2.66 of Binney & Tremaine (2008)). The total mass which is equal for both haloes is calculated using

\[ M_T = \lim_{r \to \infty} M(r) = 4\pi \rho_0 r_H^3. \]  

(4.2)

and the central densities are found to be \( \rho_{0,1} = M_T/(4\pi r_{H1}^3) \) and \( \rho_{0,2} = M_T/(4\pi r_{H2}^3) \).

Using these in Eq. 4.1 we obtain

\[ \frac{M_{H2(< r_{\text{obs}})}}{M_{H1(< r_{\text{obs}})}} = \frac{(r_{H1} + r_{\text{obs}})^2}{(r_{H2} + r_{\text{obs}})^2}. \]  

(4.3)

The \( \sigma_{10}/\sigma_9 \) in Figure 4.24 is again showing that most of the good models do not have strong tidal disturbance in the outer parts. Normally this might be due to both the orbits of the models and their halo densities. We will soon look at the orbital constraints we get from this chain. But keeping in mind that we have not made any supplementary assumptions in this set, we first use this information in order to discuss its implication on the density difference between the cored and the cusped models.

In Figure 4.24 we overplot two additional curves on the contours of \( r_h \) and \( M_h \). The red curve shows the radii that the models have for a constant density \( \rho_0 = 6.4 \times 10^6 M_\odot/kpc^3 \) as in MC2. However, this curve does not fit the models in MC3 as well. Therefore, the blue curve assumes a smaller density, \( \rho_1 = 1.1 \times 10^7 M_\odot/kpc^3 \) which fits the models better. This density limit means that the cored models in general need to be more compact as the constant density core would be affected by any disturbance within its scale length, while the cuspy density profiles are better.
protected even within the characteristic scale lengths as the increase of density towards the centres will protect them from being affected too strongly. This is further confirmed by the lack of very large $r_h$ models in this set. The $r_h$ histograms also show that the constraints on this parameter are strong and it is limited to a value of $r_h < 2 \text{kpc}$.

Finally, although the perigalactic distances of the model orbits are again correlated with the eccentricities in the MC3 set with $\gamma = 0$ models (see Figure 4.25), all the orbital parameters ($\mu_\alpha$, $\mu_\delta$, $r_p$, $e$) are less constrained in MC3 (Figure 4.23 and Figure 4.24).
An MCMC based search for dynamical models of the Carina dSph 4.6. MC3: Cored halo profile set

Figure 4.24: Contours of parameter pair densities for the MC3 set. The figure is similar to Figure 4.23 replacing the proper motions with eccentricity and $\sigma_{10}/\sigma_9$. The red curve on the $m_h$ vs. $r_h$ plot is a constant density curve with $\rho_0 = 6.4 \times 10^6 M_\odot/kpc^3$ that was found to fit the MC2 contours. The blue curve is plotted for models having $\rho_1 = 1.1 \times 10^7 M_\odot/kpc^3$ which is a density value that fits the MC3 set of models better.

Figure 4.25: The relation found for the orbital parameters: $e$, and $r_p$ found for the best models MC3 set.
An MCMC based search for dynamical models of the Carina dSph

4.6. MC3: Cored halo profile set

Figure 4.26: Orbital analysis of MC3 set with a core halo profile. The units and color coding are the same as in Figure 4.22.
Figure 4.26 investigates these relations further. All three rows in the figure show color coded distribution of the models for certain parameters together with $r_p$, $e$ and the $\chi^2$ for each of them. A first inspection of the top and middle panels shows that compact and massive models with $M_h > 10^8 M_\odot$ and $r_h < 1 kpc$ are the best fit models to the data. The bottom panel shows that these also correspond to models with $\sigma_{10}/\sigma_9 < 1.5$. However, it is also seen that models with $1 < \sigma_{10}/\sigma_9 < 1.5$ which would correspond to gently tidally heated models exist and can fit Carina’s data well. This is seen from the blue colored models on the low $r_p$ low $\chi^2$ of the bottom row. As is seen from the other two rows these correspond to the most massive and compact models, which ensures that although being affected by tides on their very close perigalactic passages, they do not completely get destroyed by the galactic tides.

Before we show the results of our next set of chains we make a last test for MC3 results. Although Figures 4.23 or 4.24 could be used to constrain the halo mass ($M_h$) and the stellar scale length ($r_s$), it is interesting that both of these parameters seem to have two peaks in their histograms. In Figure 4.27 we show all the values that were tried for these two parameters. It is seen from the figure that we cannot claim confidently at the moment that all our chains have completely converged to the best regions. Although it is possible that the secondary peaks that the chains find for both of these parameters are just as good as the primary ones, it is also possible that a couple of the chains might not have found the best regions in the parameter space yet and with a larger number of simulations these secondary peaks might become statistically insignificant. We note however, that this would not mean that these regions do not contain good values for the parameters of Carina, and the dichotomy seen as the existence of several peaks might just be explained by the
An MCMC based search for dynamical models of the Carina dSph

4.7. MC4: Shallow cusp halo profile set (\( \gamma = 0.5 \))

In this section we analysed the results of a set with cored halo satellite models, and investigated the difference between this set and MC2 with cusped halo profiles. In the next section we will look at the models with shallow central halo cusps.

**Figure 4.27**: All the stellar scale length and halo mass values used in models (including the rejected models) in the set MC3. We note that although the chains might not have converged completely yet, it is seen that the secondary regions for low halo masses and small stellar scale lengths do not belong to a single chain, i.e. several chains explore different areas.

The MC4 set of chains with \( \gamma = 0.5 \) halo profiles is the largest set of chains we have, with 4400 simulations. In the following we analyse the results for its best models that have \( \chi^2 < 3.5 \).

First, we look at the stellar parameters in this set and we do not study the stellar...
mass any further, as similarly to MC2 and MC3, it is well constrained and does not show any parameter dependance. The stellar scale length is also constrained with an upper limit \( r_s < 0.4 \text{kpc} \) (see Figure 4.28). In addition, the number of models with very small stellar scale lengths that were present in the MC2 and MC3 sets is very low in MC4.

The halo scale length is well constrained giving an upper limit of \( r_h < 1.8 \text{kpc} \) similar to the \( r_h < 2 \text{kpc} \) found in MC3. Differently than MC2 and MC3 however, the halo mass is also better constrained in this set where an upper limit of \( 5 \times 10^8 M_\odot \) at 2\( \sigma \) confidence level can be seen in Figure 4.29). Nevertheless, it might be degenerate with proper motions (Figure 4.28) and eccentricity (Figure 4.29) and is not well constrained. This is also supported by the top row of the Figure 4.31 where models with a large range of halo masses can be found to fit the data well. There are however, as in the MC3 set, a class of models with large mass and small scale length that are on very eccentric orbits with small perigalactic distances that give \( 1 < \sigma_{10}/\sigma_9 < 1.5 \).

Finally, the orbital parameters themselves are not well constrained in this set as in addition to the degeneracy between the proper motions and the halo mass seen from Figure 4.28 there are several peaks of eccentricity values seen in the histogram in this figure. The peaks also correspond to distinct perigalactic distances as expected due to the relationship which is again present between these two parameters (see Figure 4.30).

Until now we carried out our extensive parameter space search with three halo profiles of different fixed inner slopes. For completeness, in the next section, we present
4.8 MC5: Mass Follows Light set

Our final set consists of 2100 simulations run with five chains of MFL models. The density profiles are chosen to be a standard Plummer density with $\rho = \rho_0/(1 + (r^2/a^2))^{2.5}$ where $\rho_0 = 3M/(4\pi a^3)$ is the central density and $a$ is the scale length within which the density is approximately constant. For the MFL model, the scale lengths for both components of the galaxy will be the same: $r_s = r_h$. This means that the initial halo scale length is limited to a much smaller value than before, where it will be limited by the maximum of the stellar scale length $r_s \leq 0.74 \text{kpc}$ which is a large limit considering the present observed scale length of $r_s \approx 0.3 \text{kpc}$. 

---

**Figure 4.28:** The correlations between the five free parameters of the 6 dimensional MCMC space, obtained with the MC4. The figure is organised the same way with Figure 4.8. The red and blue curves correspond to the same constant densities $\rho_0$ given for the MC2 and MC3 sets.
Figure 4.29: Contours of parameter pair densities for the MC3 set. The figure is similar to Figure 4.28 replacing the proper motions with eccentricity and σ_{10}/σ_{9}.

Figure 4.30: Perigalactic distance, eccentricity relationship for the MC4 set.
An MCMC based search for dynamical models of the Carina dSph

4.8. MC5: Mass Follows Light set

**Figure 4.31:** Orbital analysis of MC4 set with a shallow cusped halo profile. The color coding and the units of the parameters are the same with Figure 4.31.
This choice of the limit for $r_s$ is also justified by our results for this set. Figure 4.32 shows that the scale length is well constrained for the Plummer models where we do not find good models with $r_s = r_h > 0.3\text{kpc}$. This constraint means that the MFL models request a more compact set of models to fit the data. It is thus not surprising that the halo mass is also constrained to $M_{\text{halo}} < 6 \times 10^7 M_\odot$, as for larger masses the compact MFL models that can fit Carina’s surface brightness scale length, would result in very high velocity dispersions that are not compatible with the data.

An interesting result of the MC5 chain is that the initial stellar masses are not as well constrained as in the sets MC2 to MC4. Although it is found that even for this set it has an upper limit of $M_s < 10^6 M_\odot$, this means that we now find good models that have lost half of their stellar mass in the last 6Gyrs. In addition $M_s$ is now weakly degenerate with $\mu_\delta$.

While $\mu_\delta$ (top panel of Figure 4.32) seems to be confined to an upper limit: $\mu_\delta < 0.2$, $\mu_\alpha$ is again not well constrained in this set. Their combinations however, results always in highly eccentric orbits with small perigalactic distances as seen in Figure 4.33 and Figure 4.34. As the perigalactic distances are often smaller than 50kpc in this set, this naturally implies that MC5 would find a larger number of models that are strongly affected by the Milky Way’s tidal forces acting on them.

This is confirmed by the large number of tidal models $\sigma_{10}/\sigma_9 > 1$ models seen in Figure 4.33. We note that there are also models with $\sigma_{10}/\sigma_9 < 1$ that simply do not get very affected by tides during our simulations. This is also confirmed by the existence of models that did not lose considerable amount of mass as they started
An MCMC based search for dynamical models of the Carina dSph

4.8. MC5: Mass Follows Light set

Figure 4.32: The correlations between the five free parameters (as \( r_s = r_h \)), obtained with the MC5. The figure is organised the same way with Figure 4.8.

Figure 4.33: As in Figure 4.32. The proper motions are replaced by \( \sigma_{10}/\sigma_{10} \) and the perigalactic distance.
An MCMC based search for dynamical models of the Carina dSph

4.9. Conclusions

In this chapter, we analysed the results of ca.13500 N-body simulations with which we have done a parameter space search in order to determine the structural proper-

---

**Figure 4.34**: Perigalactic distance vs. eccentricity relationship for MC5.

with small initial stellar masses as seen in Figure 4.32 and Figure 4.33

In Figure 4.35, we investigate the orbit-mass degeneracy further. The best models seen in the figure are the ones with either no or relatively low mass loss, as models with an initial mass \( M_* > 8 \times 10^5 M_\odot \) are the ones with a large increase of velocity dispersion in the outer bin and have larger \( \chi^2 \) values. In general the figure suggests that models with \( 5 \times 10^5 M_\odot < M_* < 8 \times 10^5 M_\odot \) where Carina would have lost between 10% and 50% of its stellar mass on very eccentric orbits in the last 6 Gyrs are favoured.

### 4.9 Conclusions

In this chapter, we analysed the results of ca.13500 N-body simulations with which we have done a parameter space search in order to determine the structural proper-
Figure 4.35: Orbital analysis of the MC5 set of MFL models.
ties and the dynamical evolution of the Carina dwarf spheroidal galaxy. Table 4.6 summarizes the results that we have found for the different sets of Markov Chains we have run for this investigation. It also includes the results of two additional tests MC1 and MCART. While none of the good models in the MC1 could reproduce the increase in the velocity dispersion in the outer parts of the models, we used this set as a diagnostic test to help us choosing the range of parameter values we use in the set MC2 to MC5. On the other hand, in MCART we used the result of a previous simulation of tidally disrupting galaxy as our observational data and tested our method on this artificial model. For MCART, we have obtained constraints on the types of orbits and the model scale lengths that are consistent with the data. Although we have found models with the right stellar and halo masses, the peak values of the distributions of these two parameters are found to be slightly lower than the the halo and stellar mass of the target model. However, despite the fact that this set of chains used models where the errors were chosen to be worse than in our science chains, we have shown that we do not rule out the right values for the parameters and are able to get constraints in a general sense.

Three of our four main chains had split power density profiles with different inner slopes for the halo densities: $\gamma = 1$, $\gamma = 0$, and $\gamma = 0.5$ for MC2, MC3 and MC4 respectively. MC5 consisted of two component Plummer models where the halo and stellar scale lengths were equal to each other, generating Mass Follows Light models.

Table 4.6 shows that one of the main differences between MC5 and the other set is that, this is the only set for which the constraint on the stellar mass is not as strong. Although each set found some models that were being tidally disrupted and some
not, MC2 to MC4 favoured models with stellar components that were not affected by tides strongly (shown with NT in Table 4.6), while MC5 have found a large number of models in which the galaxy had lost more than 10% of its stars in the last 6Gyrs.

We have also seen that the best orbital constraint could be obtained in the MC5 as the other sets had numerous models with both close and far perigalactic passages.

Finally, the stellar scale length was constrained by all of our sets to $r_s < 0.4$ kpc, and the halo scale length was constrained to $r_h < 2$ kpc with the exception of MC2. On the other hand the two scale lengths seem to be inversely proportional to each other in the cored models. Finally we found a density constraint within the halo scale length for which we showed the dependency on the halo shape as we found that the mean density within the halo scale length is smaller for the cored models than the cusped ones.

<table>
<thead>
<tr>
<th></th>
<th>MC1</th>
<th>MC2</th>
<th>MC3</th>
<th>MC4</th>
<th>MC5</th>
<th>MCART</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_\alpha$</td>
<td>P</td>
<td>–</td>
<td>P</td>
<td>–</td>
<td>–P</td>
<td>–</td>
</tr>
<tr>
<td>$\mu_\delta$</td>
<td>–</td>
<td>(&gt; 0.1)</td>
<td>P</td>
<td>P</td>
<td>(&lt; 2)</td>
<td>(&lt; 0.2)</td>
</tr>
<tr>
<td>$r_p$</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>–</td>
<td>P</td>
<td>+</td>
</tr>
<tr>
<td>$M_h$</td>
<td>P</td>
<td>P</td>
<td>2P</td>
<td>+</td>
<td>(&lt; 6 x 10$^7$)</td>
<td>+(&lt; 2 x 10$^8$)</td>
</tr>
<tr>
<td>$r_h$</td>
<td>(&lt; 2.2)</td>
<td>–</td>
<td>(&lt; 2)</td>
<td>(&lt; 2)</td>
<td>(&lt; 0.32)</td>
<td>P</td>
</tr>
<tr>
<td>$M_s$</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>&lt; (10$^6$)</td>
</tr>
<tr>
<td>$r_s$</td>
<td>*</td>
<td>(&lt; 0.4)</td>
<td>2P, (&lt; 0.4)</td>
<td>(&lt; 0.4)</td>
<td>(&lt; 0.32)</td>
<td>+,P</td>
</tr>
<tr>
<td>$\sigma_{10}/\sigma_9$</td>
<td>+NT</td>
<td>+NT</td>
<td>+T,NT</td>
<td>+NT</td>
<td>-T,NT</td>
<td>-T/NT</td>
</tr>
</tbody>
</table>

Table 4.6: P= Well defined peak. *= Not free. += Completely constrained. – = Not constrained. T=tidal(rising $\sigma$ profile), NT=Non-tidal (decreasing $\sigma$ profile). The units are [mas/year] for the proper motions, [kpc] for $r_p$, $r_h$, $r_s$ and [M$\odot$] for the masses.

Table 4.7 shows the parameters found for the best models of each set. It again
shows that apart from the MC5 set, the best models of each set agree on the structural properties of Carina. The halo mass is constrained within a factor of 3 and the scale length a factor of 1.4. The stellar mass and scale length shows very small variation between different chains. In addition the MC5 models are also in the ranges that would be accepted within the upper limits set by the other sets.

<table>
<thead>
<tr>
<th>MODEL</th>
<th>$\mu_\alpha \cos(\delta)$</th>
<th>$\mu_\delta$</th>
<th>$M_\star$</th>
<th>$M_h$</th>
<th>$r_h$</th>
<th>$r_s$</th>
<th>$r_p$</th>
<th>$e$</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C3101</td>
<td>37</td>
<td>30</td>
<td>$4.3 \times 10^5$</td>
<td>$8.7 \times 10^7$</td>
<td>0.99</td>
<td>0.372*</td>
<td>94</td>
<td>0.2</td>
<td>1.4</td>
</tr>
<tr>
<td>C1259</td>
<td>7</td>
<td>33</td>
<td>$4.4 \times 10^5$</td>
<td>$2.2 \times 10^8$</td>
<td>1.1</td>
<td>0.32</td>
<td>80</td>
<td>0.3</td>
<td>1.4</td>
</tr>
<tr>
<td>C1575</td>
<td>36</td>
<td>30</td>
<td>$4.4 \times 10^5$</td>
<td>$1.7 \times 10^8$</td>
<td>0.71</td>
<td>0.35</td>
<td>97</td>
<td>0.2</td>
<td>1.7</td>
</tr>
<tr>
<td>C1392</td>
<td>37</td>
<td>29</td>
<td>$4.4 \times 10^5$</td>
<td>$1.5 \times 10^8$</td>
<td>0.8</td>
<td>0.36</td>
<td>95</td>
<td>0.2</td>
<td>1.1</td>
</tr>
<tr>
<td>M1373</td>
<td>24</td>
<td>4</td>
<td>$65 \times 10^5$</td>
<td>$2.5 \times 10^7$</td>
<td>0.2</td>
<td>0.2</td>
<td>8.5</td>
<td>0.85</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Table 4.7: The best models of each set chosen with best $\chi^2$ models (MC1 to MC5 from top to bottom). * The value of $r_s$ was fixed for MC1. Units are [mas/century] for the proper motions, [M$_\odot$] for the stellar and halo masses and [kpc] for the scale lengths and perigalactic distance.

We emphasize that although two of our sets (MC2, MC4) found fewer models with tidal signatures this does not rule out that these tidal models could be the right models. This is clear from the result of our artificial chain where we had a tidally disrupting model that we knew was a model that was included in our parameter space. However, due to the observational uncertainties our chains did not obviously favor these models.

In the next chapter we will show more detailed analysis of four of the good models we have found in our chains that we resimulated with higher resolution.
In Chapter 4, we have presented a new MCMC tool that can be used for the mass modelling of dwarf galaxies. Because of the high number of simulations that was needed in the parameter space search, the resolution of the simulations that were performed was limited. In this chapter, we show the results of four high resolution re-simulations performed for some of the best models found in the MCMC chains. Before we explain the details of these, we emphasize that our low resolution simulations still resolved the models relatively well with $2 \times 10^5$ particles, compared to simulations in the literature which are composed of less than $10^5$, or only slightly
more, \(3.1 \times 10^5\) particles, but for a smaller parameter space coverage (see Sales (2010); Muñoz et al. (2008)).

For our high resolution simulations, we choose three models which are not considerably affected by tides and one that is. As Table 5.1 shows, all our models have good \(\chi^2\) values in the low resolution models. We use the high resolution simulations to check that the results are similar to the low resolution simulations, ie. the chains were not biased because of the resolution limit. We also monitor these simulations at different evolutionary times. Therefore, we are able to quantify the evolution of the mass for 6Gyrs of N-body integration. We also present two dimensional velocity and velocity dispersion maps for each of the models as these can be compared to the observational data and may be useful for designing future observational programs. Finally, we also show the deviations from Gaussianity in the velocity distributions in galaxies that are undergoing tidal disruption, which would need to be detected in the outer parts of galaxies with careful determination of member stars.

### 5.1 Models

The simulations in this chapter are performed using \(4 \times 10^6\) particles and a softening length of \(\epsilon = 20\) pc to obtain better resolution than the MCMC simulations presented in chapter 4. The list of all the model parameters is shown in Table 5.1. It can be seen in the table that three of the models have generally similar values for the parameters while the last one is considerably different. As it will be seen later, this latter one is also the only model that is strongly affected by Milky Way’s tidal forces acting on the satellite.
We emphasize that the models used in this chapter are not the best models that were found in the chains studied in Chapter 4 with the exception of $C^{28}_{7}$ which was the lowest $\chi^2$ model of MC2. This is due to the fact that we kept running the chains for a long time after we had chosen good models to do our high resolution investigations. In addition, in Chapter 4 we showed that there was no single good model of Carina that could rule out all the others. Furthermore, we showed in Chapter 3 that although the models do not change qualitatively, the $\chi^2$ values will not stay exactly the same once the resolution is increased. The following section shows that our choice of models is appropriate as the results are robust and both the high and the low resolution simulations of the models fit Carina dSph’s data very well.

5.2 Dynamical Evolution

Despite their similarities, model $C^{28}_{7}$, (Figure 5.2) has a larger central velocity dispersion than $C^{22}_{38}$ (Figure 5.1) due to the larger halo density in $C^{28}_{7}$. Nevertheless, these two $\gamma = 1$ models of MC2 show similar evolution to each other in the sense that in the low resolution simulations the cusps of the haloes dissipate energy.
over time, while they both survive in the high resolution simulations. However, this difference is mainly in the innermost halo points where our simulations cannot resolve and the observed profiles for the surface brightness and the line of sight velocity dispersion are in agreement between the high and the low resolution models. Both of these models have orbits with large pericentres and with small eccentricity which protect them against strong tidal effects of the Milky Way, as will be shown in the next section.

Model $GN1_{141}$ from MC4 which has a cored halo profile with $\gamma = 0$ is shown in
High resolution simulations  5.2. Dynamical Evolution

Figure 5.2: $C28_7$ model’s halo density (top left), surface brightness (bottom left), tangential velocity dispersion (bottom left), and line of sight velocity dispersion (bottom right) profiles. The color coding is the same as in Figure 5.1.

Figure 5.3 This galaxy is on a distant orbit of which the pericentre is close to Carina’s current distance, and hence is essentially immune to tides, not showing tidal evolution in the low nor the high resolution case. On the other hand, the velocity dispersion is in good agreement with Carina as the extended halo mass results in a flat velocity dispersion profile that can fit the data within the error bars.

It is seen from Table 5.1 that the above three models have stellar scale lengths corresponding to the good areas for this parameter constrained by the sets MC2, MC3 and MC4 of the previous chapter. Apart from their orbits, it is also their extended dark matter haloes that make the stellar components of these models protected against tides. This can be shown with a simple calculation of their tidal radii.
using \( r_t \approx r_p (M_{\text{model}}/3 M_{\text{MW}})^{1/3} \). The results are given in Table 5.2. It is seen that in the absence of the dark matter mass, these galaxies would show signs of tidal interaction at radii as small as 400 pc.

On the other hand, Model \( M2_{11} \) shown in Figure 5.4 is a Mass Follows Light (MFL) model with Plummer profile. This model is different than the three above in several ways. First, it is on a more eccentric orbit, of which Carina’s current distance is the apocentre, and the pericentre distance is \( r_p \approx 10 \) kpc. The lack of an extended halo together with this extreme orbit makes it more prone to tidal disruption which is seen as a break in the outer part of the surface brightness profile and rise in both the line of sight and tangential velocity dispersions.
5.2.1 Mass loss

In all of the four models, the halo masses are larger than the stellar masses. However, the models differ in the distribution of mass within the dSph. While in the models C22, C28, and GN1, most of dark matter is outside the $3r_s$: $DM_{gn1}(<3r_s) = 4.6 \times 10^7 M_\odot$, $DM_{h1}(<3r_s) = 6.1 \times 10^7 M_\odot$ and $DM_{h2}(<3r_s) = 6.2 \times 10^7 M_\odot$; the MFL model has almost all its mass within the central part $M_{M11,DM}(<3r_s) = 2.6 \times 10^7 M_\odot$. Taking into account their more compact nature as well, this allows the MFL models to generally approach Milky Way closer in without being completely
High resolution simulations

5.2. Dynamical Evolution

<table>
<thead>
<tr>
<th>MODEL</th>
<th>(r_p)</th>
<th>(M_{MW}[M_\odot])</th>
<th>(r_{TD})</th>
<th>(r_{TS})</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC2, C22\textsuperscript{38}</td>
<td>65</td>
<td>(5 \times 10^{11})</td>
<td>3.8</td>
<td>0.4</td>
</tr>
<tr>
<td>MC2, C28\textsuperscript{7}</td>
<td>80</td>
<td>(5 \times 10^{11})</td>
<td>4.2</td>
<td>0.5</td>
</tr>
<tr>
<td>MC3, GN\textsuperscript{141}</td>
<td>90</td>
<td>(5 \times 10^{11})</td>
<td>4.8</td>
<td>0.6</td>
</tr>
<tr>
<td>MC5, M\textsubscript{211}</td>
<td>10</td>
<td>(4.4 \times 10^{10})</td>
<td>0.6</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Table 5.2: The tidal radii of the high resolution models. The mass of the Milky Way is taken from model 4a in Dehnen & Binney (1998). For the top three models this is the mass within \(r = 100\text{kpc}\), while for the MFL model we used the mass within \(r = 10\text{kpc}\) as the perigalactic distance of the orbit of this model is 10kpc. \(r_{TD}\) and \(r_{TS}\) are respectively: the tidal radii that these models have for their total masses or if we only considered their stellar masses.

destroyed by tidal interactions in 6Gyrs.

Figure 5.5: Orbits and stellar masses inside 3\(r_s\) of the high resolution models. The model that fits Carina’s data well with strong tides is M\textsubscript{211} MFL model which has several close pericentre passages. It is seen from the figure that the main mass loss occurs during the pericentre passages.

Figure 5.5 shows the integrated orbits and total stellar mass enclosed within 3\(r_s\) for the four models with high resolution simulations. It is seen in the figure that the three models on orbits that have large perigalactic distances do not lose stellar mass due to the tidal forces. On the other hand, M\textsubscript{211} with close pericentre passages loses nearly 30% of its stellar mass over 6Gyrs and ends with an observable mass that is
in agreement with the observational value like the other three models. It is clear from the figure that the mass loss is due to the pericentre passages, as at epochs that the galaxy is at larger distances its mass remains nearly constant. Therefore, this is the kind of model that could explain the episodic star bursts of Carina, in relation to the orbital epochs (Pasetto et al., 2010).

5.2.2 Following the streams

Figure 5.6 and 5.7 show the density contours of the resulting galaxies in models C22\textsubscript{38} and M2\textsubscript{11}, respectively. The top and bottom panels in Figure 5.6 compare the galaxies of the high and low resolution simulations of C22\textsubscript{38}. Although the low resolution simulation galaxy’s contours are less regular, there is no obvious difference between these figures. On the other hand, the poorer sampling of the less dense areas in the outer parts show that studying the tidal tails is harder with these low resolution simulations.

This is better seen in Figure 5.7 where the data from the high resolution simulation of the M2\textsubscript{11} model are plotted on the top two rows. The result of the low resolution simulation shown on the bottom row is again very similar to the high resolution result seen in the middle row, only differing in the sampling of the outer parts. The top-most row on the other hand shows clearly that the wraps of the tidal tails are well sampled in the high resolution simulation.
5.3 Velocity distribution

So far in this chapter we have focused on the mass distribution of the models. However, the velocity information is also crucial both for modelling masses using the velocity dispersions and for detecting tides in the velocity gradients. As stars will be lost by a galaxy that is being tidally disrupted by the Milky Way, in the simplest way the stars nearer to the Milky Way will seem to have higher velocities towards it than the ones that are on the far side of the satellite galaxy. In reality the gradient will depend on the individual model and its orbit. However, a velocity gradient between the different edges of a satellite galaxy is a good confirmation of tidal disruption. In this section, we look at our high resolution simulations in more detail, comparing their velocity dispersion and mean velocity distributions on 2D maps as they would be seen in observations. We also study the distribution of velocities at different radii in order to quantify the number of stars that would be
Figure 5.7: Density contours and scatter plots for the middle (top) and low (bottom) resolution simulations for model M211. The top panel shows the model simulated with $4 \times 10^6$ particles in a larger area where all the tidally stripped stars can also be seen in tidal tails wrapping the orbit. The middle row is plotted again for the high resolution simulation. However, it shows the same range of radius covered by the bottom row and shows that no clear shape or orientation difference can be seen.
Figure 5.8: Two dimensional line of sight mean velocity (left) and velocity dispersion (right) maps of the $M2_{11}$ models. The top figures belong to the low resolution simulation from the chains while the bottom row belong to the high resolution data. The black line in the top left panel is the direction of the observed major axis of Carina with the position angle, $P.A. = 65^\circ$ [Irwin & Hatzidimitriou (1995)] plotted on our simulation data.

observed as parts tidal tails at outer radii.

Figure 5.8 shows the two dimensional maps of the line of sight velocity dispersions and the mean velocity for the low and high resolution simulations of model $M2_{11}$. The black line overplotted on the high resolution velocity map indicates the direction of the observed major axis of Carina as calculated by Irwin & Hatzidimitriou (1995). As we do not use this as a condition in our chains, we do not expect to necessarily get this right for each good model. However, as it lies perpendicular to the direction of the velocity gradient that we see in both panels, this shows that we cannot compare the line of sight velocity distribution over the major axis (see
Figure 5.9 directly to the observational data. We also note that we do not see any signs of apparent rotation within the central part of the galaxy.

The dispersion maps plotted on the right panels of Figure 5.8 are also consistent between the high and low resolution simulations that both have $\sigma_{\text{los}} = 8\text{km/s}$ in the central part, decreasing towards the outer parts. It is noteworthy, however, that as the sampling of the tidally heated parts outside 1kpc is poorer in the low resolution simulation the high dispersion areas expected to be seen outside this radius are not as dominant as in the high resolution case. Interestingly, this might explain the cold populations seen by Wilkinson et al. (2004) at large radii in Draco and Ursa Minor as the velocity distribution would be expected to be undersampled in those observations too.

Figure 5.9 show only the high resolution distribution of the velocity and velocity dispersion for the remaining three models: $C22_{38}$, $C28_7$, $GN1_{141}$. None of these models have the hot outer parts seen in $M2_{11}$. There might be a very small velocity gradient due to tidal heating in $C28_7$ in the middle panel of the figure. However, this seems to be mostly noise.

Finally, we look at the distribution of velocities at different radii in our simulations. First, Fig 5.11 and 5.12 show the velocity histograms of the $M2_{11}$ model at different radii, in the high and low resolution cases respectively. In both cases most of the velocities are distributed normally at each radius. The only deviation from the normal distribution is seen in the outermost parts of the models, inside 1.7 kpc from the centre. Two opposite signed wings on both sides of the Gaussian distributions are visible between $1.5\text{kpc} < r < 1.7\text{kpc}$ for the high resolution simulation (Fig 5.11).
Figu 5.9: The mean line of sight velocity of our model calculated along the observed major axis of Carina. Although this is of interest as it shows the magnitude of the velocity gradient to be seen within the inner kiloparsec (≈ 10km$^{-1}$kpc$^{-1}$), it is not directly comparable to observations as the model does not match Carina’s orientation on the sky.

and 1kpc < r < 1.7kpc for the low resolution simulation (Fig 5.12). This shows that the tidal tails would be detectable by the observations within 1.7 kpc.

In order to test whether these wings are a characteristic property of the particular $M_{211}$ rather than being a clear sign of tides, we test two other models. The histograms on Figure 5.13 and Figure 5.14 belong to the $C22_{38}$ model and another good $\chi^2$ model ($C33_0$) chosen among the models in the MC2 which showed tidal features. This model has only the low resolution simulation. However, it is added in this chapter as it shows that the $M_{211}$ is not an exceptional model and that other models that are affected by tides have the same velocity dichotomy in the outer parts as well. In this case the wings are visible at smaller radii starting from 850pc. As can be seen in Figure 5.13 the gradient related wings are not seen in the $C22_{38}$
5.3. Velocity distribution

Figure 5.10: Two dimensional line of sight mean velocity (on the left) and velocity (on the right) dispersion maps of our models. From top to bottom: $C22_{38}$, $C28_{7}$, $GN1_{141}$. 
5.3. Velocity distribution

**Figure 5.11:** Line of sight velocity distribution of the high resolution $M_{211}$ model, inside $r < 0.15\, kpc$ (450000 particles) green, $r < 0.28\, kpc$ (900000 particles) blue, $r < 1.5\, (450900$ particles) red, $1.5\, kpc < r < 1.7\, kpc$ (140000 particles) magenta.

**Figure 5.12:** Line of sight velocity distribution of the low resolution $M_{211}$ model, inside $r < 0.18\, kpc$ (20000 particles) green, $r < 0.3\, kpc$ (40000 particles) blue, $r < 1.0\, (24000$ particles) red, $1\, kpc < r < 1.7\, kpc$ (1560 particles) magenta.
model that does not show signs of tidal interactions.

The distribution of velocities in the outer parts of these models can tell us the ratio of the number of stars in these wings with high relative velocities, to the total number of stars that we expect to see in the outer parts of a dSph galaxy. While model $C22_7$ is completely Gaussian at all radii (see Figure 5.13) and only 4% of the stars in the outer bin are outside the $2\sigma$ limit of the velocity mean, in the models that are being tidally disrupted this ratio is much higher. In the low and high resolution $M2_{11}$ models’ distributions in Figure 5.11 and Figure 5.12, 40% and 17% of all the stars in the outer bin are outside the $2\sigma$ interval of the velocity peaks respectively. This is in agreement with our conclusion of the tidal tails being less populated in a lower resolution simulation. The final model in Figure 5.14 also has 35% of its large radii stars in the wings of the distribution. Clearly the exact ratio will depend on the individual galaxy. However, the numbers quoted here show that for galaxies
that are currently undergoing highly effective tidal disruption, we should be able to
detect these outer populations of star in the near future. Although we observe the
giant stars in the dSphs which is a population of stars to observe, using the nar-
row band imaging for pre-selection of the targets and the chemistry information to
determine the populations, we can use spectroscopic data out to great distances to
observe these tidal tails.

5.4 Conclusions

In this chapter we have presented the results of four high resolution re-simulations
that were done for good models found in our sets of MCMC simulations in Chap-
ter 4. We showed that our results are robust for higher resolution simulations of
our good models. However, it was also seen that in the cases with tidal tails the
low resolution simulations would not sample the tails sufficiently to use these as additional constraints in the simulations.

We have also compared the mass loss rates of models and were able to show that on one hand we could have gentle orbits with models that had no change in the stellar masses over the course of 6 Gyrs, while on the other hand we could match the pericentre passages of an extreme orbit to a total loss of mass of the order of 30\% for a model.

Finally, we have shown that we can distinguish between the models under the effect of strong tides from those that are not, using first two dimensional mean velocity and velocity dispersion maps. Our dispersion maps showed that the undersampling of the velocities at large radii could result in the detection of an apparently cold kinematic population. Finally, we also showed that non-Gaussianity in velocity distributions should be significant enough to be detected in the outer parts of tidally disrupting galaxies.
In this thesis, we studied the present day kinematics and dynamical evolution of the Local Group dSphs using observational data, Monte Carlo methods and N Body simulations. This work consists of two distinct parts: one in which we studied kinematic stellar sub-populations in the context of modelling the current mass contents of the dSphs and another in which we used stellar kinematics to model the time evolution of a dSph’s mass as it dynamically interacts with the Milky Way.

In Chapter 2, we presented a new data set of velocities and metallicities for the
CVn I dSph, based on spectra taken with the GMOS-N spectrograph on the Gemini North telescope. We found a mass of $4.4^{+1.6}_{-1.1} \times 10^7 M_\odot$ which is in agreement with the currently published kinematic data. We used our data set to search for a centrally concentrated, kinematically cold stellar sub-population that was found in CVn I by Ibata et al. (2006). As we could not confirm such a signal in our data, we performed Monte Carlo tests with artificial data sets in order to determine at which level of confidence we can trust the detections of stellar sub-populations in very small data sets such as we have for many dSphs. We showed that a sample size of $N < 100$ is usually too small to draw meaningful conclusions about the sub-structures, except for the very extreme cases where the sub-structures have very distinct properties such as large differences in the mean metallicities or very different velocity dispersions. In the special case of CVn I this is one of the reasons we could not find a conclusive answer to whether a cold sub-population existed: The previously claimed sub-populations in this dSph seemed to have such different velocity dispersions from each other that they should be detectable even in small data sets. However, we also showed that, if a cold population has a relatively small ratio of stars compared to a hotter population in the same system, this is harder to detect. This means that for CVn I the study should be repeated by looking for these distinct populations in new observations that include all 42 stars from both of the data sets.

In the second part, we took a step further and studied the evolution of a dSph in time, using N-body simulations. We showed that the uncertainties in the observations caused additional complications in modelling the dynamical evolution of the dSphs. By searching for dynamical models that can reproduce the observed data of the Carina dSph, we investigated the differences between models with different dark
Conclusions

matter halo profiles. Our Markov Chain Monte Carlo algorithm provided us with a way of uncovering the degeneracies between the observed and model parameters, while looking for the best-fit models.

Overall, taking into account all the different halo types (cusped, cored, Mass Follows Light models) we found a high number of probable models that could match the data within the current observational constraints. In addition, we found that even for a fixed halo profile, would each provide us with several models that can fit the data. However, we showed that the different halo profiles provided us with constraints on the properties of these galaxies such as the scale length and the densities. In particular, we showed that there is a clear density difference between the cusped ($\rho_0 = 6.4 \times 10^6 M_\odot/kpc^3$) and the cored ($\rho_0 = 1.1 \times 10^7 M_\odot/kpc^3$) models due to the latter being more compact in order to survive tidal forces exerted by the Milky way’s host potential.

However, in the case of Carina, we showed that except for the Mass Follows Light models, whichever halo profiles we used the best models that we found had similar masses ($M_{DM} \approx 10^9 M_\odot$), scale lengths ($r_h < 2\text{kpc}$, and $r_s < 0.4\text{kpc}$) and orbits with ($r_p \approx 90\text{kpc}$, $0.2 < e < 0.3$). Since we find that more eccentric orbits are favoured in the Mass Follows Light models, it is reasonable to think that once the proper motions of the dSphs are better constrained we can determine the plausibility of this class of models.

Furthermore, if the Mass Follows Light profiles that provided most of the models that are being tidally disrupted, can be ruled out, this would mean that our main favoured models for Carina are those with large perigalactic distances (considering
Conclusions

only the best $\chi^2$ models given at the end of Chapter 4. This would mean that the effect of the pericentre passages during the orbit might not be strong enough to trigger the several bursts of star formation seen in Carina.

Finally we used the results of our MCMC based analysis for choosing some good models that we studied in more detail. Using the high resolution re-simulations of these models we have investigated the observable differences between the models that are being affected by tides and those which are not. Apart from the velocity gradients that we could identify in the “tidal” models, we also showed the the apparent existence of cold populations in the outer parts of the observed dSphs can be due to the undersampling of the velocities in these areas. In addition we found that while the velocities in the inner parts of all these models followed Gaussian distributions, in the outer parts, out to $r < 1.7\text{kpc}$, more than $\approx 20\%$ of the stars fall in broad wings on either sides of these velocity Gaussians’ $2\sigma$ limits. This means that the kinematic data would have to be treated differently to detect these: instead of a direct velocity cut to define the member stars, one would need to take into account chemical abundances and the orientation of the tails as a further membership criterion.

To conclude, we can say that both parts of this work showed that despite having a large amount of information about the evolution and structures of the dSphs, the stellar populations need to be modelled with more and better constrained observational data: especially kinematic data.


6.1 Future plans

While in this thesis, we studied the kinematics of Carina and CVn I dSphs in detail, more importantly we have developed two methods that can be used to study dSphs in general. The use of Monte Carlo experiments in order to determine the significance of the stellar sub-population detections is a simple method that should be used to test the sub-population detection in other low luminosity systems. Furthermore, the information from the stellar sub-populations can be added to our MCMC algorithm and would provide additional constraints on the dynamical models that can be used to match the data.

The MCMC can also be used as an initial diagnostic method to find good regions in the parameter space as have been done in this thesis. But the higher resolution re-simulations of the good models could be performed including both the sub-structure dynamics and gas physics to account for star formation and gas loss/gain processes.

Another interesting application of our method would be to study, the evolution of the tidal streams that result from our models in more detail. In Chapter 5 we saw that the distribution of the velocities in the tails depends on the model. If we can predict the the number of stars and velocities that would be observed in different models, with the high precision data from GAIA we might be able to add these to our algorithm as another constraint and find the right dynamical models for dSphs.

Finally, we could use the algorithm to study the dynamical evolution of the Milky Way, where the simulations would be made find the models of the Milky Way that could evolve to match the distribution of the Globular Clusters and the dSphs around
Conclusions

6.1. Future plans

it.
References


Cayrel, R., 1988, IAUS, 132, 345


Dehnen, W., 2002, JCP, 179, 27.


Larson,D. 2010arXiv1001.4635L

176


177


Powel R., www.atlasoftheuniverse.com


Tonry J., Davis M., 1979, AJ, 84, 1511


Valenzuela, O., Rhee G., Klypin A.,


