A Bayesian/MCMC Approach to Galaxy Modelling: NGC 6503

by

David Puglielli

A thesis submitted to the
Department of Physics, Engineering Physics and Astronomy
in conformity with the requirements for
the degree of Doctor of Philosophy

Queen’s University
Kingston, Ontario, Canada
December 2009

Copyright © David Puglielli, December 2009
Abstract

We use Bayesian statistics and Markov chain Monte Carlo (MCMC) techniques to construct dynamical models for the spiral galaxy NGC 6503. The constraints include surface brightness profiles which display a Freeman Type II structure; HI and ionized gas rotation curves; the stellar rotation, which is nearly coincident with the ionized gas curve; and the line of sight stellar dispersion, which displays a $\sigma$–drop at the centre. The galaxy models consist of a Sérsic bulge, an exponential disc with an optional inner truncation and a cosmologically motivated dark halo. The Bayesian/MCMC technique yields the joint posterior probability distribution function for the input parameters, allowing constraints on model parameters such as the halo cusp strength, structural parameters for the disc and bulge, and mass-to-light ratios. We examine several interpretations of the data: the Type II surface brightness profile may be due to dust extinction, to an inner truncated disc or to a ring of bright stars; and we test separate fits to the gas and stellar rotation curves to determine if the gas traces the gravitational potential. We test each of these scenarios for bar stability, ruling out dust extinction. We also find that the gas cannot trace the gravitational potential, as the asymmetric drift is then too large to reproduce the stellar rotation. The disc is well fit by an inner-truncated profile, but the possibility of ring formation by a bar to reproduce the Type II profile is also a realistic model. We further find that the
halo must have a cuspy profile with $\gamma \gtrsim 1$; the bulge has a lower $M/L$ than the disc, suggesting a star forming component in the centre of the galaxy; and the bulge, as expected for this late type galaxy, has a low Sérsic index with $n_b \sim 1 - 2$, suggesting a formation history dominated by secular evolution.
Statement of Co-Authorship

This work has been carried out in conjunction with my supervisor, Larry Widrow, and Stéphane Courteau. Material in chapters 1, 4, 7, 8 and 9 is mainly taken from a paper that has been submitted to *The Astrophysical Journal*. For this research, I used galaxy models previously implemented by Larry in conjunction with John Dubinski and the $N$–body code from David Stiff. Unless otherwise noted, all work contained herein is the author’s.
Acknowledgements

This work is the culmination of seven years of research, and I would not have been able to complete it without the support of a number of people. First, my parents, whose love and support have been irreplaceable and who have always encouraged me to shoot for the stars (I went a little further and landed among the galaxies). I would like to thank my supervisor, Larry Widrow, for putting up with my meandering style of work and guiding me into finally producing a coherent piece of research. I thank Stéphane Courteau for many fruitful discussions and for pushing me to dig deeper into the literature and develop a greater understanding of my own research. I owe a debt of gratitude to Roelof Bottema, Kristine Spekkens, Peter Teuben, Victor Debattista and John Dubinski for a number of very useful and informative discussions. I must also thank Joseph MacMillan, who developed a framework for analyzing $N$–body simulations that I contributed to, for pushing me away from Fortran to learn C and for answering my many questions about Linux. Thanks also to David Stiff for his $N$–body code, Doug McNeil for being a great TA partner, Kelly Foyle for lending me the star formation version of GADGET and Pascal Elahi for many discussions that may or may not have had anything to do with this work. I’d also like to thank James Silvester for being a wonderful officemate and Nate and Jess for all the pies.

I cannot forget to thank Ricardo and Chuck at the counselling service and Drs.
Finch and McNevin in psychiatry for helping me when I felt that I could not continue because of OCD. I suspect I would have accomplished much more had I not waited so long to seek treatment, but I will gladly accept no longer spending 6+ hours per day on my obsessions.

My eternal gratefulness goes to a number of people who have kept me from going insane and allowed me to take, however hesitantly, my first steps emerging from that dark and lonely space. Among them are Justin, John, Cristina, Marc, Mark, Danny, Ash, Daro, Beau and Vinh. Special thanks go to Jason for being there from the beginning, to Jay for many wonderful, funny and illuminating conversations, and to Jon and Joshua for their insight. I am especially grateful to Johny, for teaching me what it means to be human, for pushing me outside my comfort zone and for a great many fascinating and delightful discussions. Finally, a huge thanks and cwtch to Terrance, for reasons far too numerous to list here. He is the most wonderful human being I know, and I dedicate this work to him.
# Table of Contents

Abstract i

Statement of Co-Authorship iii

Acknowledgements iv

Table of Contents vi

List of Tables ix

List of Figures x

Chapter 1: Introduction 1

Chapter 2: Galaxy Dynamics 6

2.1 Orbital dynamics 7

2.2 Dynamical models and $N$–body simulations 8

2.3 $N$–body methods 14

2.4 Asymmetric drift 16

Chapter 3: Structure of Disc Galaxies 18

3.1 Bulges and discs 19
3.2 The structure of halos ........................................ 25
3.3 Matter distribution in disc galaxies ......................... 30

Chapter 4: Our Lucky Specimen: NGC 6503 .................. 33
4.1 Observations ............................................... 33
4.2 Prior models for NGC 6503 ................................ 45

Chapter 5: GalactICS model ................................. 49
5.1 The GalactICS bulge and halo ............................ 50
5.2 The GalactICS disc ....................................... 51
5.3 The inner-truncated disc .................................. 53
5.4 Summary ................................................... 55

Chapter 6: Bayes’ Theorem and Markov chain Monte Carlo .... 57
6.1 Maximum likelihood ....................................... 58
6.2 Bayes’ theorem .......................................... 59
6.3 The likelihood function .................................. 62
6.4 Pros and cons of MCMC ................................. 64
6.5 Ensuring convergence ................................... 66
6.6 An improved MCMC algorithm ......................... 67
6.7 Our MCMC runs ......................................... 69

Chapter 7: Results: Constraining the Physics ............... 72
7.1 Fit quality ................................................. 73
7.2 Testing the rotation curve fits ............................. 78
7.3 Testing dust extinction ................................... 85
List of Tables

4.1 Summary of scenarios and data fit. ......................... 44

5.1 Model parameters and calculated quantities. .............. 56

6.1 Acceptance rates for each scenario. ....................... 71

7.1 Posterior means and errors for each parameter. .......... 74

8.1 $B/T$ for each scenario. .................................. 104
## List of Figures

3.1 Types of surface brightness profiles. ........................................... 20

4.1 NGC 6503 in the optical. ...................................................... 34

4.2 Surface brightness of NGC 6503. ............................................. 36

4.3 Log-log plot of the surface brightness. .................................... 37

4.4 Rotation curve for NGC 6503. ................................................. 39

4.5 Velocity dispersion for NGC 6503. .......................................... 42

7.1 Time series trace for three scenarios. ....................................... 75

7.2 Time series trace for the fourth scenario. ................................. 76

7.3 Autocorrelations for each scenario. ........................................ 76

7.4 Model rotation curve fits. ...................................................... 79

7.5 Model surface brightness fits. ................................................. 80

7.6 Model velocity dispersion fits. ................................................. 81

7.7 The disc mass–$Q$ PDF. ......................................................... 82

7.8 The $X – Q$ PDF. ............................................................... 83

7.9 Bar formation in each scenario. .............................................. 86

7.10 Surface brightness residuals. ............................................... 88

7.11 Bar strength versus radius. ................................................. 90
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1</td>
<td>The $R_d - R_\sigma$ PDF.</td>
<td>93</td>
</tr>
<tr>
<td>8.2</td>
<td>The $z_d$ PDF.</td>
<td>95</td>
</tr>
<tr>
<td>8.3</td>
<td>The $R_h - \alpha$ PDF.</td>
<td>96</td>
</tr>
<tr>
<td>8.4</td>
<td>The $n_b - r_b$ PDF.</td>
<td>97</td>
</tr>
<tr>
<td>8.5</td>
<td>The $\gamma$ PDF.</td>
<td>100</td>
</tr>
<tr>
<td>8.6</td>
<td>A model for which $\gamma = 0$.</td>
<td>101</td>
</tr>
<tr>
<td>8.7</td>
<td>The bulge mass--$(M/L_R)_b$ PDF.</td>
<td>102</td>
</tr>
<tr>
<td>8.8</td>
<td>The disc mass--$(M/L_R)_d$ PDF.</td>
<td>103</td>
</tr>
<tr>
<td>9.1</td>
<td>Velocity dispersion ratio for each scenario.</td>
<td>111</td>
</tr>
<tr>
<td>9.2</td>
<td>Asymmetric drift dependence on several factors.</td>
<td>115</td>
</tr>
<tr>
<td>9.3</td>
<td>Surface brightness in several bands for NGC 6503.</td>
<td>117</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

The joint analysis of galaxy surface brightness profiles and rotation curves has traditionally yielded rich insights into galaxy physics. Surface brightness profiles allow for bulge-disc decompositions, while extended rotation curves constrain mass models (van Albada et al., 1985; Palunas and Williams, 2000; de Blok et al., 2008; Kuzio de Naray et al., 2008). However, mass models are subject to degeneracies due in part to uncertainties in the mass-to-light ($M/L$) ratios of the baryonic components (e.g., Maller et al. 2000, Dutton et al. 2005), so the individual contributions from the disc, bulge and halo are uncertain. Breaking these degeneracies requires more data, such as line-of-sight (LOS) velocity dispersions or colour; this also allows the creation of dynamical models (Rix et al., 1997; Gebhardt et al., 2000; Widrow et al., 2003; Baes and Dejonghe, 2004; Thomas et al., 2007). These provide the initial conditions required to carry out $N$-body simulations, which are needed to test stability to bars and other non-axisymmetric perturbations. In addition, non-circular motions complicate the interpretation of gas rotation curves (Rhee et al., 2004; Valenzuela et al., 2007). Stellar rotation curve data may help address these complications. A data set
that includes surface brightness profiles, gas and stellar rotation curves and stellar velocity dispersions can, as we shall demonstrate, provide constraints on important galaxy parameters, along with information about stability against bar formation.

Such a data set exists for the isolated dwarf Sbc galaxy NGC 6503. Gas and stellar rotation curves were measured by de Vaucouleurs and Caulet (1982), Begeman (1987) and Bottema (1989) (hereafter B89); B89 also measured the stellar velocity dispersion profile and surface brightness profiles in the $B$– and $R$–bands. There are no strong asymmetries in galaxy structure (such as a bar), so it is a good candidate for analytic axisymmetric models. However, further investigation reveals several peculiarities. The galaxy possesses a sharp drop in velocity dispersion near the centre (known as a ‘$\sigma$–drop’). The gas and stellar rotation curves are nearly coincident despite the standard asymmetric drift formula’s prediction that stars should noticeably lag behind the gas. Finally, the surface brightness profile displays four distinct regions: a central peak, a flat region indicating a Freeman (1970) Type II profile, an inner exponential of scale length $\sim 1$ kpc and a shallower outer exponential.

In this paper, we present several scenarios corresponding to differing interpretations of the data. For each scenario, we construct dynamical models that reproduce these features. We decompose the surface brightness using a Sérsic bulge and an inner-truncated light profile of the type used by Kormendy (1977). Inner truncated disc profiles may be due to dust extinction against an exponential disc (MacArthur et al., 2003); they may be intrinsic to the density profile; or they may be due to a ring of bright stars which does not trace the mass. We explore these possibilities by testing models numerically under each scenario for bar formation. The coincidence of gas and stellar rotation curves also deserves attention as it suggests far less asymmetric
drift than predicted. Thus, we wish to understand whether gas follows circular orbits that trace the gravitational potential, or whether it has its own asymmetric drift. We investigate this issue by testing two different ways of fitting the rotation curves: by fitting the stellar rotation alone and using the asymmetric drift to calculate the circular velocity; and by fitting the gas alone assuming it traces the circular velocity.

Our aim is to construct dynamical models for each of these scenarios and test them via $N$-body simulations. We use the GalactICS model from Widrow et al. (2008) (hereafter WPD) defined in terms of distribution functions for the disc, bulge and halo. The halo allows for a cusp or a core, and the bulge follows a Sérsic (1968) profile. The disc is initially exponential, but we also modify it to allow for the possibility of an inner-truncation.

Finding a suitable set of models to fit the data is not a straightforward exercise because of the large number of model parameters. A number of techniques have been developed to determine best-fit parameters in such complex spaces – notably, maximum likelihood techniques which involve minimizing the $\chi^2$ value. One promising approach which improves on the maximum likelihood method employs Bayes’ theorem and a Markov chain Monte Carlo technique to survey the relevant parameter space. Bayes’ theorem provides a method of determining the posterior probability of a hypothesis based on both the likelihood function and prior information about the input parameters; MCMC provides for a way to survey the parameter space. In conjunction, the two techniques provide the joint posterior probability distribution function (PDF) of the multidimensional parameter space; marginalizing over any parameter provides the marginal posterior probability distribution function for it, yielding the
formal mean and error bars. This technique is very flexible and yields realistic constraints on the model parameters, including (but not limited to) the halo cusp, the disc and bulge mass-to-light ratios, mass estimates for each component, inner truncation parameters, and bulge Sérsic index. Similar techniques have been used, for example, to determine cosmological parameters (Tegmark et al., 2004; Percival, 2005; Corless and King, 2008).

In this work we build dynamical models for NGC 6503 using this method, obtaining a set of best-fitting model parameters. We fit several data sets for this galaxy simultaneously in order to obtain a relatively complete picture of the galaxy; however, the lack of a bar provides a further constraint. Therefore, we also obtain the stability parameters $X$ and $Q$ for our models. The $X$–parameter determines global stability to multi-armed modes (Toomre, 1981) and the $Q$–parameter determines stability locally (Toomre, 1964). As a purely phenomenological matter, $Q$ can also determine global stability to bars (Athanassoula and Sellwood, 1986). We test the stability of our models using $N$–body simulations. In this work, we focus on $Q$ to characterize global disc stability, because it is more widely used to parameterize stability in the literature and it uses the radial velocity dispersion, which is more accessible than the epicycle frequency.

Our approach allows us to investigate numerous other properties of the galaxy. The Sérsic index is a direct indicator of whether a bulge is a pseudobulge (Fisher and Drory, 2008), as pseudobulges nearly all have Sérsic indices below 2, while bulges that resemble small ellipticals tend to have higher Sérsic indices. We also explore the possibility that the bulge is related to luminous nuclear clusters found near the centres of many disc-dominated late-type galaxies, and derive $M/L_R$ ratios for the disc and
bulge. Most interestingly, we can constrain the cuspiness of the halo. Cosmological simulations consistently demonstrate that halos are expected to have cuspy density profiles (e.g., Navarro et al. 1997; Moore et al. 1999), but observations of the rotation curves of low surface brightness (LSB) galaxies were once thought to show cored halos, birthing the so-called core catastrophe (Blais-Ouellette et al., 2001; de Blok et al., 2001). However, there are considerable difficulties involved in inferring the cusp value from rotation curve measurements (Hayashi et al., 2004; Rhee et al., 2004; Valenzuela et al., 2007). For example, gas rotation curve data near the centre of the galaxy may not trace the gravitational potential; using the stellar rotational velocity, where available, may be a better way to infer the cusp value (Pizzella et al., 2008). Further, noncircular gas motions at the centre (as may be caused by a triaxial halo) may skew the interpretation of the rotation curve, depending on the orientation of the disc (Simon et al., 2005). With both gas and stellar rotation curves, NGC 6503 provides an excellent opportunity to investigate these issues.

The organization of this thesis is as follows. Chapters 2 and 3 discuss background information: galaxy dynamics in chapter 2 and the structure of disc galaxies in chapter 3. We discuss the observational properties of NGC 6503 in chapter 4, along with prior work done to model the galaxy. A description of the GalactICS model is found in chapter 5. In chapter 6, the reader will find background information on Bayesian statistics and MCMC, along with a description of our approach. We constrain the physics of the models in chapter 7 using stability studies, and in chapter 8 we discuss the other results we obtain. Discussions of several interesting matters are presented in chapter 9. Chapter 10 is the conclusion, which includes a description of future projects. The appendix contains a proof that MCMC populates the posterior PDFs.
Chapter 2

Galaxy Dynamics

Galaxies are composed of multiple components. The visible components are the disc and bulge, and there is an unseen component made up of dark matter in a halo surrounding the visible components. The primary constituents of the luminous matter are stars and gas; the primary components of the dark matter are unknown, but likely to be some kind of subatomic particle. The key characteristic of both stars and dark matter is that the constituent particles are collisionless, i.e., each particle moves as though it were in the mean potential generated by the system, so that nearby encounters are not important. The vast majority of the mass of disc galaxies is collisionless, so an equally accurate title for this chapter could have been *Dynamics of Collisionless Systems*.

In this chapter, we outline the fundamental concepts of galaxy dynamics, star orbits, dynamical models and $N$–body simulations. The primary reference for this material is the classic text by Binney and Tremaine, *Galactic Dynamics*. 
2.1 Orbital dynamics

The study of galaxy dynamics begins with the study of orbits. We first define the concept of an integral of motion: an integral of motion is a quantity which is preserved along orbits in a given potential. Formally, we have

\[ I[x(t_1), v(t_1)] = I[x(t_2), v(t_2)]. \]  

(2.1)

Integrals of motion govern the geometry of orbits in a given potential. In a static potential, the energy \( E = \frac{1}{2}v^2 + \Phi \) is an integral of motion, and in an axisymmetric potential there are at least two integrals: \( E \) and \( L_z \), the vertical component of the angular momentum.

Orbits in static spherical potentials, such as the bulge or halo, are quite simple. These orbits conserve energy and the three components of the angular momentum \( \mathbf{L} \), so any given orbit lies in a plane known as the orbital plane. The orbits of stars in such potentials follow rosette patterns, in which the radial and azimuthal coordinates of a star oscillate.

The orbits of the disc are well described by approximately circular, approximately planar star motions. The important quantities are defined in relation to deviations from circular, planar motion; we define \( x \equiv R - R_0 \), where \( R_0 \) is the radius of the guiding centre of the orbit and is constant. By expanding the potential in a Taylor series and truncating it, we then have

\[ \ddot{x} = -\kappa^2 x \]  

(2.2)

\[ \ddot{z} = -\nu^2 z \]  

(2.3)

where \( \kappa \) is the epicycle frequency and \( \nu \) is the vertical frequency. Thus, orbits in disc galaxies can be approximated as harmonic oscillations about a circle in the \( z = 0 \)
plane. This is known as the epicycle approximation, and it is crucial to modelling
disc galaxies. Eq. 2.3 admits a third integral of motion, given by $E_z = \dot{z}^2 + \nu^2 z^2$
the energy in the $z$–direction; in the limit of zero thickness, this integral is exact
but becomes approximate at nonzero thicknesses. Nonetheless, it remains a good
approximation in real disc galaxies and we will employ it in our modelling.

While the vertical frequency $\nu$ rarely shows up elsewhere and serves mainly to
motivate the interpretation of disc galaxy orbits, the epicycle frequency is fundamen-
tal, as it helps to determine both the tangential velocity dispersions and stability
properties of the disc. The epicycle frequency may be written as

$$\kappa^2 = \left( R \frac{d\Omega^2}{dR} + 4\Omega^2 \right)$$ (2.4)

where $\Omega$ is the angular frequency of star orbits. Thus, the rotation curve offers
a measure of the epicycle frequency. In practice, however, the derivative in this
expression may lead to large errors when working with experimental data, so this
expression must be used with caution.

### 2.2 Dynamical models and $N$–body simulations

The evolution of collisionless systems is governed by the collisionless Boltzmann equa-
tion (CBE), given by

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f - \nabla \Phi \cdot \frac{\partial f}{\partial \mathbf{v}} = 0$$ (2.5)

where $\mathbf{v}$ is the velocity field of the stars and $\Phi$ is the gravitational potential. $f$ is the
distribution function (DF), or the phase space density of stars; it is the number of
stars per unit of phase space. By integrating the CBE over all velocities, we obtain
the Jeans equations

\[ \frac{\partial \rho}{\partial t} + \frac{\partial (\rho \langle v_i \rangle)}{\partial x_i} = 0 \]  

\[ \rho \frac{\partial \langle v_j \rangle}{\partial t} + \rho \langle v_i \rangle \frac{\partial \langle v_j \rangle}{\partial x_i} = -\rho \frac{\partial \Phi}{\partial x_j} - \frac{\partial (\rho \sigma_{ij}^2)}{\partial x_i} \]

where \( \rho \) is the mass density of stars, the brackets indicate the mean of the \( i \)-th velocity component \( v_i \), and the summation convention is used. The last term includes a stress tensor \( \rho \sigma_{ij}^2 \), where \( \sigma_{ij} \) is the \( ij \)-th component of the velocity dispersion tensor and is given by \( \sigma_{ij}^2 = \langle v_i v_j \rangle - \langle v_i \rangle \langle v_j \rangle \). The Jeans equations govern the evolution of stellar systems, and are analogous to the continuity equation and Euler equation in fluid dynamics.

Evidently, the solution of the CBE cannot be determined analytically in most cases, because \( N \), the number of particles, is usually a very large number; hence the need to find approximate solutions using \( N \)-body simulations. This technique brings up a number of issues regarding how to generate the initial conditions for \( N \)-body simulations, and how to ensure that, once the system is evolved, the resulting evolution is an accurate approximation of the evolution of the corresponding real system. We first examine the former issue, and defer a discussion of \( N \)-body methods to the next section.

Construction of dynamical models requires a method to specify the velocities of the particles; \( N \)-body simulations depend on the reliability of the adopted dispersions. The Jeans equations are often solved to yield the velocity dispersions, as is done in Binney et al. (1990) and Cinzano and van der Marel (1994). Hernquist (1993) generalized this method by taking moments of the CBE. The resulting distributions are approximations, because the Jeans equations and the CBE typically do not yield analytical solutions. Consequently, \( N \)-body simulations need time to settle; in the
process, the prescribed densities and velocities are altered, making it difficult to accurately evolve models when a specific set of parameters is desired. This is a problem in early modelling of NGC 6503, as we will see in §4.2.

Hernquist’s approach leads to relatively small deviations from equilibrium, but it can be improved upon. A more precise method of generating full dynamical models that specifies the densities and velocities simultaneously is to use the phase space distribution functions for the system of interest. The benefit of running $N$–body simulations with initial conditions determined by DFs is that the initial conditions exactly satisfy the CBE, and therefore no settling is required.

### 2.2.1 Distribution functions in the literature

DFs for spherical systems are commonly available. Spherical systems admit four integrals of motion: the energy $E$ and the three components of the angular momentum $\mathbf{L}$. However, if the velocity dispersion is isotropic, that is, $\sigma_r = \sigma_\theta = \sigma_\phi$, then the DFs become functions of $E$ only; in spite of their simplicity, these are of great practical importance. Stellar polytropes satisfy $f(E) \propto (-E)^n$; these systems have densities identical to gaseous polytropes, for which $p \propto \rho^\gamma$, provided $\gamma = 1 - \frac{1}{n}$. A special case is the isothermal sphere, for which $n \to \infty$ and $f(E) \propto \exp(-E/\sigma^2)$. The density of the isothermal sphere is given by

$$\rho(r) = \frac{\rho_0}{1 + (r/r_c)^2}$$  \hspace{1cm} (2.8)

where $\rho_0$ and $r_c$ are a characteristic density and radius, respectively. In some cases, models require a DF from a given density; in such cases, the resulting DFs may not be analytic. Realistic bulges and halos require this technique, discussed in chapter 5.
Finding DFs for axisymmetric systems such as discs or flattened halos can be rather complicated, but it has been done; in fact, because of the complexity in finding DFs for discs, it is only recently that DF-based models have found a prominent place in galaxy modelling. Such DFs rely on at least two integrals of motion, $E$ and $L_z$. The DF for the Kuz'min-Toomre disc (Toomre, 1963), for which $\Sigma(R) \propto (R^2 + a^2)^{-3/2}$, has been investigated by Miyamoto (1971), Kalnajs (1976) and Athanassoula and Sellwood (1986). The DF for the Mestel (1963) disc, for which $\Sigma \propto R^{-1}$, is given in Binney & Tremaine as $f \propto L_z^t \exp(-E/\sigma^2)$ where $t$ is constant. DFs for the three-dimensional axisymmetric logarithmic potential of Binney (1981)

$$\Psi = -\frac{1}{2}v_0^2 \log \left( R_c^2 + R^2 + z^2/q^2 \right)$$

(2.9)

where $q$ is the axial ratio of the equipotentials, were obtained by Evans (1993); they too rely on two integrals of motion and are nothing more than sums of exponentials, so they are quite simple.

A planar DF for axisymmetric discs discovered by Shu (1969) is given by

$$F(E, L_z) = \frac{\Omega(R_c) \Sigma(R_c)}{\pi \kappa(R_c) \sigma_R^2(R_c)} \exp \left[ \frac{E_c(L_z) - E}{\sigma_R^2(R_c)} \right]$$

(2.10)

where the energy $E$ and angular momentum $L_z$ are integrals of motion, $R_c$ is the radius of a circular orbit with angular momentum $L_z$ and $E_c$ is the energy of such an orbit. Kuijken and Dubinski (1995) (hereafter KD) construct DFs for disc-bulge-halo models using the Evans DF for their bulges and halos, and extend Shu’s DF for planar discs to three dimensions with three integrals of motion (the DF is given in chapter 5). This is an application of the Jeans theorem, which states that any function of three isolating integrals of motion in a given potential will exactly solve the CBE (Binney and Tremaine, 2008). The KD models were modified by Widrow and Dubinski (2005)
and WPD with more realistic halos and bulges. These are known as the GalactICS models, and are described in detail in chapter 5.

### 2.2.2 Disc stability

Constructing dynamical models is useful precisely because it is possible to test the models for stability; fits to the observed data are constrained not just by the data but by the presence, or lack thereof, of an obvious bar. Bar stability is usually parametrized in terms of the $X$– and $Q$–parameters. The $X$–parameter determines global stability to multi-armed modes, and is given by

$$X(R) = \frac{\kappa^2 R}{2\pi m G \Sigma}$$  \hspace{1cm} (2.11)

where $\kappa$ is the epicycle frequency, $m$ is the number of arms ($m = 2$ for a bar) and $\Sigma$ is the surface density (Toomre, 1981). The $Q$–parameter determines stability locally, and is given by

$$Q(R) = \frac{\sigma_R \kappa}{3.36 G \Sigma}$$  \hspace{1cm} (2.12)

where $\sigma_R$ is the radial velocity dispersion (Toomre, 1964). $Q < 1$ implies local instability to axisymmetric perturbations, but as a purely phenomenological matter, $Q$ can also determine global stability to bars – Athanassoula and Sellwood (1986) find that $Q$–values of $2 - 2.5$ operate as a global cutoff to bar formation. The $X$–value has a less well-defined stability cutoff, however. Note that these parameters strictly apply only to thin discs, and discs with realistic scale heights may have lower stability cutoffs. In this work, we focus on $Q$ to characterize global disc stability because it is more widely used to parameterize stability in the literature and it uses the radial velocity dispersion, which is more accessible than the epicycle frequency (cf. Eq. 2.4 and comments following).
Several other empirical measures of bar stability exist in the literature; for example, Ostriker and Peebles (1973) developed a criterion based on the ratio of disc kinetic energy to total gravitational energy, and Efstathiou et al. (1982) determined a criterion based on the maximum rotational velocity. These studies provide important clues to the conditions under which bars are induced or suppressed; Ostriker and Peebles found that an added gravitational potential, in the form of a dark matter component, could inhibit bar formation, hence providing an early indication of the existence of dark halos; Efstathiou et al. corroborated this result. An additional potential in the form of a central bulge may also inhibit bar formation. Moreover, it has long been known that very warm discs (that is, discs with a relatively large amount of random motion, as evinced by a large $Q$–value) are not susceptible to global asymmetries (Hohl, 1971).

Unfortunately\(^1\), things are not that simple. Athanassoula (2002) finds that halos can actually induce bar formation, leading to very strong bars. The culprit is resonant interactions between the halo and disc particles – an effect that is absent when using a rigid halo potential. Thus, fully self-consistent simulations are required to thoroughly investigate bar stability, as sometimes even discs with massive halos are susceptible to the formation of bars. An additional complication is that no bar studies have been performed on inner-truncated discs, so the applicability of bar formation criteria such as those discussed above are in question when applied to our models.

\(^1\)Or fortunately, depending on one's point of view.
2.3 $N$–body methods

A wide variety of astrophysical problems are amenable to $N$–body simulations; correspondingly, a wide variety of $N$–body techniques have been developed to study these problems. Collisionless galaxy modelling lends itself to direct summation or treecode force calculations. In these methods, the gravitational force on each particle is calculated by summing the individual forces due to all other particles; in a treecode, the force calculation is accelerated by using a single term for blocks of distant particles, either by using the total block mass centred on the block’s centre of mass, or by using multipole expansions to account for the internal structure of these blocks (Barnes and Hut, 1986). Since this process must be repeated for $N$ particles, direct summation run times go as $O(N^2)$, so it is clearly infeasible for high–$N$ runs; treecodes reduce the run time to $O(N \log N)$, and are much more frequently used. The fastest force summation codes, however, are known fast multipole codes; they are a variant of treecodes whose run time is nearly $O(N)$ (Dehnen, 2000). We use this algorithm as implemented by Stiff (2003). Dehnen’s fast multipole algorithm exploits the fact that nearby particles obtain the same force from distant blocks; the idea is that, between two distant cells, one can obtain the multipole expansion of the potential (and hence the force) generated by the first cell at all positions in the second cell. Thus, the forces from the first block on the second and vice versa may be computed in one step. This process not only significantly speeds up the force calculation, but it also ensures conservation of momentum. The result is $O(N)$ scaling for large $N$. It is worth noting, however, that there is some confusion in the literature regarding how general this result is, as some implementations seem to run as $O(N \log N)$ (Capuzzo-Dolcetta and Miocchi, 1998); however, Stiff’s code is based on Dehnen’s algorithm and runs very
Collisionless codes are good at modelling galaxies because the actual number of particles is very high for any galaxy; thus, each star may be assumed to move in the mean potential generated by the other stars without being affected by two-body encounters. Because the simulated $N$ will usually be smaller than the real $N$, a softening length $\epsilon$ needs to be implemented to mitigate the effect of two-body encounters. However, collisionless codes are not good at modelling smaller systems, such as globular clusters, in which $N \sim 10^5$; over time, the effects of two-body encounters in such system do become important. These are known as collisional systems, because the collisions between nearby stars have an impact on the structure and evolution of the system. The time scale on which two-body encounters are expected to impact the resulting evolution is given by the relaxation time:

$$t_{\text{relax}} = \frac{N}{8 \ln N} t_{\text{cross}}.$$  \hspace{1cm} (2.13)

Here, $t_{\text{cross}} = R/v$ is the crossing time. The relaxation time measures the time taken for the trajectory of a star in a collisionless system to change significantly because of two-body encounters from that in a perfectly smooth potential. When $N \sim 10^6$, as is typical of many galaxy simulations today, $t_{\text{relax}} \sim 10$ Gyr, so our simulations are reliable if they are run for less than roughly 10 Gyr (which, indeed, they are). In a real galaxy where $N \gtrsim 10^9$, two-body encounters are essentially irrelevant. In globular clusters, however, $t_{\text{relax}}$ may be an order of magnitude or more smaller, so collisional effects are important in these systems over the course of the age of the universe.

If collisional systems are more complicated than collisionless systems, then dissipative systems – i.e., systems with gas – are the most complicated. Unlike star
orbits, gas orbits cannot cross; instead, they shock. They are also subject to a host of other effects, such as turbulence, radiative heating and cooling, etc. Modelling gas is therefore more complicated; one popular method is to treat gas as particles with a viscosity, as is done in smoothed particle hydrodynamics (SPH). There are many implementations of SPH; the most notable may be GADGET (Springel et al., 2001; Springel, 2005). To model a galaxy with gas, separate stellar and gas discs are required; the evolution of the stellar component is determined using the standard treecode approach, and SPH is used to determine the gaseous evolution. SPH lends itself to additional physics; besides heating and cooling, star formation may be included by establishing a density threshold for the formation of star particles out of the gas, and supernova feedback can also be implemented. Such effects are undoubtedly necessary for a full understanding of galaxy evolution; for analyzing specific phenomena such as bar formation, they are not always needed, so we confine ourselves to collisionless simulations to determine if discs are stable to bars. However, correctly reproducing the $\sigma-$drop cannot be done with collisionless simulations (Bottema and Gerritsen, 1997); full physics simulations are required (Wozniak et al., 2003). More details are provided in §4.1.3.

2.4 Asymmetric drift

One particularly important application of the Jeans equations is the calculation of the asymmetric drift of a rotating disc. Asymmetric drift is the difference between the mean tangential velocity of the stars and the circular velocity implied by the potential, and it occurs because the stars have nonzero velocity dispersion. The magnitude of the asymmetric drift may be estimated, under a few assumptions, using the Jeans
equations in cylindrical coordinates. We assume a steady state, axisymmetric disc. At \( z = 0 \), the relevant equation reads

\[
\frac{R}{\rho} \frac{\partial (\rho \langle v^2_R \rangle)}{\partial R} + R \frac{\partial \langle v_R v_z \rangle}{\partial z} + \langle v^2_\phi \rangle - R \frac{\partial \Phi}{\partial R} = 0
\] (2.14)

where \( \langle v_\phi \rangle \) is the mean rotational velocity of the stars. The circular velocity is \( v^2_c = R (\partial \Phi / \partial R) \), the asymmetric drift is defined by \( v_a \equiv v_c - \langle v_\phi \rangle \) and we have that \( v^2_c - \langle v_\phi \rangle^2 = v_a (2 v_c - v_a) \simeq 2 v_a v_c \). The value of \( \partial \langle v_R v_z \rangle \)/\( \partial z \) depends on the orientation of the velocity ellipsoid; the two extreme cases are cylindrical alignment and spherical alignment. The velocity ellipsoid is said to be cylindrically aligned if the principal components of the velocity dispersion tensor \( \sigma_{ij} \) are aligned parallel to the cylindrical coordinate axes; in this case, \( \partial \langle v_R v_z \rangle \)/\( \partial z = 0 \). If the velocity ellipsoid is spherically aligned, we have \( \partial \langle v_R v_z \rangle \)/\( \partial z \simeq (\langle v^2_R \rangle - \langle v^2_z \rangle)/(z/R) \). Thus, the asymmetric drift equation simplifies to

\[
v_a = \frac{\langle v^2_R \rangle}{2 v_c} \left[ \frac{\sigma^2_\phi}{\langle v^2_R \rangle} - 2 \frac{\partial \ln \rho}{\partial \ln R} + \frac{1}{2} \left( \frac{\langle v^2_z \rangle}{\langle v^2_R \rangle} \pm \frac{\langle v^2_z \rangle - 1}{\langle v^2_R \rangle} \right) - \frac{3}{2} \right]
\] (2.15)

(cf. Eq. 4.228 of Binney & Tremaine 2008), where the sign ambiguity captures the range of possible orientations of the velocity ellipsoid.
Chapter 3

Structure of Disc Galaxies

Let us now examine the structure of disc galaxies; we motivate these discussions by focussing on their observational properties. Observations of galaxies start with surface brightness profiles, in which the light from a galaxy is broken up into elliptical annuli and integrated in each annulus to provide the azimuthally averaged surface brightness profile. Surface brightnesses are defined in terms of magnitudes per unit area on the sky. Magnitudes, in turn, are logarithms of the actual light emanating from a celestial source. Typically, individual observations are restricted to a specific range of wavelengths, so surface brightnesses are limited to specific bands, labelled as $U$, $B$, $V$, $R$, $I$ and so forth; $B$ and $V$ are the optical bands. Thus, a surface brightness profile is essentially a log plot of the light coming from the galaxy as a function of radius.

The other observational driver of galaxy structure investigations is the rotation curves, in which the tangential velocity of some kind of ionized gas is obtained along the line of sight by locating emission peaks in the spectrum, providing the velocity of the gas in the galaxy as a function of radius. This technique is known as emission
line kinematics, and a number of tracers are used for this purpose, most notably hydrogen–α (Hα), hydrogen–β (Hβ), nitrogen II (NII), oxygen III ([OIII]) and CO. Each of these is a specific atomic or molecular transition, and they are driven by temperature and density. In particular, Hα and NII arise out of hot gas in star-forming regions. Also used for rotation curve observations (especially of the outer regions of disc galaxies) is neutral hydrogen (HI). These observations use the 21 cm line to trace the rotation. HI is frequently used to obtain the rotation curve beyond the optical radius because there are no star-forming regions here to provide ionized gas kinematics, but it tends to suffer from beam smearing in the inner regions, limiting its usefulness inside the optical radius.

3.1 Bulges and discs

To analyse the structure of disc galaxies, surface brightness profiles are typically decomposed into disc and bulge components. There are two basic types of surface brightness profiles (Freeman, 1970): Type I profiles, in which there are two components on the magnitude plot, with the inner component at a higher slope; and Type II profiles, in which the two components are separated by a flatter portion in between; the ‘break’ between the flat portion and the outer component may be located in the inner or outer part of the disc. A Type III profile, in which a shallower exponential follows outside the middle exponential attributed to the disc, is sometimes employed as an additional category. Fig. 3.1, taken from Pohlen et al. (2008), shows an example of each profile, with an outer Type II break. (The break in NGC 6503 is an inner break; the figure is for edification only.) Traditionally, the inner components have
CHAPTER 3. STRUCTURE OF DISC GALAXIES

Figure 3.1: Type I surface brightness profile on the left, Type II in the centre and Type III on the right.

been fitted using the de Vaucouleurs (1948) law, also known as the $R^{1/4}$ law:

\[ I(r) = I_0 \exp \left(-b(r/r_e)^{1/4}\right) \]  \hspace{1cm} (3.1)

where $r_e$ is the half-light radius, $b = 7.67$ is a constant and $I$ is the surface brightness. The de Vaucouleurs law fits early-type galaxies well – these galaxies tend to have an evolutionary history dominated by mergers. The Sérsic profile (Sérsic, 1968; Ciotti, 1991) is a generalization of the de Vaucouleurs law for arbitrary exponent:

\[ I(r) = I_0 \exp \left(-b(r/r_e)^{1/n}\right) \]  \hspace{1cm} (3.2)

where $I_0$ is the central intensity, $n$ is the index and $b$ is now dependent on $n$. The Sérsic law offers a better fit to bulges in late type galaxies, for which $n \sim 1 - 2$ rather than 4. Note that $n = 1$ is an exponential profile.

Why is the distinction among bulge fits interesting? Bulges are distinguished on the basis of structural and dynamical parameters, which in turn may be used to constrain their formation histories. Classical bulges are dynamically hot and strongly resemble elliptical galaxies. However, there also exist pseudobulges, which tend to be dynamically cold (i.e., rotationally dominated); their morphology often resembles
that of disc galaxies (Kormendy, 1993); and their surface brightnesses are well approximated by pure exponentials (Andredakis and Sanders, 1994; MacArthur et al., 2003), a result that robustly correlates with late type galaxies (Courteau et al., 1996). They also tend to display nuclear bars, rings or spirals (Carollo et al., 1997), while classical bulges are relatively free of such features.

The flatness of bulges increases with Hubble type, suggesting flatness as a possible diagnostic for pseudobulges (Kormendy and Kennicutt, 2004). However, a better diagnostic may be the Sérsic index itself, as Fisher and Drory (2008) find that Sérsic indices are bimodal about $n = 2$, and that pseudobulges occupy the $n < 2$ side of the distribution. A considerable body of research links changes in the Sérsic index to numerous other galaxy parameters, both bulge parameters and parameters of the discs that host them; for example, Sérsic index has been found to correlate with surface brightness and effective radius of the bulge (Graham et al., 1996; Khosroshahi et al., 2000). Andredakis et al. (1995) find that Sérsic index correlates with Hubble type, and Drory and Fisher (2007) also find that blue galaxies tend to have pseudobulges while red galaxies tend to have classical bulges, further confirming a correlation with Hubble type. Moreover, there exist strong correlations for stellar colours, stellar populations, and scale lengths between the bulges and discs of late type galaxies (Courteau et al., 1996; Peletier and Balcells, 1996; MacArthur et al., 2003; Fisher, 2006).

These correlations strongly suggest that bulges and pseudobulges have distinct formation histories, and that pseudobulge formation is linked to the properties of the host discs. Classical bulges likely form out of violent merging processes in the early
universe; features such as spirals and rings are destroyed by this process, so the formation process for pseudobulges cannot involve any major mergers. The alternative is secular evolution, in which slow processes internal to the galaxy generate the bulge. These processes rely on nonaxisymmetric features, especially bars, to redistribute energy and angular momentum, in the process funnelling gas inward where star formation can then occur, forming a bulge-like structure. This scenario is supported by observations of late type galaxies in which barred galaxies are seen to have larger gas concentrations and higher star formation rates in the centre (Sheth et al., 2005; Fisher, 2006).

Such a formation pattern is seen in hydrodynamical simulations of disc galaxies, and the resulting surface brightness profiles are well fit by double exponential profiles (Mayer and Wadsley, 2004; Debattista et al., 2006).

Meanwhile, the disc surface brightness profile is taken to be an exponential, as, crudely speaking, nearly all disc galaxy surface brightness profiles display linearly on a log plot over a large range in radius. We therefore have

\[ I(R) = I_0 \exp\left(-\frac{R}{R_d}\right) \]  

where \( R_d \) is the photometric disc scale length. Thus, bulge-disc decompositions often employ Sérsic or de Vaucouleurs profiles for the bulge and exponential profiles for the disc.

The tendency of discs to assemble into exponential profiles arises out of analytical models of structure formation in the early universe (Lin and Pringle, 1987; Dalcanton et al., 1997; Firmani and Avila-Reese, 2000; Ferguson and Clarke, 2001; Dutton, 2009). In these models, the baryons and primordial halos are assumed to be tidally torqued by their neighbours, giving rise to nonzero angular momentum; the baryons then cool and collapse into a disc while mostly conserving their initial angular momentum. It
appears that viscous processes play a key role, as the interplay of viscous evolution of
the disc and star formation leads naturally to exponential discs, provided the timescale
for angular momentum transport is large compared to that for collapse, and that the
timescales for viscous evolution and star formation are comparable. However, the
models still suffer from defects, as the exponential nature of these discs appears
over a smaller range in radius than actually observed, and their central densities are
larger than for pure exponentials. Taking account of more physical effects such as
supernova-driven galactic outflows improves the situation somewhat (Dutton, 2009),
reproducing exponential discs but the central cuspiness of discs remains.\footnote{Taken
at face value, this suggests that some ‘bulges’ seen in surface brightness profiles may
simply be part of the disc.}

\section*{3.1.1 Inner truncated profiles}

A subset of disc galaxy surface brightness profiles display flattened regions just inside
the linear disc part that make this conventional fitting technique inappropriate –
these are Freeman Type II profiles. Kormendy (1977) addressed this by truncating
the inner part of the disc. The inner truncated profile is given by

\begin{equation}
I(R) = I_0 \exp\left[-(R/R_d + (R/R_h)^\alpha)\right]
\end{equation}

where $R_h$ is the turnover radius and $\alpha$ is the cutoff index. Kormendy (1977) finds
that $\alpha = 3$ provides a good fit, but this result does not appear to be rigourously
established. Kormendy argued that the expression is merely a mathematical device
to improve the fits to his galaxies; Baggett et al. (1998) and Anderson et al. (2004)
also find that the Kormendy profile provides a better fit to their sample of galaxy
surface brightness profiles, but further argue that this reflects a real structural change
in the central parts of such galaxies. In 80% of cases, their truncated galaxies also exhibit a bar, but about 10% of their sample of non-barred galaxies also exhibit inner truncations, the reasons for which are not entirely clear.

How do we know the truncations are a real part of the surface brightness profile? We can eliminate a few possibilities. They are certainly not measurement artifacts, as Anderson et al. find that there is no correlation between truncations and Hubble-type morphological features, nor does the presence of truncations depend on the bulge fitting function. They are not likely to be due to dust extinction either, as there are no dust lanes observed in galaxies with Type II profiles, and multiband photometry of Type II galaxies reveals little change across different bands, contrary to the expected changes in the case of dust extinction (Héraudeau and Simien, 1996). There are also kinematic indications in some rotation curves of the existence of truncations (Einasto et al., 1980; Sofue et al., 1999), negating the dust extinction possibility. Thus, disc truncations indicate a genuine dearth of disc stars.

The causes of inner-truncated discs are up for debate. Freeman (1970) suggested that protogalactic material in the centre of a collapsing disc may predominantly form a bulge population, leaving a depleted or nonexistent disc in the centre; Pfenniger and Norman (1990) found that interactions between the bulge and disc could cause disc to mix with the bulge and obtain the dynamical characteristics of the bulge, leaving the appearance of a depleted central disc. However, given the results of Anderson et al., it seems clear that bars are a major factor. While bars can cause a redistribution of disc stars that can produce an inner-truncated surface brightness profile (especially when the bar is oriented perpendicularly to the galaxy’s axis of inclination), the existence of unbarred inner-truncated discs and nontruncated barred systems complicates the
analysis and indicates that the connection between truncations and bars is certainly not one-to-one.

If there is a connection between barred and non-barred inner-truncated profiles, the obvious question is whether bars can be destroyed. It was once thought that central mass concentrations (such as supermassive black holes (SMBH) or star clusters) at the centres of barred galaxies could weaken and eventually destroy bars by orbit scattering or altering the phase space structure of bar orbits (Norman et al., 1996). This introduces a scenario in which bar formation is a recurrent process, as most galaxies are observed to have SMBHs at their centres, and the possibility that surface brightness profiles of galaxies with destroyed bars may retain the structural properties of bars. However, Shen and Sellwood (2004) have shown that SMBHs (and other types of central mass concentrations) cannot destroy bars unless they are much more massive than currently observed. There is another method by which bars could be destroyed: Bournaud et al. (2005) find that bars cause gas to funnel into the centre of the galaxy, which in turn can weaken and destroy bars by angular momentum redistribution. The result is often a noticeable ring, which would cause flattening in the surface brightness profile. Thus, destroyed bars may leave signatures of their onetime presence in this fashion, creating an inner-truncated surface brightness profile.

3.2 The structure of halos

There are several reasons to suspect that galaxies possess a component consisting primarily of unseen dark matter (and that this component also permeates the entire universe). One famously important indicator is the existence of flat rotation curves of
disc galaxies to distances far exceeding their optical limits\(^2\) (e.g., Rubin et al. 1980; Rubin et al. 1985; Persic and Salucci 1995; Persic et al. 1996). Flat rotation curves are not universal – in low surface brightness galaxies, the rotation curves often appear to be rising to the outermost observed point (de Blok et al., 1996; Swaters et al., 2009). In the absence of dark matter, rotation curves would decline as \(r^{-1/2}\) outside the optical radius, so a dark matter component is postulated to resolve the discrepancy. Rotation curves are therefore decomposed into contributions from the disc, bulge and halo (and, sometimes, a nucleus or supermassive black hole). However, because the contribution of the luminous components to a given rotation curve is uncertain, the density profile of the halo is not easily gleaned, so it must be either guessed or obtained from simulations of structure formation.

Halos were once often modelled as cored structures such as the isothermal sphere (Begeman et al., 1991), given by Eq. 2.8. Note that the central density of the isothermal sphere converges; this is known as a core. The motivation behind this choice is that, in a chaotic dark matter collapse scenario, the conditions for violent relaxation are satisfied, and the resulting equilibrium is the isothermal sphere (Lynden-Bell, 1967). However, cosmological simulations suggest a different profile, such that \(\rho \propto r^{-\gamma}\) as \(r \to 0\), and that \(\rho \propto r^{-3}\) as \(r \to \infty\) (Navarro et al. 1997, Moore et al. 1999). The density profiles are typically fitted by expressions of the type

\[
\rho(r) = \frac{\rho_0}{r^\gamma(1 + r/r_0)^{3-\gamma}}
\]

where \(\rho_0\) and \(r_0\) are the scale density and radius, respectively. The density divergence at the centre is known as a cusp.

\(^2\)Of course, there exist other explanations for flat rotation curves such modified Newtonian dynamics (MOND), which posits a change to the law of gravity to account for flat rotation curves; in this work we simply assume the the dark matter hypothesis is valid.
In their seminal work, Navarro et al. (1997) (hereafter NFW) find that this density profile is universal, and that $\gamma = 1$. Both of these assertions have been challenged and are constantly being refined, especially as the resolution of later cosmological simulations has improved. The assumption of universality was found to be invalid by Jing & Suto (2000). More recently, Navarro et al. (2008) find that the profiles for cosmological halos exhibit subtle changes from universality that cannot be captured by a two-parameter fitting formula such as Eq. 3.5. They use the Einasto (1965) formula, which contains a third shape parameter to fit their halos:

$$\ln\left(\frac{\rho(r)}{\rho_{-2}}\right) = (-2/\alpha) \left[\left(\frac{r}{r_{-2}}\right)^{\alpha} - 1\right]$$

where $r_{-2}$ locates the maximum of the $r^{2}\rho$ profile, $\rho_{-2} = \rho(r_{-2})$ and $\alpha$ is a shape parameter. For the purposes of this work, the small deviations from universality are not expected to have a significant effect on our models.

The cusp value is a more interesting matter. Different researchers all agree that there is a cusp, but not on the number. Moore et al. (1999) find that $\gamma = 1.5$. Other recent simulations also show cusp values larger than 1 (e.g., Diemand et al. 2005), but Navarro et al. (2008) use very high resolution simulations to show that the cusp value is closer to $\gamma = 0.9 \pm 0.1$. The latter authors also find that the cusp value is not constant, and becomes shallower as $r \to 0$.

While all cosmological simulations are in agreement that dark matter halos tend to display cusped centres, the observational situation remains unsettled. A great deal of work has focused on fitting the inner rotation curves of low surface brightness (LSB) galaxies with cored and cusped profiles to determine which provide better fits. Much research (Flores and Primack, 1994; Blais-Ouellette et al., 2001; de Blok et al., 2008) has found that cored profiles offer better fits to the observed rotation curves,
although even here doubts have been raised about the quality of the fits, and some work suggests that both cored and cusped halos may be consistent with the observed data (Swaters et al., 2003; Simon et al., 2005; Spekkens et al., 2005). More recent work has focused on the problems inherent in the data collected for this purpose and in physical effects that can impact the results in important ways.

The relevant issues with observed rotation curves are related to measurement problems and a failure to account for the physical effects that limit the applicability of the models used to judge the fits. Inevitably, certain assumptions must creep into the procedures used to derive density profiles from rotation curves; the resolution to the controversy almost certainly involves determining how valid these assumptions are. To wit, here are the assumptions made by de Blok et al. (2008) when deriving their density profiles: the LSB galaxies are purely dark matter dominated; the gas rotation is circular; and the galaxies are symmetric (with the dark matter halos being spherically symmetric). In addition, although not stated, they assume that the gas traces the gravitational potential, as they use Poisson’s equation directly to solve for the density profile.

The first of these has been confirmed to not significantly affect the resulting derived density profiles; even if it does, providing for a significant disc mass would lower the halo mass interior to any given radius, and therefore the cusp value as well. It is less clear that the other assumptions are valid, each of which has undergone intense scrutiny in recent years. The impact of noncircular gas motions is not well understood, but it is known that simulated dark matter halos tend to be prolate or triaxial rather than spherical. Thus, it is possible that a triaxial potential would generate noncircular gas motions; Oh et al. (2008) find that the magnitude of noncircular gas motions is
CHAPTER 3. STRUCTURE OF DISC GALAXIES

not large enough to resolve the discrepancy between observed cores and theoretical cusps. However, triaxial potentials would not be the only source of noncircular gas motions; in massive discs, the growth of spiral or bar modes are also likely to generate noncircular gas motions, and relying on observations of gas motions in these galaxies is questionable. Of course, in LSB galaxies, the halo dominates and this is not likely to be an issue because LSB discs are rather light. However, there is a complication in that studies of halo evolution in the presence of baryons tend to favour halos that become nearly oblate (Gnedin et al., 2004; Abadi et al., 2009), so the possible sources of noncircular motions in halo dominated galaxies are not so clear.

Simon et al. (2005) emphasize another problem with conventional dark matter profile fitting: by using only two candidate profiles (pseudoisothermal and NFW), some important subtleties are missed. In particular, the possibility that three parameter fitting formulae (e.g., the Einasto profile) are better than the two-parameter NFW profile is not addressed by early work in which the NFW fits to observed rotation curves are inferior, and Simon et al. (2005) show that a three parameter fit from cuspy halos do a better job of fitting the data. Their fits employ ‘hybrid’ profiles of the type in Eq. 3.5 with $\gamma = 0.5$, finding that they provide better fits than NFW as well, suggesting that shallow cusps may be likely in many halos.

The assumption that gas rotation traces the gravitational potential may be the most problematic. The derived value of the cusp is highly sensitive to the slope of the rotation curve used to derive it, and the possibility that gas does not trace the potential will impact the slope of the rotation curve near the centre. Systematic studies of gas deviations from circularity have only recently been undertaken; Pizzella et al. (2008) measure gas and stellar rotation curves for 6 LSB galaxies to examine
this issue. They find that the gas rotation suffers from numerous issues, including non-circular motions and warps. By contrast, the star rotation is much more well-behaved, making it a more reliable indicator of halo dynamics and easier to model. In NGC 6503 itself, the gas speed drops below that of the stars at about 1.5 disc scale lengths. B89 speculates that this is due to distortion by a spiral arm; this may be plausible if spiral arms cause the gas to dissipate rotational energy. Thus, we will be testing both modelling options here, as cusp values from stellar rotation curves have received little investigation up to now.

### 3.3 Matter distribution in disc galaxies

While rotation curves can be decomposed into multiple components, as noted, typically a large range of $M/L$ values may fit a given rotation curve. A decades-old debate concerns whether or not discs are ‘maximal’, that is, whether or not discs are as massive as possible, consistent with the rotation curves of spiral galaxies. The rotation curve $v_c$ for a lone exponential disc is given by (Freeman, 1970)

$$v_c^2(R) = 4\pi \Sigma_0 R_d y^2 [I_0(y) K_0(y) - I_1(y) K_1(y)]$$  \hspace{1cm} (3.7)

where $\Sigma_0$ is the central surface density, $R_d$ is the disc scale length, $I$ and $K$ refer to the modified Bessel functions of the first and second kind, and $y = R/2R_d$. This equation is a maximum at $R = 2.2R_d$, and maximal discs typically yield $v_{\text{disc}}/v_{\text{tot}} = 0.85 \pm 0.10$ at this radius (Sackett, 1997).

The maximal disc hypothesis is put forward by van Albada et al. (1985), who measure the flat HI rotation curve for NGC 3198, an inclined nearby Sc galaxy, to 11 disc scale lengths. Their argument is twofold: first, the shape of the rising part of the
rotation curve matches what is expected from the disc light distribution, suggesting that the disc dominates the mass in the inner portion and hence that halos only dominate in the outer portions; and second, the small scatter of the Tully-Fisher relation suggests that the luminous matter and not the dark component determines the maximum rotation. The maximal disc hypothesis is attractive because it suggests a ‘rule’ for how galaxies assemble that limits the available free parameters, and may imply a consistent formation scenario; otherwise, different galaxies could, in principle, have widely differing disc to halo mass ratios, with no consistent way to determine what the ratio is in any given galaxy. The maximum disc hypothesis is therefore useful for developing mass models of disc galaxies without needing to fine tune parameters for proper fits; in a sense, this is an application of Occam’s Razor.

An attractive hypothesis, unfortunately, is not necessarily a correct hypothesis, and the current consensus is that the maximal disc hypothesis is probably invalid for most galaxies (although some lines of evidence appear to conflict). As noted in §2.2.2, lone discs tend to be unstable, and one way to stabilize them is to add a dark component. While a large fraction of disc galaxies are barred, the presence of unbarred galaxies militates against the maximal disc hypothesis because lone discs, or disc dominated galaxies, tend to be strongly unstable to bars. Bottema & Gerritsen (1997) found this to be the case for NGC 6503, by evolving a maximal disc and showing that a bar promptly forms; Fuchs (1999) also argues against a maximal disc in NGC 6503 because such a disc would require a $Q$–value less than 1. This does not imply that all discs are submaximal, of course, but it does point to a possible lack of a general rule governing the mass distribution in disc galaxies. However, massive halos may, in fact, promote bar growth thanks to resonant interactions with the disc
Thus, even barred galaxies may be halo-dominated inside the optical radius, and it is entirely plausible that barredness cannot be used to diagnose maximal disc systems.

Another (clever) argument against maximal discs comes from Courteau and Rix (1999), who point out that the residuals of the Tully-Fisher relation may be used to constrain the relative contribution of the disc to the rotation curve. For an exponential disc with no halo, we have $\partial V^2 / \partial R_d = -0.5$ from Eq. 3.7. This can be compared to the correlations between residuals of the Tully-Fisher, colour magnitude, and size luminosity relations, which yield $\partial V^2 / \partial R_d \sim -0.15 \pm 0.15$, which is not consistent with a lone disc. By calculating rotation curves of model galaxies with exponential discs and NFW halos, they find $\partial V^2 / \partial R_d$ as a function of $v_{\text{disc}} / v_{\text{tot}}$ at $2.2 R_d$. This constrains the disc contribution to $0.6 \pm 0.1$, substantially lower than required for a maximal disc, and thus these authors find that, on average, nonbarred HSB galaxies cannot possess maximal discs.

Other lines of evidence, however, suggest that discs are adequately fit by maximal discs, as Kent (1986) and Palunas and Williams (2000) find that dark matter is unneeded to fit many rotation curves within the optical radius. Additionally, constraints on halo masses may be obtained from the observed pattern speeds of bars. These speeds are generally quite large (Gerssen et al., 2003), whereas bars embedded in dark halos are expected to slow down by dynamical friction (Debattista and Sellwood, 2000), suggesting that halos of barred galaxies have low density in the centre.
Chapter 4

Our Lucky Specimen: NGC 6503

4.1 Observations

NGC 6503 is an isolated dwarf Sc galaxy at a distance of 5.2 Mpc (Karachentsev and Sharina, 1997) and is inclined at 74°. By isolated is meant that it has no tidal companions (apart from a possible faint nearby dwarf), and by dwarf is meant that both its size and luminosity are small compared to typical galaxies of its class. It displays a small, bright central bulge and no signs of strong asymmetric central structure (e.g., no obvious bar) apart from some spirality; it is therefore an ideal candidate for full dynamical modelling. An image of NGC 6503 in the optical is found in Fig. 4.1. In this chapter we review photometric and kinematic observations of the galaxy and highlight several key issues that arise. Throughout this document, we fix the distance at 5.2 Mpc and convert all angle measurements to lengths using this distance; we therefore have that 1 kpc = 40′′ at this distance.
Figure 4.1: NGC 6503 in the optical. By Liverpool Telescope/JMU.
4.1.1 The surface brightness profile

Photometric imaging by B89 provides one-dimensional surface brightness profiles in the $B -$ and $R -$ bands, both of which display a sharp rise within 50 pc and a Freeman (1970) Type II hump between 50 pc and $\sim 2$ kpc. Beyond that is an exponential profile. The central flattening is accompanied by a slight reddening in the $B - R$ profile (Fig. 4.2). A circumnuclear Hα ring is observed, indicating a star-forming region (Knapen et al., 2006). The outer exponential is remarkably straight in spite of the large error bars, and a regression line fit to those points yields a projected photometric scale length of 1.18 kpc. To correct for inclination and scale height, we manually adjust the intrinsic photometric scale length to reproduce the projected photometric scale length, finding a value of 1.3 kpc. This outer profile extends to several kpc (Héraudeau and Simien, 1996).

From Eq. 3.2, we see that the inner slope of a plot of the log of the surface brightness versus the log of the radius yields a value for the Sérsic index $n$. Fig. 4.3 shows the log-log plot of the surface brightness. The inner few points are almost straight, approaching a line with a slope $\sim 1$, suggesting a Sérsic index of $\sim 1$.

These surface brightness properties invite several interpretations. There are three possibilities for the origin of Type II profiles: first, they may be caused by dust extinction; second, they may reflect the formation of bright stars that do not trace the mass; and third, they may be intrinsic to the mass distribution of the disc. In NGC 6503, the reddening in the inner portion may support the dust extinction hypothesis (Bottema and Gerritsen, 1997) (hereafter BG97). BG97 adopt an exponential disc and assume the flattening in the surface brightness is due to dust (see §4.2 for a discussion of their work). As Fig. 4.1 shows, dust is indeed visible in NGC 6503. The
Figure 4.2: Surface brightness and colour profile from Bottema (1989). The $R$–band surface brightness data is in red and the $B$–band surface brightness data is in blue in the top panel, and the colour $B - R$ is in the bottom panel. Error bars smaller than the symbols are not shown.
Figure 4.3: Log-log plot of the surface brightness. The green line is determined by the slope of the inner two points and has a slope of \( \sim 1 \). The black curve is the exponential fit found in Fig. 4.2. The central magnitude \( \mu_0 \) is estimated from the slope of the inner points in Fig. 4.2 to be \( 17.6 \) mag arcsec\(^{-2} \).

The second possibility relates to bar formation. Baggett et al. (1998) and Anderson et al. (2004) show that many galaxies with Type II profiles can be fit using Kormendy’s (1977) inner truncated profile, Eq. 3.4. Most of these galaxies have a bar, but a significant minority do not. These authors argue that the surface brightness flattening is intrinsic to the galaxy rather than being caused by dust. It is known
that bars can induce ring formation in galaxies, which may contribute to the Type II profile flattening, and it is also known that circumnuclear rings correlate strongly with barred galaxies (Knapen, 2005). Moreover, simulations by Foyle et al. (2008) suggest that the flattening effect could also be reproduced by the existence of a central bar, as the resulting reorganization of light can flatten the surface brightness profile over the bar’s extent. Furthermore, while NGC 6503 does not appear to have a bar, it may be possible for bars to dissolve by gas inflow, which transfers angular momentum from the gas to the bar (Bournaud et al., 2005). The ring currently seen in NGC 6503 may therefore be a relic of a now dissolved bar. In a scenario of ring formation by bar destruction, the outer exponential represents the ‘true’ disc, and the Type II hump is mainly due to the formation of bright stars that do not trace the mass; Foyle et al. (2008)’s results suggest that the Type II profile could be reproduced by the emergence of a bar in a disc of scale length 1.18 kpc. Therefore, it is incumbent to investigate the surface brightness profile by fitting only the points exterior to 2 kpc.

The third possibility is that the Type II profile is caused by a genuine inner mass truncation, and therefore we also consider a model that fits the full surface brightness profile with an inner truncated disc. This will produce a better fit to the hump than to the outer points; along with the first scenario, this scenario provide a direct comparison with the modelling of BG97, who did not account for the outer exponential. Thus, we test three different ways to fit the surface brightness: we fit the whole surface brightness by assuming an inner truncated disc; we assume an exponential disc with an inner truncated light profile to model simply the effect of dust extinction (a full dust model is outside the scope of this work); and we fit only the inner ($R \lesssim 0.1$ kpc) and outer ($R > 2$ kpc) points to account for the true disc
Figure 4.4: Rotation curve data for NGC 6503. The Hβ data are filled red circles, the stellar rotation data are filled blue triangles (Bottema 1989) and the HI data are hollow purple squares (Begeman 1987). The black curve is a fit using Eq. 4.1 to the gas data and the green curve is a fit to the stellar data. The fitting parameters are: for the black curve, $v_0 = 117 \, \text{km} \, \text{s}^{-1}$, $R_c = 0.89 \, \text{kpc}$ and $\zeta = 2.1$; and for the green curve, $v_0 = 108 \, \text{km} \, \text{s}^{-1}$, $R_c = 0.95 \, \text{kpc}$ and $\zeta = 3.2$.

being external to the Type II hump.

### 4.1.2 Rotation curves

Several authors have measured rotation curves for NGC 6503. HI observations by Begeman (1987) and Greisen et al. (2009) reveal a remarkably flat rotation curve to $\sim 20 \, \text{kpc}$ deviating by no more than $\sim 4 - 5\%$, and a maximum tangential velocity of $\sim 120 \, \text{km} \, \text{s}^{-1}$. B89 gives Hβ and [OIII] observations of the inner rotation curve, as well as the pure stellar rotation curve, while de Vaucouleurs and Caulet (1982)
supply Hα data. Where they overlap, all data sets are in agreement with each other. The stellar rotation curve largely coincides with the Hβ and [OIII] curves, with an average difference of approximately 2 km s\(^{-1}\). At about 1.5 kpc the stellar rotation exceeds the gas rotation, possibly owing to the effect of a spiral arm. For our models, we fit the HI, Hβ and stellar data sets. These are found in Fig. 4.4, along with fits made using the fitting formula

\[ v(R) = \frac{v_0}{(1 + (R_c/R)\zeta)^{1/\zeta}} \]

where \(v_0\) is an asymptotic velocity, \(R_c\) is a projected length scale and \(\zeta\) is a shape parameter governing the sharpness of the ‘turnover’ to the asymptotic velocity (Courteau, 1997). The fit suggests an asymptotic tangential velocity of 117 km s\(^{-1}\) and shows how close the two rotation curves are.

The coincidence of stellar and gas rotation curves needs to be accounted for. If the gas rotation traces the gravitational potential, the difference between the two rotation curves provides a measure of the asymmetric drift. The asymmetric drift for an exponential disc with a cylindrically aligned velocity ellipsoid may be obtained from Eq. 2.15, yielding

\[ v_a = \frac{\sigma_R^2}{2v_c} \left[ \frac{2R}{R_d} + \frac{1}{2} \left( \frac{R \partial v_s}{v_s \partial R} - 1 \right) \right] \]

where \(R_d\) is the disc scalelength, \(v_s\) is the tangential velocity of the stars and \(\sigma_R\) is the radial velocity dispersion. When applying this equation to the observed Hβ rotation curve the resulting asymmetric drift is on the order of 6–10 km s\(^{-1}\) (depending on the adopted central radial dispersion, which lies somewhere between 40 and 60 km s\(^{-1}\) for consistency with the LOS dispersion). This is much larger than the observed difference between the gas and stellar curves; a possible cause of the discrepancy is
that gas does not trace the gravitational potential. Traditionally, gas is assumed to travel on circular orbits, tracing the gravitational potential, and thus the mass. The assumption of circularity in gas rotation curves underlies much of the work on mass modelling in the literature. However, gas has its own asymmetric drift (its dispersion is non-zero), and non-circular motions can impact the gas rotation curves of disc galaxies (see, e.g., Hayashi et al. 2004; Valenzuela et al. 2007), reducing the observed velocity relative to the actual circular velocity.

Stellar kinematic data allow us to model the asymmetric drift using the stellar rotation curve and thereby obtain the true circular velocity. The rotation curve is therefore open to multiple interpretations. In this work we model the rotation in two ways: we fit the gas rotation by assuming that it is the same as the circular velocity; and we fit the stellar rotation by simulating the LOS observations of the inclined disc. Simulating the LOS along a line of sight is particularly important for stellar kinematic measurements because of the non-negligible scale height of the stellar disc, which leads to integration effects as the line of sight passes through regions of the disc that are not on the plane. This effect is less important for the gas disc, which is expected to have a smaller scale height.

### 4.1.3 The LOS velocity dispersion

The LOS velocity dispersion profile for NGC 6503 along the major axis is found in B89; there is a $\sigma$-drop within 200 pc and a normal exponential structure outside that range, with a central exponential dispersion extrapolated to about 55 km s$^{-1}$ (B89), as shown in Fig. 4.5. We have fit the data to an exponential plus an inverse Gaussian
Figure 4.5: LOS stellar dispersion data from Bottema (1989) in red and an exponential + inverse Gaussian (Eq. 4.3) fit to the data in black. The fitting parameters are given by $\sigma_0 = 59.5 \, \text{km s}^{-1}$, $R_\sigma = 1.16 \, \text{kpc}$, $A = 0.59$ and $B = 0.12 \, \text{kpc}$.

in the centre:

\[
\sigma(R) = \sigma_0 \exp\left(-\frac{R}{R_\sigma}\right) \left[1 - A \exp\left(-\frac{R^2}{2B^2}\right)\right]
\]  

(4.3)

where $A$, $B$, the extrapolated central dispersion $\sigma_0$ and projected scale length $R_\sigma$ are the fitting parameters. The fit suggests a scale length $R_\sigma$ of 1.16 kpc and an extrapolated central velocity dispersion of 59.5 km s$^{-1}$, which will provide a point of comparison for our results in §8.1. Note that we do not attempt to model the physical processes which may be responsible for the $\sigma-$drop, but we do determine structural parameters and effective $M/L_R$ ratios for the bulge.

NGC 6503 was the first galaxy shown to have a $\sigma-$drop. A growing body of data suggests that $\sigma-$drops are found in numerous spiral galaxies (Márquez et al.,
2003; Comerón et al., 2008; de Lorenzo-Cáceres et al., 2008), although Koleva et al. (2008) argue that some $\sigma$—drops may be observational artifacts. BG97 suggest that the $\sigma$—drop is due to a small central, possibly star forming component capable of continually cooling the centre. A similar explanation for the drop is proposed by Wozniak et al. (2003) and Wozniak and Champavert (2006), who argue, using full physics simulations, that the presence of dynamically cold gas funnelling into the centre is responsible for forming dynamically cold stars that dominate the observed LOS kinematics. Comerón et al. (2008)’s sample of $\sigma$—drop galaxies correlates with higher incidence of nuclear dust spirals, H$\alpha$ rings and Seyfert fraction, which would be consistent with a cold star-forming component in the centre due to gas inflow. Meanwhile, de Lorenzo-Cáceres et al. (2008) suggest that inner nuclear bars may be responsible for at least some of the central dispersion drops found in the literature.

### 4.1.4 Four scenarios to test

The peculiarities discussed above invite several possible lines of enquiry. To account for the different interpretations, we test four different scenarios. In scenario K (for Kormendy), we fit the full surface brightness with an inner truncated Kormendy disc, and we fit the stellar rotation curve. In scenario KL (Kormendy light), we fit the surface brightness using an inner truncated light profile atop an exponential disc to account for dust extinction, and fit the stellar rotation as in scenario K. In scenario KG (Kormendy gas), we fit the surface brightness as in K but fit the gas rotation by assuming that the gas travels on circular orbits, without fitting the stellar rotation data. Finally, in scenario E (exponential), we fit the surface brightness excluding the points that constitute the Type II hump, and fit the stellar kinematic data as in K.
Data Scenario

<table>
<thead>
<tr>
<th>Data</th>
<th>K</th>
<th>KL</th>
<th>KG</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R-$band surface brightness</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>$R &lt; 0.1$ kpc, $R &gt; 2$ kpc</td>
</tr>
<tr>
<td>$\text{H} \beta$ rotation</td>
<td>$&gt;3$ kpc</td>
<td>$&gt;3$ kpc</td>
<td>Full</td>
<td>$&gt;3$ kpc</td>
</tr>
<tr>
<td>HI rotation</td>
<td>$&gt;3$ kpc</td>
<td>$&gt;3$ kpc</td>
<td>Full</td>
<td>$&gt;3$ kpc</td>
</tr>
<tr>
<td>Stellar rotation</td>
<td>Full</td>
<td>Full</td>
<td>None</td>
<td>Full</td>
</tr>
<tr>
<td>Stellar dispersion</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
</tr>
</tbody>
</table>

Table 4.1: Summary of each scenario and the radial range over which each data set is fit for each scenario.

Scenario E therefore alleviates the need for an inner truncated disc. In all cases, we fit the LOS velocity dispersion and the HI data outside 3 kpc. These runs, and the fitting procedures, are described in more detail in chapter 7; a summary of the fitted data is found in Table 4.1. We avoid modelling the Type II hump itself or the colour profile because this would require accounting for stellar populations, for which we do not yet have a good model; these would also add more parameters to the models, likely increasing the computation time significantly.

4.1.5 A cautionary note

There are difficulties involved in measuring a number of observables for highly inclined galaxies. In particular, measurements of the ionized gas are subject to beam smearing because the finite aperture covers a large range in azimuth, and to non-Gaussianity in the emission peak spectra, making it difficult to accurately determine the true rotation. Furthermore, long slit spectra are susceptible to non-axisymmetric motions generated by spirals. The HI data that we use have been corrected for beam smearing (and display a slightly larger velocity in Fig. 4.4 than the H$\beta$ data). Stellar kinematics are less affected by these considerations. Where possible, we evaluate the model data
by inclining the model galaxy and sampling along the line of sight (§6.3 contains more detail), but in the absence of a gas disc model this cannot be done for the gas observations.

Furthermore, dust (clearly visible in Fig. 4.1) may impact the surface brightness significantly; we can test for dust extinction causing the Type II profile, but dust is found throughout the galaxy. As noted, we do not include a full dust model, but we do not expect the presence of dust to alter the structure of the surface brightness, provided it is not distributed too unevenly. We would then expect the dust to affect only our derived $M/L_R$ ratios, which would be somewhat higher than the true values.

### 4.2 Prior models for NGC 6503

Prior research modelling NGC 6503 has been published by B89 and BG97. B89 present dynamical models for NGC 6503 to fit the data published therein, while BG97 construct dynamical models of NGC 6503 and evolve them to test for bar stability. Both assume that the mass distribution of the disc is

$$
\rho(R, z) = \rho_0 \exp(-R/R_0) \text{sech}^2(z/z_0) \tag{4.4}
$$

with a scalelength of $R_0 = 1$ kpc (van der Kruit and Searle, 1981a). The model is an isothermal disc, for which the velocity dispersions obey

$$
\sigma_z = \sqrt{\pi \Sigma(R) z_0} \tag{4.5}
$$

$$
\sigma_R = \sigma_z/0.6 \tag{4.6}
$$

$$
\sigma_\phi = \sigma_R \sqrt{B/(B - A)} \tag{4.7}
$$

where $\Sigma(R)$ is the projected surface density, $z_0$ is the scale height of a sech$^2$ disc, and $A$ and $B$ are the Oort constants. The relationship between $\sigma_R$ and $\sigma_\phi$ comes from the
epicycle equations; however, the relationship between $\sigma_R$ and $\sigma_z$ is an extrapolation from the observed value in our solar neighbourhood. It is not clear if there is a theoretical justification for this ratio that would cause it to hold throughout the disc of a normal galaxy; see §5.2 for a fuller discussion.

B89 adopted this model and a second model in which $Q$ is held constant throughout the disc, in accordance with Carlberg and Sellwood (1985). For this model, the radial dispersions obey

$$\sigma_R \propto \frac{P}{\kappa} \exp(-R/R_0) (B(B - A))^{-1/2}. \quad (4.8)$$

In both cases, the line-of-sight velocities and velocity dispersions are calculated. B89 finds that both models yield reasonable fits (the former is slightly better), but cannot reproduce the central dispersion drop. To reproduce a borderline stable disc where $Q = 1.7$ (Sellwood and Carlberg, 1984), B89 finds that the disc $M/L$ must be $1.7 \pm 0.3$ in the $B$-band.

BG97 model the galaxy in more detail and provide stability studies; they adopt the gas rotation curve as the fundamental input to their simulations and fit an isothermal, cored halo profile to the rotational data from B89. Because of the inner reddening, these researchers argue that the Type II hump is due to dust extinction, and therefore treat the disc as a pure exponential with scale length $40''$. They vary the disc to halo mass ratio in order to investigate the stability properties of the galaxy and to verify that a cold centre could form or persist, by (for example) decreasing the scale height near the centre. In their model, $Q$ for each disc declines slightly from the lowest disc to halo mass ratio to the highest.

While the dispersions of the isolated disc scale with the square root of the masses (see Eqs. 4.5 and 4.6 above) the process of embedding the disc in a dark halo modifies
the dispersions; BG97 find that simply embedding the isolated disc in the halo allowed the settling process at the beginning of the $N$–body simulations to rearrange the dispersions itself.\textsuperscript{1} The result is an initial central outflow in the simulations, leading to a decrease in the observed dispersions for the heavier discs. This is a settling effect, but it more-or-less naturally accounts for the observed dispersions of NGC 6503. BG97 find that their two lightest discs remain stable to bars and do a reasonably good job of reproducing the observed dispersions.

To rectify the asymmetric drift inconsistency described in §4.1.2, BG97 find that, provided the disc mass is low enough, the initial settling process could rearrange mass so as to maintain the coincidence between the two rotation curves over the course of the galaxy’s evolution. Of course, as the galaxy disc mass increases, the appearance of spiral arms and bar structure will cause the dispersions to increase considerably and therefore increase the asymmetric drift, an effect that may also be aggravated by numerical two-body scattering. It is also worth noting that it is not possible to discern, from the results that BG97 present, whether or not their drift is small enough to be consistent with the observations.

\textsuperscript{1}BG97 attempted to account for this: $\sigma_z$ is modified by multiplying by a function of the halo contribution to the rotation curve to the disc contribution, i.e., $F(\epsilon)$ where

$$\epsilon = \left( \frac{\rho_{z=0}^{\text{halo}} - \frac{1}{4\pi R} \frac{\partial (v_R^2)}{\partial R}}{\rho_{z=0}^{\text{disc}}} \right).$$

(4.9)

The result is that when $\epsilon$ is less than 0 (as it could be near the centre where $\frac{\partial}{\partial R} v_R^2$ is very large) the dispersion can decrease upon embedding, while large values of $\epsilon$ cause the dispersions to increase. Because the calculated value of $F(\epsilon)$ was incomplete at the time of writing, this factor was dropped for the models discussed in that paper and therefore the isolated disc dispersions were used (Bottema, private communication).
4.2.1 How can we improve these models?

There are multiple interpretations of the surface brightness and rotation curve data that merit investigation, as delineated earlier in this chapter. Additionally, the models used by BG97 may be made more realistic by employing a realistic halo and bulge; by adding an inner-truncated Kormendy disc to allow more accurate modelling of the surface brightness; and by decoupling the vertical dispersion from the radial dispersion. The GalactICS models, described in the next chapter, possess these advantages. The new halo can be used to obtain the cusp value and the new bulge can be used to determine the effective properties of the bulge, such as its radius and $M/L_R$ ratio. Furthermore, using models based on distribution functions, as we do here, ensures that settling effects, such as seen by BG97, are not observed when evolving the system using $N$–body simulations and therefore cannot affect the dynamics.
Chapter 5

GalactICS model

The GalactICS models are multicomponent galaxy models from Kuijken and Dubinski (1995), modified by Widrow et al. (2008). These models consist of a bulge, disc, and halo, each defined by its distribution function. The DFs are described in terms of several parameters that can be varied to produce a particular model. They have been further modified to reproduce a Kormendy hole; the description of these modifications follows the description of the standard GalactICS models.

The Jeans theorems are used to ensure that the models are in equilibrium. The Jeans theorems state that any function of the integrals of motion will solve the CBE, and that any galaxy with regular orbits and incommensurable frequencies can be represented by DFs that are a function of three isolating integral of motion. A spherical system, such as the bulge or halo, may be defined by a DF which is a function of only the energy $E$, provided that its velocity dispersion tensor is isotropic (otherwise, additional integrals of motion are required, usually the components of the angular momentum vector). However, the models must reproduce a target density; thus, we need to obtain a distribution function from the density distribution, which
is accomplished using Eddington inversion. Eddington’s inversion formula is given by
\[
f(\mathcal{E}) = \frac{1}{\sqrt{8\pi^2}} \left[ \int_0^{\mathcal{E}} \frac{d\rho}{d\psi^2} \frac{d\psi}{\sqrt{\mathcal{E} - \psi}} + \frac{1}{\sqrt{\mathcal{E}}} \left( \frac{d\rho}{d\psi} \right)_{\psi=0} \right]
\]
(5.1)
where \(\Psi = -\Phi + \Phi_0\) is the relative potential, \(\mathcal{E} = \Psi - \frac{1}{2}v^2\) is the relative energy and \(\Phi_0\) is a constant satisfying the boundary condition that \(\Psi \to \Phi_0\) as \(r \to \infty\).

### 5.1 The GalactICS bulge and halo

The bulge is designed to reproduce the Sérsic profile (Sérsic, 1968; Ciotti, 1991) upon projection, for which the intensity is given by
\[
I(r) = I_0 \exp \left( -b \left( \frac{r}{r_b} \right)^{1/n} \right)
\]
(5.2)
where \(r\) is the projected radius, \(I_0\) is the central intensity, \(r_b\) is the half-light radius, \(n\) is the index and \(b\) is a parameter dependent on \(n\). The expression is deprojected using an Abel integral equation (Binney and Tremaine, 2008) to yield the desired density distribution
\[
\rho_{\text{bulge}}(r) = \rho_b \left( \frac{r}{r_b} \right)^{-p} \exp \left[ -b \left( \frac{r}{r_b} \right)^{1/n} \right]
\]
(5.3)
where \(r\) is now the spherical radius and \(p = 1 - 0.6097/n + 0.05563/n^2\) yields the Sérsic profile (Prugniel and Simien, 1997). We define a characteristic velocity scale \(v_b\) for the bulge by
\[
v_b = \left[ 4\pi nb^{n-2} \Gamma(n(2-p))r_b^2 \rho_b \right]^{1/2}
\]
(5.4)
and use this as an input parameter instead of \(\rho_b\). This yields three parameters: the characteristic radius \(r_b\), the characteristic velocity \(v_b\), and the Sérsic index, which
we now designate $n_b$. In addition, a rotation parameter is introduced to account for bulge rotation.

For the halo we adopt a profile in accordance with cosmological simulations of structure formation suggesting that halos must be cusped (e.g., Navarro, Frenk & White 1996, 1997; Moore et al. 1999; Diemand et al. 2005), given by

$$\rho = \frac{\rho_h}{(r/a_h)^\gamma (1 + r/a_h)^{3-\gamma}}$$  \hspace{1cm} (5.5)

where $\rho_h$ and $a_h$ are the scale density and radius, respectively. A characteristic velocity scale $\sigma_h$ is introduced according to $\sigma_h^2 = 4\pi a_h^2 \rho_h$. For practical purposes, the density profile is truncated using an error function, so the preceding equation is multiplied by erfc \( \left( \frac{r-r_h}{\sqrt{2\delta}} \right) \), where $\delta$ is the truncation width. The cusp value $\gamma$ is of particular interest – for a NFW profile, $\gamma = 1$. We allow the cusp value to vary here for two reasons: first, to provide maximum flexibility in fitting the observed data so that we may investigate how cusped the halo of this galaxy is; and second, simulations of halo formation suggest that cusp values other than 1 may be appropriate (Moore et al., 1999) and that NFW profiles may not be universal (Jing and Suto, 2000; Navarro et al., 2008).

### 5.2 The GalactICS disc

The GalactICS disc is a Kuijken and Dubinski (1995) disc, which employs two integrals of motion: the energy $E$ and the angular momentum $L_z$, and an approximate third integral describing the energy in the vertical motions, $E_z$. The reason for this is that, by the Jeans theorem, any function of three isolating integrals of motion in a given potential will exactly solve the CBE (Binney and Tremaine, 2008). The DF is
given by
\[
f(E_p, L_z, E_z) = \Omega(R_c) \sqrt{2\frac{\pi^{3/2}}{\kappa(R_c)\sigma^2(R_c)\sigma_z(R_c)}} \exp \left[ \frac{E_c(R_c) - E_p}{\sigma^2(R_c)} - \frac{E_z}{\sigma^2_z(R_c)} \right]
\] (5.6)
where \( E_p \equiv E - E_z \) is the energy in planar motions (and is an integral since \( E \) and \( E_z \) are integrals as well), \( R_c \) is the radius of a circular orbit with angular momentum \( L_z \), \( E_c \) is the energy of such an orbit, and the tilde functions are target distributions. These are chosen to match observations of disc galaxies; the radial density distribution is exponential the vertical distribution is a \( \text{sech}^2 \) distribution,
\[
\rho(R, z) = \rho_0 \exp(-R/R_0)\text{sech}^2(z/z_0)
\] (5.7)
where \( R_0 \) is the photometric scale length and \( z_0 \) is the scale height. Consistent with observations of edge-on disc galaxies, \( z_0 \) is constant with radius (van der Kruit and Searle, 1981a).

The third integral \( E_z \) allows \( \sigma_z \neq \sigma_R \); observations of the solar neighbourhood suggest that \( \sigma_z/\sigma_R = 0.6 \), which would be impossible if the DF were a function of only two integrals of motion. However, it is not clear that \( \sigma_z/\sigma_R = 0.6 \) holds at all radii in a general disc galaxy. Gerssen et al. (1997) and Gerssen et al. (2000) find that the ratio is closer to 0.7 for NGC 488 and to 0.85 for NGC 2985, and Westfall et al. (2007) find that the ratio is also high for NGC 3949 and NGC 3982 (the latter’s vertical dispersion is larger than its radial dispersion). However, Ida et al. (1993) and Shiiudsuka and Ida (1999) use the theory of two body scattering and orbit integrations to show that \( \sigma_z/\sigma_R \sim 0.6 \) is roughly correct provided that \( \kappa/\Omega \lesssim 1.5 \); higher values of \( \kappa/\Omega \) imply higher \( \sigma_z \) values. This condition is probably satisfied in the centres of galaxies, and it would therefore be unreasonable to extend a fixed ratio of velocity dispersions across the entire disc.  

\[\text{In fact, one might surmise that the reason for the central outflow in BG97’s heavier discs is the}\]
Therefore, the GalactICS model decouples the vertical and radial dispersion. The vertical dispersion is given by the vertical potential gradient and scale height. The radial velocity dispersion is given by

$$\sigma_R^2 = \sigma_0^2 \exp(-R/R_\sigma)$$  \hspace{1cm} (5.8)

and the tangential dispersion is given by the epicycle equations, i.e.

$$\sigma_\phi = \frac{\kappa}{2\Omega} \sigma_R$$  \hspace{1cm} (5.9)

where $\kappa$ is the epicycle frequency (Eq. 2.4). Note that the vertical dispersion, contrary to what BG97 assumed, need not be 60% of the radial dispersion.

The GalactICS disc models also include a parameter to control the scale length of the velocity dispersion profile. Although observations by Bottema (1993) suggest that $R_d = R_\sigma$, there is no clear theoretical reason why this should necessarily be the case. Here we treat them independently to provide maximum flexibility in fitting all data sets.

### 5.3 The inner-truncated disc

As noted in §3.1.1, the Type II hump may be due to dust extinction, or it may be intrinsic to the structure of the disc. We can account for the effect by treating the disc as having an inner truncation (Eq. 3.4). This is done in two ways: by treating the standard exponential disc as having an inner truncated light profile in which inner lack of light is due to dust extinction, and by redefining the disc density to have an inner truncated profile. The former method allows us to directly model the effect of incorrect ratio of velocity dispersions there.
dust extinction on an exponential disc in an *ad hoc* fashion, and the latter accounts for the case in which the flattening is intrinsic to the density profile.

In the case of dust extinction, the Kormendy cutoff expression is

\[
I(R) = I_0 \exp\left[\frac{-R}{R_d} + \left(\frac{R_h}{R}\right)^\alpha \right] \tag{5.10}
\]

where \( R_h \) is the turnover radius, \( \alpha \) is the cutoff index and \( R \) is the major axis radius because the profile is applied to the projected profile. Kormendy (1977) finds that \( \alpha \approx 3 \) yields good fits to many galaxies, but we allow this value to vary.

The original KD model is constructed so that any axisymmetric density profile may be used instead of a standard exponential, which we exploit to implement a full inner truncated Kormendy disc. Eq. 5.7 now becomes

\[
\rho(R, z) = \rho_0 \exp\left[\frac{-R}{R_d} + \left(\frac{R_h}{R}\right)^\alpha \right] \operatorname{sech}^2\left(\frac{z}{z_0}\right) \tag{5.11}
\]

where \( \rho \) refers to the mass density. In this instance, we are assuming that disc has a dearth of mass in the centre instead of merely a dearth of light coming from the central mass, and the Kormendy profile is applied to the full three-dimensional profile rather than the projection.

The Kormendy profile has nothing to say about the velocity dispersion of such a disc; indeed, Eqs. 4.5 and 4.6 suggest that the dispersion would approach zero at the centre. Applying the Kormendy profile therefore requires us to choose a realistic velocity dispersion profile for this galaxy. We have conducted separate simulations showing that Kormendy discs embedded in cuspy halos are stable, provided they possess the same exponential velocity dispersion as outlined above for the exponential disc. If we instead assume that the radial dispersions decrease towards the centre, the resulting models are grossly out of equilibrium and the dispersion quickly increases in the centre to an exponential-type profile, in the process wiping out the central hole.
5.4 Summary

The GalactICS model modified to include inner truncation has two further parameters for the disc: the turnover radius $R_h$ and the cutoff index $\alpha$. We complete the parameter input set with mass-to-light ratios for the bulge and disc. We assume that the mass-to-light ratios are constant with radius, and use the same mass-to-light ratio for the surface brightness fit and the stellar kinematic fits. Although the frequency range over which the stellar observations were taken are not identical to the $R$–band, using different $M/L$ ratios for the surface brightness and stellar kinematic invites complications owing to correlations between mass-to-light ratios in different bands. We do not expect the error introduced by this simplification to be significant, in any case.

Thus the free parameters are: The five halo parameters; the disc mass, scale height, scale length, and Kormendy parameters; all four bulge parameters; the central radial dispersion and dispersion scale length; and the disc and bulge mass-to-light ratios. In addition, each data set has an associated noise parameter that is also allowed to vary (see §6.3). We also fix the halo to be nonrotating and truncate the disc at 5 kpc (well outside the measured range). The galaxy inclination $i$ is also fixed at 74°.

A list of the relevant parameters that are fit, as well as further calculated quantities that each run is made to output, is presented in Table 5.1. The units we use are defined so that $G = 1$; the unit of length is the kpc and the unit of velocity is 100 km s$^{-1}$. This yields a unit of mass equal to $2.325 \times 10^9 M_{\odot}$; for simplicity we have converted the masses back to $10^9 M_{\odot}$ in the figures that follow.
### Chapter 5. Galactics Model

#### Table 5.1: Table of the input parameters for the MCMC chains, as well as calculated quantities that are output for each model.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_h$</td>
<td>kpc</td>
<td>Halo truncation radius</td>
</tr>
<tr>
<td>$v_h$</td>
<td>$100 \text{ km s}^{-1}$</td>
<td>Halo characteristic velocity</td>
</tr>
<tr>
<td>$a_h$</td>
<td>kpc</td>
<td>Halo characteristic radius</td>
</tr>
<tr>
<td>$\delta$</td>
<td>kpc</td>
<td>Halo truncation width</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Dimensionless</td>
<td>NFW-type profile index</td>
</tr>
<tr>
<td>$M_d$</td>
<td>$10^9 M_\odot$</td>
<td>Exponential disc mass</td>
</tr>
<tr>
<td>$R_d$</td>
<td>kpc</td>
<td>Photometric disc scale length</td>
</tr>
<tr>
<td>$h_d$</td>
<td>kpc</td>
<td>Disc scale height</td>
</tr>
<tr>
<td>$n_b$</td>
<td>Dimensionless</td>
<td>Sérsic index</td>
</tr>
<tr>
<td>$v_b$</td>
<td>$100 \text{ km s}^{-1}$</td>
<td>Bulge characteristic velocity</td>
</tr>
<tr>
<td>$r_b$</td>
<td>kpc</td>
<td>Bulge characteristic radius</td>
</tr>
<tr>
<td>$\sigma_0$</td>
<td>$100 \text{ km s}^{-1}$</td>
<td>Central velocity dispersion</td>
</tr>
<tr>
<td>$b_{\text{rot}}$</td>
<td>Dimensionless</td>
<td>Bulge rotation parameter</td>
</tr>
<tr>
<td>$(M/L_R)_d$</td>
<td>$M_\odot/L_\odot$</td>
<td>Disc mass to light ratio</td>
</tr>
<tr>
<td>$(M/L_R)_b$</td>
<td>$M_\odot/L_\odot$</td>
<td>Bulge mass to light ratio</td>
</tr>
<tr>
<td>$R_\sigma$</td>
<td>kpc</td>
<td>Radial dispersion disc scale length</td>
</tr>
<tr>
<td>$R_h$</td>
<td>kpc</td>
<td>Kormendy cutoff radius</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Dimensionless</td>
<td>Kormendy cutoff index</td>
</tr>
</tbody>
</table>

#### Other Quantities

- Disc mass: $10^9 M_\odot$ Real disc mass
- Bulge mass: $10^9 M_\odot$ Model bulge mass
- Halo mass: $10^9 M_\odot$ Model halo mass within 20 kpc
- $Q$: Dimensionless Toomre local stability parameter
- $X$: Dimensionless Toomre global stability parameter

#### Fixed Quantities

- $R_{\text{trunc}} = 5$ kpc Disc truncation radius
- $\delta_{\text{trunc}} = 0.5$ kpc Disc truncation width
- $i = 74^\circ$ Galaxy inclination
- Distance = 5200 kpc Distance to galaxy
Chapter 6

Bayes’ Theorem and Markov chain Monte Carlo

A wide range of astrophysical problems are amenable to Bayesian analysis, and Markov chain Monte Carlo methods are an excellent, versatile and precise way to implement it. As noted in the introduction, Bayesian/MCMC methods are widely used in cosmology for a variety of applications (Tegmark et al., 2004; Percival, 2005). The technique is also used for detecting black hole and neutron star binary inspirals from interferometric gravitational radiation, because of the enormous amount of data collected by such interferometers and the high dimensionality of the parameter space (Cornish and Crowder, 2005; Röver et al., 2006; Cornish and Porter, 2006; van der Sluys et al., 2008), and finds widespread use in planetary astrophysics to model the orbits of extrasolar planets (Ford, 2005; Gregory, 2007) and in stellar physics to characterize stellar oscillations (Brewer et al., 2007; Benomar, 2008). MCMC techniques have gained less traction in extragalactic astronomy, but Kelly et al. (2008) use this technique to derive parameters for the galaxy luminosity function, finding that the
maximum likelihood method yields too-narrow confidence intervals. Other applications include gravitational lensing studies (Barnabè and Koopmans, 2007; Jullo et al., 2007), in which MCMC is used to constrain mass profiles; Corless and King (2008) use MCMC to constrain parameters for triaxial halo models using lensing data. In this chapter we explain Bayes’ theorem and MCMC techniques and discuss their applications to our problem of interest.

6.1 Maximum likelihood

Before discussing Bayesian statistics, we briefly explain the logic of a maximum likelihood method. The likelihood function is used to determine which model parameters are ‘most likely’, based on the observed data, but it does not generally correspond to the most probable parameters. The distinction occurs because likelihood refers to parameters obtained from a known experimental outcome, which is the reverse of a true probability; the likelihood function $\mathcal{L}(B|A)$ is considered to be a conditional probability $P(A|B)$ viewed as a function of $B$ with $A$ held constant. If $p$ is the vector of model parameters, $f_p$ is the known univariate probability density function and $D_1\ldots D_N$ is the sample of observed data points, then the likelihood function $L$ is defined as

$$\mathcal{L}(p) = \prod_{i=1}^{N} f_p(D_i)$$ (6.1)

provided the data points are independent, a reasonable assumption for our data. $\mathcal{L}$ can then be maximized using any number of multivariate optimization methods. When $f_p$ is Gaussian – i.e., when the errors for the observed data points are normally distributed – the result is the usual problem of minimizing the $\chi^2$ distribution to obtain the best fit parameters, once the log of Eq. 6.1 is taken. The minimization
may be done by many methods, such as a downhill simplex methods or gradient based
methods, such as the Levenberg-Marquadt method. Thus, $\chi^2$ minimization is a special
case of the maximum-likelihood method, albeit one with a high degree of applicability.
It is also possible to determine error contours (e.g., Efstatíthiou et al. 1988).

6.2 Bayes’ theorem

As the name implies, maximum likelihood methods only supply the likelihood; ideally
we would prefer to have the full posterior probability distribution of the parameters
$p$. This is where Bayes’ theorem comes in. Bayes’ theorem is a result from probability
theory that allows us to determine the probability distribution of a set of variables
using prior information available to us. The PDF of the model parameters is written
as $P(p|D)$, where, as before, $p$ is the vector of model input parameters and $D$ is the
vector of observational data. Bayes’ theorem is written

$$P(p|D) = \frac{P(p)P(D|p)}{P(D)} \quad (6.2)$$

where $P(p)$ is the prior probability if the input parameters, $P(D|p)$ is the likelihood of
the data given a specific set of model parameters $p$ and $P(D)$ is the prior probability of
the data, which functions as a normalization. (Note that we have switched notations
for the likelihood; $P(D|p)$ is the conventional notation when discussing Bayesian
statistics.)

Meanwhile, a Markov chain is defined as a stochastic process in which the sequence
of random variables have the property that past and future states are independent:

$$P(X_{n+1}|X_1, X_2 \ldots X_n) = P(X_{n+1}|X_n) \quad (6.3)$$
where the $X_i$ are the random variable and $P(X_i|X_j)$ is the probability of obtaining $X_i$ given $X_j$. Thus, in a Markov chain the value of the random variable at the next step depends only on the current step. The Markov chain Monte Carlo technique makes use of Markov chains to construct a stochastic process that samples a target density distribution. Thus, the idea behind MCMC is to sample the posterior probability distribution for the $n$–dimensional parameter space that defines our galaxy models.

There are numerous methods for constructing the chain; the one we use here is a Metropolis-Hastings algorithm (Metropolis et al., 1953; Hastings, 1970), which constructs a random walk in parameter space to sample the posterior PDF. The idea is to use a proposal distribution $Q$ from which to sample the next step of the chain; the Metropolis-Hastings algorithm is described thus:

1. Sample a proposed point $p^*$ from the proposal distribution $Q(p|p^*)$, where $p$ is the current value of the parameter vector

2. Evaluate the ratio

$$r = \frac{P(p^*|D) Q(p|p^*)}{P(p|D) Q(p^*|p)}$$  \hspace{1cm} (6.4)

3. If $r \geq 1$, accept $p^*$ as the next step of the chain. Otherwise, select a random variable $\alpha$ from $(0,1)$ and accept $p^*$ if $\alpha \leq r$, otherwise reject $p^*$ and let the next step of the chain simply be $p$ again

The proposal distribution $Q$ may be a Gaussian centred on $p$; in this case, the proposals are sampled from the area surrounding the current point $p$, and the proposal distribution is symmetric: $Q(p|p^*) = Q(p^*|p)$, so $r$ simplifies to $r = P(p^*|D)/P(p|D)$. We use this form of the Metropolis algorithm.
Now we may elucidate the details of our technique. We make use of Bayes’ Theorem (Eq. 6.2) to derive the probability distribution functions for each parameter that governs our models. We define each component on the right side of Eq. 6.2 as follows: the priors for all model parameters, $P(p)$, are taken to be unity; the likelihood function $P(D|p)$ is taken to be a $\chi^2$ function, described in detail below; and the normalization is ignored since it cancels in the division. The left hand side of Eq. 6.2 refers to the joint posterior probability distribution function (PDF) of the parameter space $\{p\}$.

We define, for the $i$th input parameter, a step-size $s_i$ and draw a random variable $g_i$ from a Gaussian distribution. If $p_i$ are the old parameter values and $p_i^*$ are the new parameter values, then the new parameters for the new step size are given by

$$p_i^* = p_i + g_i s_i$$

(6.5)

where $M$ is the total number of parameters that are allowed to vary and $1 \leq i \leq M$. Once $p_i^*$ are calculated, the likelihood function – i.e., the $\chi^2$ – is calculated and the Metropolis algorithm used to determine whether to accept or reject this point. At the next step of the chain, the process is repeated. The result is a sequence of parameter coordinates, some of which have been accepted, some of which have been rejected. Initially, the likelihood is very small, leading to rapid decline in the $\chi^2$ and high acceptance rates. After a period of ‘burn-in’, the $\chi^2$ values stabilize about a minimum and populate it. This chain contains all the relevant information about the posterior probability distribution of the input parameters. The marginal probability distribution functions are obtained by essentially making a histogram of the resulting parameter. Formally, we integrate out the other parameters:

$$P(p_i|D) = \int dp_1 dp_2 \ldots dp_{i-1} dp_{i+1} \ldots dp_M P(p|D)$$

(6.6)
for one parameter, and similarly for more parameters. For two parameters, marginalizing supplies the two-dimensional PDF, which can be represented as a contour plot in two dimensions.

6.3 The likelihood function

While it is possible to use a standard $\chi^2$ for the likelihood function, a subtle issue comes into play when dealing with the error bars. The error bars do not account for every possible source of error; in many cases they should be taken as no more than formal experimental error, unable to capture systemic errors. Moreover, some data sets do not have reported error bars, such as a large range of the surface brightness observations for NGC 6503. Furthermore, the model used may lack crucial features that are captured by the data; for example, an axisymmetric disc model cannot reproduce a wavy surface brightness profile which may indicate spirality in the disc. To alleviate these problems, we use augmented error bars by introducing a noise parameter $f$ for each data set. The new errors $\sigma_i^2$ are given by $\sigma_i^2 = \sigma_{0i}^2 + f^2$, where $\sigma_{0i}$ are the original experimental errors on each point. The noise parameters are identical for each data set; for NGC 6503 we fit four data sets and so require four noise parameters. The noise parameters act to increase the error bars and also provide errors on the surface brightness data where none exist. The likelihood function is

$$L = (2\pi)^{-N/2} \prod_{i=1}^{N} \sigma_i^{-1} \exp \left[ \frac{(d_i - m_i)^2}{2\sigma_i^2} \right].$$

(6.7)

Because of the error factor in the denominator, the error bars are prevented from enlarging arbitrarily. Thus, after burn-in, a sort of equilibrium is reached in which the tendency of the error bars to increase is balanced by the likelihood being driven
too high as they enlarge.

Finally, to complete the description of MCMC, we must explain how the likelihood is calculated and how we use the observations as input. The model values $m_i$ are calculated at every step of the MCMC run. For the gas rotation curve, we assume that the observed H$\beta$ curve traces the potential and therefore use the model potential to derive the gas rotation curve, modified for the galaxy’s inclination angle. For the stellar rotation curve and dispersion profile, we sample the distribution function along the line of sight represented by each data point from B89 to find the line of sight velocity and velocity square; at each point sampled along the LOS, the disc and bulge velocities are weighted for the mass-to-light ratios. Finally, for the surface brightness we sample the distribution function along elliptical annuli around the inclined galaxy centre whose axis ratio is determined by the inclination angle (i.e., $q = \cos i = 0.28$) to determine the elliptically averaged density and thus the surface brightness in mag arcsec$^{-2}$. We emphasize that in no cases have we corrected the observed profiles for inclination; all inclination effects are accounted for by rotating the model galaxy and sampling along the line of sight, except for the calculation of the circular velocity which simply requires a factor of $\sin i$.

In principle, we could calculate these quantities by generating the $N$–body realization, tilting it and taking quantities along a given line of sight, but this is computationally prohibitive to do at every step of a MCMC chain – in the process of generating a $N$–body realization from new input parameters, the CPU time for the calculation of particle positions and velocities exceeds the time taken for the calculation of the distribution function by an order of magnitude. We are, however, able to obtain reliable results by sampling the distribution function along the line of sight,
provided the sampling occurs along a sufficiently large number of points; the main
difficulty lies in resolving the centre correctly, especially for the purpose of detecting
the central dispersion drop. If too few points are sampled along the line of sight, the \( \sigma \)-drop is not correctly reproduced, even if all other results match the results
obtained by generating the \( N \)-body realization.

Note the distinction between maximum likelihood methods and Bayesian ap-
proaches; only if all priors are assumed to be unity do the likelihood and the posterior
probability equal each other, and results from maximum-likelihood and Bayes/MCMC
approaches should not differ significantly. However, in our view the use of Bayes’
theorem is still preferable because of its ability to fully sample the joint posterior
probability distribution of the parameters; thus, it is often possible to find substructure
in the distribution that is not apparent from maximum likelihood methods. Of
course, using non-uniform priors is an added benefit that cannot be captured by
maximum-likelihood methods.

### 6.4 Pros and cons of MCMC

With that in mind, let us enumerate the benefits and drawbacks of Bayes/MCMC

1. The full posterior probability distribution is sampled, which, in principle, per-
   mits the chain to detect substructure or multiple peaks in the likelihood func-
   tion, and allows us to determine formal errors easily.

2. It is easily modified to include more parameters if needed.

3. It tends to burn-in fairly rapidly.
4. The PDFs for other quantities can be calculated from the PDFs; for example, \( Q \) and \( X \).

5. There is some evidence that it provides more realistic error constraints than maximum likelihood methods (Kelly et al., 2008).

The primary disadvantages, meanwhile, are twofold. First, there is no guarantee of uniqueness to the discovered solutions, or even nondegeneracy in the random variable – i.e., multiple likelihood peaks are possible in any given parameter space. Strictly, the problem here is the likelihood function, and MCMC can handle it slightly better than maximum likelihood methods; a judicious choice of the step size and proper tweaking to obtain the correct acceptance rate mitigates this problem. Second, and related to the first, although MCMC chains tend to converge quickly, it is not so simple to verify that the converged chains have fully populated the joint posterior probability distribution, and it is not so simple to interpret cases in which the chain ‘spreads around’ the parameter space. Thus, formally ensuring convergence may require prohibitively long lengths of time. However, in practice this problem is not so severe; only a short period of time is required to find good models, and after a reasonable length of time we are in fact able to obtain good constraints for most parameters. This does not preclude the possibility that a MCMC chain passed through multiple likelihood peaks after burn-in, but it does suggest that there no reason to prefer anyone one such peak over the other when the \( \chi^2 \) values are nearly equal for all peaks.

A final point is worth mentioning. The acceptance rate of the MCMC chain is a barometer of the quality of the sampling process; obviously, if the acceptance rate is very low then it will take a very long time to fully sample the PDFs, so this is not
desirable. However, an acceptance rate which is too high may mean that the size of the steps \( s_i \) is too small. If \( s_i \) is arbitrarily small, we expect an acceptance rate of 50\% (because of the random number \( \alpha \) that mediates the Metropolis algorithm), so the ideal acceptance rate must be less than 50\%. The appropriate balance is found to be approximately 23\% by Roberts et al. (1997) for parameter spaces with dimensions > 2 (for one and two-dimensional parameter spaces the figure is 45\%). Therefore, we strive for an acceptance rate of 23\% in our chains.

### 6.5 Ensuring convergence

A number of tests for convergence of MCMC chains to the joint PDF of the parameter space have been proposed, with varying levels of rigour. Perhaps the simplest is visual examination of the time series trace of the random variable – extended periods of no change in the \( \chi^2 \) indicate poor mixing, so inspection often suffices to verify good mixing. A more rigorous measure of convergence is the autocorrelation function of \( k \)-th order, which measures correlations among the trace of a single parameter \( k \) steps apart. The autocorrelation function is given by

\[
A_k = \frac{\text{Cov}(\theta_t, \theta_{t+k})}{\text{Var}(\theta_t)} = \frac{\sum_{t=1}^{n-k} (\theta_t - \langle \theta \rangle)(\theta_{t+k} - \langle \theta \rangle)}{\sum_{t=1}^{n-k} (\theta_t - \langle \theta \rangle)^2} \tag{6.8}
\]

where \( \text{Cov}(\theta_t, \theta_{t+k}) \) refers to the covariance matrix element, \( \text{Var}(\theta_t) \) is the variance of the parameter \( \theta \) and \( \langle \theta \rangle \) is mean of the parameter over the course of the chain. We have \( A_k \leq 1 \); the closer \( A_k \) is to 1, the more strongly correlated every \( k \) values of the chain are. In practice, autocorrelations of 0.9 – 0.99 are not unusual for MCMC chains. A graph of \( A_k \) versus \( k \) is expected to yield an exponential decline.
Another test of convergence is to simply take ‘cuts’ of the chain at different intervals after burn-in (for instance, the first 20% and last 20%) and verify that the parameter means and error bars are similar (Geweke, 1992). This test, and other tests designed to verify full burn-in and mixing, fail to account for the possible existence of multiple minima, for which the means at different points in the chain will naturally be different. Furthermore, strong correlations between parameters will limit the effectiveness of this test because correlations tend to limit the mobility of the parameters (for example, masses and mass to light ratios are strongly correlated). Thus, using this test for specific parameters is a poor way to verify good mixing when the posterior PDFs are complex and possibly multimodal. On the other hand, when a parameter does satisfy this test, it provides a very strong indication that the parameter is properly mixed. However, while individual parameters may not be properly mixed, the autocorrelation and Geweke tests are much better applied to the $\chi^2$ values, which do display good mixing; see §6.7.

In practice, the existence of correlations makes it difficult to ensure proper mixing over a reasonable timescale, so a compromise is required between obtaining good statistics by ensuring formal convergence and computational constraints. We discuss these issues in more detail in §6.7.

### 6.6 An improved MCMC algorithm

There are two possible improvements to the algorithm described in §6.2. The first allows the chain to sample larger regions of parameter space before burn-in while speeding up the process of finding the global likelihood maximum, and is termed simulated annealing. In this technique, the Metropolis ratio $r$ (Eq. 6.4) is modified
by raising it to a power smaller than 1, i.e.,

\[ r = \left( \frac{P(p^*|D) Q(p|p^*)}{P(p|D) Q(p^*|p)} \right)^{\frac{1}{T}} \quad (6.9) \]

where \( T \) is a ‘temperature’ for the parameter space. \( T \) will typically start very high and decline as the chain progresses. The result is that, when \( T = 1 \), the chain should be burned in. This technique tends to ‘hone in’ on the global maximum more precisely and more quickly than the basic Metropolis-Hasting algorithm, which may take a very long time to find the best model.

The second improvement addresses the issue of correlated input parameters. Correlated parameters make the mixing process more difficult because an additional condition must be satisfied between two parameters at each step if the chain is to progress; for example, masses and mass-to-light ratios are tightly correlated (the contour plots in chapter 8 will make this obvious), but the jumps for each parameter at each step are not likely to lie along this correlation. One solution is to generalize the jumping rule, Eq. 6.5, as follows

\[ p^* = p + S \cdot G \quad (6.10) \]

where \( p \) and \( p^* \) are the current and proposed parameter vectors, \( G \) is the vector of random variables, and \( S \) is a matrix which, in the special case of Eq. 6.5, is diagonal. \( S \) may be given by the covariance matrix, which relates the correlations of various parameters; the elements of this matrix would then be given by the correlations found in a test MCMC run of a few thousand steps.

These two techniques may be merged by allowing staggered temperature decreases from a given initial value, and calculating the covariance matrix anew at each new temperature until the temperature reaches 1. The method works as follows: the
initial temperature is obtained by exploring the parameter space on the first pass and determining the variance of the likelihood value at each step. This temperature is fixed on the next pass, which lasts $n$ steps. At the end of the second pass, the temperature is halved and the covariance matrix is calculated. Then, the covariance matrix (which is no longer diagonal) is used in Eq. 6.10 for the next pass, which also lasts $n$ steps. The covariance matrix is calculated anew at the end of each pass and the temperature halved until it reaches 1, at which point the standard Metropolis algorithm takes over with the final covariance matrix used in Eq. 6.10. This method yields better acceptance rates than the standard Metropolis algorithm, and we employed it for the last MCMC run we conducted, which tested scenario E.

6.7 Our MCMC runs

As noted in chapter 4, we test four different interpretations of the data with four different MCMC runs. In all runs, we fit the stellar dispersion. For scenarios K, KL and KG, we fit the full $R$–band surface brightness; for scenario E, we fit the innermost eight surface brightness data points and the outer six data points, completely excluding the effect of the Type II hump. In scenario K, we model the galaxy using a disc with a Kormendy density profile and fit B89’s stellar rotation curve data to the inner region (within 3 kpc) and Begeman’s (1987) HI data outside that range. In scenario KG, we wish to compare the effect of different rotation curve fits on the fitted parameters, and therefore replace B89’s stellar rotation curve data with B89’s H$\beta$ data and treat it as if it were tracing the circular velocity. For scenario KL, we wish to test the hypothesis that dust in the central region may cause the Type II surface brightness profile, so we use a purely exponential disc mass profile (Eq. 5.7) and a
light profile that is inner truncated (Eq. 3.4), to mimic the effect of dust extinction; here, we fit rotation curves as in scenario K. In scenario E, we fit the rotations as in scenario K, and fix the disc scale length at 1.3 kpc.

Scenarios K, KG and KL are allowed to run for approximately 200 000 steps in total, although the steps used for the data analysis vary by run because some runs take longer to fully burn in than others. The acceptance rate data for each run is found in Table 6.7. Our acceptance rates hover between 10 and 20%, lower than the 23% ideal rate but adequate to obtain good statistics. One difficulty in obtaining ideal acceptance rates lies in managing the step sizes of each parameter in such a complex parameter space – step sizes too large will lower the acceptance rate while step sizes too small will increase it. In the former case, the length of time required to sample the space correctly is prohibitive; in the latter case, there is no guarantee that the chain will do a good job of covering the full region of the minimum. Note that the total time for each run was roughly two months on a standard desktop computer. Some parameters are not well mixed (in particular, the bulge scale length and velocity), but most display reasonably good constraints as the figures in the following sections will show, and even in cases where the parameters display poor mixing it is possible to obtain good constraints.

By contrast, the method described in §6.6 was used to derive the PDFs for scenario E. This method dramatically lowers the time required to locate a global minimum, and the PDFs are evidently well-mixed in the figures that follow. We ran the final chain for 20 000 steps and chose the dataset from the last 12 000 steps to use in the analysis. This data set took less than a week to compile, but required about two months of experimentation to find the right way to model the surface brightness
because of the ambiguity in how to cut the surface brightness profile.
Chapter 7

Results: Constraining the Physics

In all three runs, we are able to fit the observed data well and obtain the posterior values for each parameter. Our results are summarized in Table 7.1, which lists the average posterior value for each fitted parameter with error bars for each run. The quoted errors are 1σ confidence intervals.

In Fig. 7.1 are found the time series traces for the $\chi^2$ for scenarios K, KL and KG. The $\chi^2$ values are negative because the $\chi^2$ expression (the log of Eq. 6.1) includes a term $\log \sigma$, which is required because of the variable noise parameters. The full time series traces are plotted but full burn-in does not occur until $\sim 20000 - 60000$ steps into the chain. It is only after these points that the random variable begins to fixate about a stable value. The $\chi^2$ for scenario K and KL are roughly equal, and slightly lower than for scenario KG, although we caution that the interpretation of $\chi^2$ is less straightforward because of the variable noise parameters. In the case of scenario KL, two distinct minima appear in the trace; the first minimum corresponds to a marginally lower $\chi^2$ value than the rest of the chain. There is perhaps a degeneracy in this run; several PDFs display multiple peaks, and multiple maxima are clearly
visible in some of the two-dimensional marginal probability distributions in this and the next chapter. There is enough substructure in the time series traces to suggest that the other two runs also passed through multiple likelihood peaks after burn-in, although the effect is not so pronounced as for scenario KL (except in Fig. 7.7).

The time series trace for scenario E is found in Fig. 7.2. Because of the modified technique used for this run, burn-in occurs much more quickly and there is little evidence, either from Fig. 7.2 or the PDFs plots, that this run passed through multiple peaks.

We display autocorrelations for the $\chi^2$ values in Fig. 7.3 as a function of $k$. This shows that the correlations decrease with $k$, as expected, and the autocorrelations are reasonable (as noted, $A_k \gtrsim 0.9$ are not unusual for a MCMC chain). Combined with the good constraints obtained for most parameters, we infer that we found a good compromise between the need to ensure proper mixing and the computational constraints. In the case of scenario E, the modified technique leads to dramatically lower autocorrelations in spite of the short run time, confirming the soundness of the new approach.

7.1 Fit quality

We select an example from each scenario and demonstrate the quality of the fit in Figs. 7.4, 7.5 and 7.6. The surface brightness is broken down between the bulge and the Kormendy light profile, and the rotation curve panel displays the circular velocity and stellar rotation. Scenarios K, KL and KG do an excellent job of reproducing the surface brightness profile within 2 kpc, as expected; Scenarios K, KL and E fit the stellar rotation curve well, and scenario KG fits the gas rotation curve well. Note that
### Table 7.1: A table of the final posterior values for each input and calculated parameter with 1σ error bars. The units are as given in Table 5.1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Run</th>
<th>K</th>
<th>E</th>
<th>KL</th>
<th>KG</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r_h )</td>
<td></td>
<td>38.3</td>
<td>±4.1</td>
<td>33.7</td>
<td>±1.3</td>
</tr>
<tr>
<td>( v_h )</td>
<td></td>
<td>2.47</td>
<td>±0.02</td>
<td>2.38</td>
<td>±0.02</td>
</tr>
<tr>
<td>( a_h )</td>
<td></td>
<td>6.66</td>
<td>±0.49</td>
<td>8.69</td>
<td>±0.66</td>
</tr>
<tr>
<td>( \gamma )</td>
<td></td>
<td>1.22</td>
<td>±0.04</td>
<td>1.43</td>
<td>±0.03</td>
</tr>
<tr>
<td>( M_d )</td>
<td></td>
<td>5.77</td>
<td>±0.83</td>
<td>3.12</td>
<td>±0.43</td>
</tr>
<tr>
<td>( R_d )</td>
<td></td>
<td>0.71</td>
<td>±0.02</td>
<td>1.3</td>
<td>fixed</td>
</tr>
<tr>
<td>( z_d )</td>
<td></td>
<td>0.14</td>
<td>±0.01</td>
<td>0.24</td>
<td>±0.02</td>
</tr>
<tr>
<td>( n_b )</td>
<td></td>
<td>1.06</td>
<td>±0.07</td>
<td>1.93</td>
<td>±0.10</td>
</tr>
<tr>
<td>( v_b )</td>
<td></td>
<td>0.43</td>
<td>±0.04</td>
<td>0.51</td>
<td>±0.03</td>
</tr>
<tr>
<td>( r_b )</td>
<td></td>
<td>0.16</td>
<td>±0.02</td>
<td>0.35</td>
<td>±0.02</td>
</tr>
<tr>
<td>( \sigma_0 )</td>
<td></td>
<td>0.37</td>
<td>±0.03</td>
<td>0.34</td>
<td>±0.03</td>
</tr>
<tr>
<td>( b_{rot} )</td>
<td></td>
<td>0.56</td>
<td>±0.22</td>
<td>0.05</td>
<td>±0.03</td>
</tr>
<tr>
<td>( (M/L_R)_d )</td>
<td></td>
<td>1.86</td>
<td>±0.33</td>
<td>1.60</td>
<td>±0.21</td>
</tr>
<tr>
<td>( (M/L_R)_b )</td>
<td></td>
<td>0.87</td>
<td>±0.12</td>
<td>1.23</td>
<td>±0.11</td>
</tr>
<tr>
<td>( R_a )</td>
<td></td>
<td>2.57</td>
<td>±0.29</td>
<td>0.77</td>
<td>±0.05</td>
</tr>
<tr>
<td>( R_h )</td>
<td></td>
<td>0.69</td>
<td>±0.11</td>
<td>0.73</td>
<td>±0.10</td>
</tr>
<tr>
<td>( \alpha )</td>
<td></td>
<td>0.90</td>
<td>±0.14</td>
<td>0.56</td>
<td>±0.10</td>
</tr>
</tbody>
</table>

|          |     | Disc mass | 3.19 | ±0.59 | 3.03 | ±0.39 | 4.79 | ±0.49 | 2.95 | ±0.33 |
|          |     | Bulge mass | 0.070 | ±0.014 | 0.15 | ±0.02 | 0.045 | ±0.007 | 0.0038 | ±0.0023 |
|          |     | Halo mass | 61 | ±2 | 61 | ±2 | 59 | ±2 | 61 | ±2 |
|          |     | \( Q \) | 1.62 | ±0.24 | 0.69 | ±0.15 | 0.99 | ±0.14 | 1.58 | ±0.28 |
|          |     | \( X \) | 1.88 | ±0.46 | 2.84 | ±0.57 | 0.47 | ±0.08 | 1.51 | ±0.16 |
Figure 7.1: Graph of the $\chi^2$ value as a function of step for scenario K (top), scenario KL (middle) and scenario KG (bottom). The red lines indicate the points at which burn-in is reached.
CHAPTER 7. RESULTS: CONSTRAINING THE PHYSICS

Figure 7.2: Graph of the $\chi^2$ value as a function of step for scenario E. The red line indicates the point at which burn-in is reached.

Figure 7.3: Graph of the autocorrelation $A_k$ as a function of $k$ for scenario K (black), scenario KL (green), scenario KG (red) and scenario E (blue).
scenarios K, KL and E cannot fit the Hβ data and that scenario KG cannot fit the stellar rotation data. The dispersion profile fits are excellent for scenario KL and KG; the central dip is present in all models and the outer region is well reproduced. The dispersion profiles for scenario K are not as well fit, owing to a lower $\sigma_0$, but a dip is still present in this data. The outer region fit appears to overshoot the points slightly; the cause of this is the large $R_\sigma$ for these models. We have verified that decreasing $R_\sigma$ to 50% of the mean value produces better fits to the dispersion without affecting the remaining fits, suggesting that the cause of the poorer dispersion fit in the outer region is a local likelihood minimum. Scenario E displays a different type of dispersion profile, in which the bulge dominates the profile and the $\sigma$–drop is essentially absent. It is important to note that, in all runs, there is very little difference in the fits after burn-in – thus, the fits are virtually identical across all models for a given run, even in the case of KL where at least two distinct minima are apparent. Therefore, in the absence of a reliable way to distinguish one minimum from another, we supply the global average and error bars for all runs in Table 7.1.

Fig. 7.4 also displays the breakdown in rotation curve by halo, disc and bulge components. Since the disc is inner truncated in scenarios K and KG, the disc rotation curves for those runs start at a radius larger than 0, and the lack of disc influence on the rotation curve creates a kink at $\sim 0.2$ kpc which is missing from scenarios KL and E. The kink is small and probably not detectable in the observed rotation curve. The breakdown also starkly demonstrates how dominant the disc is in the inner region; for scenarios K, KG and E it appears to share dominance with the halo (and in these specific models it is slightly subdominant), but in scenario KL the disc is much more dominant than the halo in the inner parts. This suggests that scenario KL produces
models very close to maximal disc models.

### 7.2 Testing the rotation curve fits

Runs K and KG enable us to determine how changing the rotation curve fit influences the PDFs. Examining the rotation curve fits in Fig. 7.4, it is apparent that scenario K cannot fit the Hβ/HI data and scenario KG cannot fit the stellar rotation data. Because asymmetric drift is built into the GalactICS models and no consideration is made for possible sources of gas asymmetric drift, it is likely that scenario K represents more realistic models for NGC 6503. Both runs fit the HI data external to \( \sim 3 \) kpc, since there is not expected to be any asymmetric drift at these radii.

To examine the stability of the models for these scenarios to bar formation, we consider their \( Q \)– and \( X \)– values. Fig. 7.7 shows the marginal 2-dimensional PDF for the disc mass and the disc \( Q \), evaluated at its minimum (usually found at around \( 1 - 1.5 \) disc scale lengths). In addition, Fig. 7.8 shows the marginal probability distribution for the minimum disc \( Q \)– and \( X \)– values.

For scenario K, we find that the marginal PDF spans a large range in disc mass and \( Q \), forming a relatively narrow trough that spans from \( Q \sim 2.3 \) at disc masses of \( \sim 2.1 \times 10^9 M_\odot \) in the upper left to \( Q \sim 1.3 \) at disc masses of \( \sim 4.4 \times 10^9 M_\odot \) in the lower right. The \( Q-X \) plot shows that \( X \) lies between 1 and 3, and appears to correlate weakly with \( Q \), in the sense that models with high \( Q \) also tend to have high \( X \) and vice-versa.

Scenario KG, meanwhile, displays a similar pattern to scenario K in both plots with significant overlap, except in the higher mass range. The plot reveals that the best fitting models for this run span a large range in \( Q \), and here some models
Figure 7.4: Graphs of the rotation curves for representative models of each scenario. Red identifies the Hβ points, purple identifies the HI data and blue represents the stellar rotation points as in Fig. 4.4. Green is the stellar model observed velocity curve while black represents the circular velocity curve. The orange curves show the rotation curve breakdown by component: halo (solid), disc (dashed) and bulge (dotted). The scenario K and KG rotation curves display a kink in the black curve at $\sim 0.2$ kpc, owing to the inner truncation. Although not shown, the HI fit extends to 20 kpc (that is, to the outermost observed data point).
Figure 7.5: Graphs of the surface brightness profile for scenarios K and E. Orange is the bulge contribution to the surface brightness, green is the disc contribution, black is the total and red identifies the observed $R$-band surface brightness. All models from scenarios K, KL and KG show fits similar to the top panel.
Figure 7.6: Graphs of the LOS dispersion for scenarios K, KG, and E. The model dispersion is in black and the observed dispersion is in red. The bulge contribution to the dispersion is in orange and the disc contribution is in green. Fits for scenario KL are similar to those for scenario KG.
Figure 7.7: The 2-D marginal probability distribution of the disc mass versus the minimum $Q$ of the disc. Black refers to scenario K; green refers to scenario KL; red refers to scenario KG; blue refers to scenario E. The blue-outlined stars correspond to the parameters of the models chosen to test stability, and the purple circles correspond to the points used by BG97. The thick lines enclose the $1\sigma$ confidence region, and the thin lines enclose the $2\sigma$ confidence region.
Figure 7.8: The 2-D marginal probability distribution of the minimum $X$ versus minimum $Q$ of the disc. Colours and symbols are as in Fig. 7.7.
(albeit very few) extend down to below $Q = 1$. However, all models for scenario KG are confined to $X < 2$, unlike for scenario K where a large portion of the models reside between $X = 2$ and $X = 3$. Unlike for scenario K, $X$ does not appear to be significantly correlated with $Q$. However, for neither scenario K nor KG do we see $X$ drop below 1.

### 7.2.1 Bar stability of best fit models

We selected several models from the range of models found in the maximum likelihood region and evolved them forward in time to examine their stability to bars. To do this, we employ Dehnen’s (2000) $N$-body algorithm implemented by Stiff (2003), a fast multipole method that produces nearly $O(N)$ scaling. All models are evolved for 5 Gyr with a time-step of 0.5 Myr and a softening parameter of 25 pc. We use 500K particles for the disc, 50K for the bulge and 1M for the halo. The bar evolution is quantified using $A$, the magnitude of the second Fourier mode, which measures the strength of two-armed asymmetries (Considère and Athanassoula, 1988):

$$ A = \left| \sum_{j=1}^{N} \frac{\exp(2i\theta_j)}{N} \right| $$

where $\theta_j$ is the azimuthal coordinate of the $j$th particle in the disc. The bar evolution is found in Fig. 7.9.

The results show which region of the disc mass–$Q$ parameter space are most susceptible to bar formation. For scenario K, we find that much of the region is at least mildly bar unstable, with the lower right region more strongly bar unstable and the upper left only showing a very mild bar instability. The growth of all bars in this run is gradual, with bars only starting to become discernable after $\sim 2$ Gyr.
Additionally, the central hole gradually vanishes in the bar unstable models as stars stream into the centre during the process of bar formation.

For scenario KG, we find something similar – most of the available parameter space is bar unstable, with the low mass-high $Q$ part of the plot susceptible only to mild bar formation. Thus, it is not possible to distinguish between scenarios K and KG on the basis of bar stability. Therefore, we refer back to the rotation curve fits; no models for scenario KG can reproduce the stellar rotation curve because the asymmetric drift is much too large. Therefore, we reject models from scenario KG as being inconsistent with the data.

### 7.3 Testing dust extinction

To test the hypothesis that dust extinction may be responsible for the Type II surface brightness profile, we conduct a MCMC run in which the disc mass was assumed to be exponential but the luminosity profile was made to be inner-truncated, as in Eq. 3.4. In Fig. 7.4 we can see that the model does a good job of fitting the stellar rotation; as expected, the circular velocity for these models is always greater than the Hβ data, however.

Fig. 7.7 displays the marginal two-dimensional posterior probability distribution for the disc mass and $Q$ for scenario KL in green. Not surprisingly, the disc masses here are much larger than for scenarios K and KG, since there is no hole in the mass distribution. However, the radial dispersions $\sigma_R$ and epicycle frequencies are constrained by the data, so $Q$ must be lower for these models than in either scenario K or KG. The best fit region extends well below the $Q = 1$ line, which suggests that some of the best fit models are grossly unstable. Meanwhile, the $Q - X$ plot shows
Figure 7.9: The Fourier amplitude $\ln A$ of several models obtained from each scenario. The colour pattern orange-green-blue-red corresponds to highest to lowest $Q$ of the models identified by stars on Figs. 7.7 and 7.8.
that all models from scenario KL have $X < 1$, suggesting very strong instability to global modes.

To discern whether or not these models are realistic at all, we must determine whether or not galaxies with $Q < 1$ are possible. Unstable galaxies generate spiral structure that heat stars as they propagate outwards, while continual cooling could occur if gas infall onto the surface of the galaxy were substantial enough. Such gas would contribute to star formation. Fuchs (1999) has conducted a maximum disc analysis for NGC 6503 that suggests $Q < 1$ implies a star formation rate of $40 \, M_\odot \, \text{yr}^{-1}$, which contradicts the result from the observed Hα flux of $1.5 \, M_\odot \, \text{yr}^{-1}$ (Kennicutt et al., 1994). Therefore, it is unlikely that this cooling method is important in NGC 6503. In the absence of star formation, the main instability for a disc in which $Q < 1$ is not the bar instability – it is disc fragmentation. The likely evolution for such models is to eventually recombine into a single stable disc, but not before the stellar dispersions have increased dramatically.

To verify that these models are not realistic, we selected several models from scenario KL and evolved them as in the previous section to determine stability to bars. The evolution of the two-armed Fourier mode is found in Fig. 7.9; clearly, the models for scenario KL are far more susceptible to global instabilities than for either scenarios K or KG, as the growth of the bar mode is rapid and nearly immediate from the start of the simulations. We therefore rule out models from scenario KL as being unable to properly model the galaxy; this cuts against the interpretation that the Type II brightness profile is due to dust extinction.
CHAPTER 7. RESULTS: CONSTRAINING THE PHYSICS

7.3.1 Testing the outer exponential fit

Finally, we test the hypothesis that the outer 2 kpc of the surface brightness profile represents the true disc, by only fitting the interior points and the external \((R > 2\) kpc) points of the surface brightness profile, corresponding to scenario E. The outer cut is made at 2 kpc; although the regression line slope of the outer points corresponds to \(R_d = 1.18\) kpc, \(R_d\) is subtly affected by the inclination and differences in scale height. We adopt a value of \(R_d = 1.3\) kpc and fix the disc \(M/L_R\) ratio to 0.53\(M_d\); we do not include the external disc in our \(\chi^2\) calculations because the subtle dependence of the outer surface brightness on scale height and inclination makes finding a well mixed data set difficult. The inner cut is trickier; we have found that small changes to the number of inner points fit may lead to large changes in the resulting bulge parameters. To rectify this issue, we plot the residuals of the surface brightness from

![Figure 7.10: Residuals of the R-band surface brightness for the exponential disc fit shown in Fig. 4.2.](image)

The residuals show a pattern that suggests deviations from the exponential fit, indicating potential issues with the model or the data. The plot helps in validating the assumptions and improving the model fit.
a regression line fit of $R_d = 1.18$ kpc (Fig. 7.10) and only fit the points that lie above the maximum of the Type II hump. At these points, the surface brightness is most assuredly bulge dominated. The stellar rotation is fit as in scenarios K and KL, along with the HI data in outer part, and the velocity dispersion is also fit as in the other three runs.

For this scenario, we are assuming that the evolution of disc would generate a ring of star formation that could reproduce the Type II hump. The process of generating this profile would alter the rotation and dispersions, however, and we caution that the kinematics that result may not fit the data as well as the direct fits presented here using MCMC. Accounting for this would require ad hoc selections of what points to fit, or excluding one or more data sets entirely; for example, excluding the stellar rotation and only fitting the HI data outside 3 kpc. Such a process is probably no more reliable than simply fitting the full data sets, evolving a test model from the PDFs and then determining how well the rotation curve and dispersion profile fit the observed data after evolution. Because we are assuming that a bar is supposed to form in this scenario, unstable models would not rule out this interpretation. However, models with $Q < 1$ must be tested separately.

As in scenario KL, the stellar rotation curve is well fit while the circular velocity is larger than the H$\beta$ curve (Fig. 7.4). The surface brightness fit is very good over the ranges we wish to fit; the inner eight points and outer exponential are well reproduced (Fig. 7.5). The velocity dispersion profile looks different from the profile for the other scenarios because the bulge contributes significantly to the profile. Because of this, the $\sigma-$drop is not properly reproduced (Sérsic profiles do possess a slight central decline in the velocity dispersion – see Ciotti 1991 – but this feature cannot reproduce the
observed $\sigma-$drop). Thus, we posit that either an additional nuclear component is responsible for the $\sigma-$drop, or that bar formation in this scenario would generate gas inflow that could reproduce the $\sigma-$drop. The disc contribution to the velocity dispersion is comparatively muted, resulting in a lower overall disc dispersion.

From Fig. 7.7, we see that scenario E does not follow the same trend as the other scenarios; these models have lower disc masses and $Q$, suggesting a strong bar instability in scenario E, but Fig. 7.8 shows that these models have high $X$. Additionally, the minimum $Q$ in these models is reached at larger radii than in the other scenarios (typically, at $\sim 2 - 2.5R_d$), and they have more massive bulges (§8.4), which may also help inhibit the bar instability.
As before, we selected several models and tested them for bar stability; the results are found in the bottom right panel of Fig. 7.9. While these models quickly form spiral structure, bar formation is gradual at the low $X$ end of the best fit region and there is no evidence of bar formation at larger $X$, in spite of the very low $Q$. Because we do not include gas in these simulations, it is not possible to simulate possible bar destruction by gas inflow, but we can examine the length of the bar that forms. In Fig. 7.11 we plot the bar strength $A$ as a function of radius at different times for the lowest $-X$ model. This provides an estimator of the length of the bar. The bar length is $\sim 1.5$ kpc at 1.5–2 Gyr before increasing in length, which is consistent with the location of the Type II hump. We conclude that scenario E is a realistic scenario for the formation of the bar, and hence a realistic model for the galaxy.
Chapter 8

More Results: Properties of the Galaxy

In this section we delineate the numerous other results that we can glean from the MCMC runs. Because models from scenarios KG and KL are ruled out, we focus mainly on the results for scenarios K and E; however, we also emphasize some of the important results by discussing all four runs here. For completeness, results for scenarios KG and KL are found in Table 7.1 and in the relevant figures.

8.1 The disc

We begin by comparing the photometric and dispersion scale lengths. We find from scenario K that the dispersion scale length is significantly larger than the photometric scale length, as seen in Fig. 8.1; this is contrary to the equality generally assumed. The mean $R_\sigma = 2.57$ kpc is also considerably larger than the best fit $R_\sigma$ of 1.16 kpc found in §2, while the photometric scale length is slightly less than the ‘fit by eye’ value of
Figure 8.1: The 2-D marginal probability distribution of the photometric scale length $R_d$ versus the dispersion scale length $R_\sigma$. Colours are as in Fig. 7.7. The dotted line identifies the $R_d = R_\sigma$ line.
1 kpc from B89. Note that the dispersion scale length is not so well constrained; this is because of the size of the error bars in the velocity dispersion data, and a wide range of $R_\sigma$ may produce good fits, as Fig. 8.1 clearly shows. The large $R_\sigma$ is due to the low $\sigma_0$, the extrapolated central dispersion – only $37 \pm 3$ km s$^{-1}$, much less than found by direct fit in §2. In the absence of a surface brightness fit, a larger $\sigma_0$ and lower bulge $M/L_R$ ratio would produce a better fit, but the surface brightness constrains the available bulge light significantly. In scenario E, by contrast, the dispersion scale length of $R_\sigma = 0.75$ kpc is much lower than the adopted photometric scale length of $R_d = 1.3$ kpc. Here $\sigma_0$ is also very low, but the bulge dispersion is high enough to contribute much more to the dispersion profile than in the other scenarios, as is apparent from Fig. 7.6. Both scenarios suggest that the disc dispersions are not well constrained because of the large error bars on the observed data.

For scenario KL and KG, we find that $R_\sigma$ lies around 1 kpc, while $R_d$ hovers around 0.8 kpc in both cases. These results are rather more in line with the common assumption that the two quantities are about equal, and are closer to the best fit $R_\sigma$ of 1.15 kpc found in §2. This is because $\sigma_0$ is much higher in these runs, and the bulge is quite small and so cannot dominate the LOS dispersion.

For the scale height $z_d$, we find a value of $0.14 \pm 0.01$ kpc in scenario K and $0.24 \pm 0.02$ kpc in scenario E (Fig. 8.2). Both of these values are about one-fifth of the photometric scale length in scenarios K and E respectively, and are consistent with observations of the scale height in edge-on disc galaxies (Kregel et al., 2005).

The posterior values for the Kormendy parameters – the hole radius $r_h$ and the Kormendy index $\alpha$ – are well constrained; see Fig. 8.3. We find $r_h = 0.69 \pm 0.11$ kpc and $\alpha = 0.90 \pm 0.14$. This contrasts with Kormendy’s (1977) result finding that
Figure 8.2: The PDF for the disc scale height $z_d$. Black refers to scenario K; green refers to scenario KL; red refers to scenario KG; blue refers to scenario E.

$\alpha = 3$ provided a good fit to most galaxies (the same value was used by Baggett et al. 1998 for their fits). This result was, as far as we can tell, not rigorously established. A value of $\alpha = 3$ leads to a very strong hole, and therefore a very large and strong bulge is needed to contribute to the surface brightness profile. However, a large bulge would likely make the dispersion profile fit worse because the bulge is not likely to extend farther out than the highest dispersions at $\sim 300$ pc. We note that all three runs suggest $\alpha \ll 3$; scenario KL gives $\alpha = 0.56 \pm 0.10$ and scenario KG yields $\alpha = 1.06 \pm 0.11$. 
Figure 8.3: The 2-D marginal probability distribution of the Kormendy cutoff radius $R_h$ versus the cutoff index $\alpha$. Colours are as in Fig. 7.7.
Figure 8.4: The 2-D marginal probability distribution of the Sérsic index $n_b$ versus the radius of the bulge $r_b$. Colours are as in Fig. 7.7.
8.2 The bulge

The results for the bulge parameters are found in Fig. 8.4. For scenario K, we find a bulge radius $r_b$ of $0.16 \pm 0.02$ kpc and a Sérsic index $n_b$ of $1.06 \pm 0.07$, while scenario E yields $r_b = 0.35 \pm 0.02$ kpc and $n_b = 1.93 \pm 0.10$. Additionally, for scenario KL, the Sérsic index is $0.91 \pm 0.16$, and for scenario KG, we have $n_b = 0.86 \pm 0.06$. In other words, scenarios K, KL and KG suggest that NGC 6503’s bulge is close to purely exponential. In accordance with Fisher & Drory (2008), who find that bulge Sérsic index distributions are bimodal about $n = 2$ with pseudobulges occupying the lower part of the distribution, this suggests that NGC 6503’s bulge may be classified as a pseudobulge. Scenario E suggests a cuspier bulge, but even here $n_b \ll 4$, suggesting a tendency towards a pseudobulge structure as well. Note also that the regression line fit $n_b$ given in §2.1 is consistent with the derived $n_b$ for scenarios K, KL and KG.

Pseudobulges typically resemble small discs embedded in larger discs – their kinematics tends to be more rotationally dominated and their morphology resembles a flattened spheroid more than an elliptical galaxy. They are thought to be slowly generated via internal secular evolution processes – the gradual buildup of gas in the centre of the galaxy – rather than violent mergers (see, e.g., Kormendy et al. 2006).

Because NGC 6503 is an isolated disc galaxy, its evolution is likely to be more susceptible to internal secular processes rather than mergers, and it is likely, based on the Sérsic index, that the bulge formation history involves few, if any, mergers.\footnote{Other possible formation scenarios for pseudobulges are briefly reviewed in Fisher and Drory (2008); they involve gas rich mergers or gravitational interactions that can drive gas into the centre. However, because NGC 6503 is isolated we do not believe such mechanisms are likely in this galaxy.}

The bimodality that we find in Sérsic index is interesting; it can be traced to the presence, or lack thereof, of an inner truncation. The underlying disc contributes to
the surface brightness in scenario E, negating the need for the bulge to account for
the full surface brightness inside $\sim 0.2$ kpc. Hence, in this case the bulge is cuspier.
The difference is visible in Fig. 7.5, where the bulge fit for scenario E shows a steeper
slope than for scenario K, corresponding to a higher Sérsic index.

8.3 Halo cusps

The halo cusp $\gamma$ is well-constrained; the PDFs are found in Fig. 8.5. All runs suggest
that the halo is cuspy, a result that is therefore quite robust. In the case of
scenarios K and KL, using the stellar rotation curve in the inner part of the galaxy
suggests that a cuspy profile is required; scenario K yields $\gamma = 1.22 \pm 0.04$, and
scenario KL yields $\gamma = 0.91 \pm 0.18$. Scenario K yields a higher cusp value because the
disks in this run lack central mass, and therefore the halo cusp must be stronger to
generate the same force required to reproduce the rotation curve as observed. Note
from Fig. 8.5 that, although two different cusp values fit the data for scenario KL,
both suggest cusps. For scenario KG, we find that $\gamma = 1.08 \pm 0.04$. This is nearly
consistent with a NFW profile; we note that this cusp is flatter than for the equivalent
run that fits the stellar rotation curve, scenario K, but it is still cuspy, because the hole
in the centre of the disk forces a concentration of mass in the halo as described above.
Meanwhile, scenario E suggests a very strongly cusped halo, with $\gamma = 1.43 \pm 0.03$.

To check this result, we ran a MCMC chain in which the cusp value was fixed to 0,
and examined the resulting fits. The resulting rotation curve fit is found in Fig. 8.6;
the $\chi^2$ values for this run are reasonably good, but the actual fit is clearly very poor.
There is a pronounced kink in the rotation curve that is not observed, and neither the
gas nor the stellar rotation curves are properly fit. We therefore reject models with
Figure 8.5: The PDF for the halo cusp strength $\gamma$. Colours are as in Fig. 8.2.

$\gamma = 0$ as being unable to correctly reproduce the observed rotation curve, establishing the robustness of the result that $\gamma > 0$.

### 8.4 Masses and mass-to-light ratios

The halo mass is necessarily limited by the outermost HI data point at 800", so the mean halo mass we find should be regarded as a lower limit. The outer rotation curve is the primary determinant of the halo mass, so we do not expect significant variation in halo masses across the different runs. We find a halo mass of $\sim 60 \pm 2 \times 10^{10}M_\odot$ within 20 kpc across all scenarios. For scenario K, we obtain a disc mass of $3.2 \pm 0.6 \times 10^9M_\odot$ and a bulge mass of $7.0 \pm 1.4 \times 10^7M_\odot$, which is only 2% of the
Figure 8.6: A model from a MCMC run in which the cusp is fixed to 0. Colours are as in Fig. 7.4.

disc mass for this run. Results are similar for scenario KL, while the bulge mass is much lower for scenario KG. In scenario E, we find a disc mass of $3.0 \pm 0.4 \times 10^9 M_\odot$ and a bulge mass of $15 \pm 2 \times 10^7 M_\odot$, which is twice the mass found in scenario K.

Unsurprisingly, in all cases the halo mass accounts for $>90\%$ of all the galaxy mass, while the bulge mass is very small – too small to significantly impact the stability of the disc, except in scenario E. Whether or not the bulge masses are consistent with other values found in the literature depends on whether the bulge is interpreted to be a nuclear cluster (§9.2). Additionally, data from Greisen et al. (2009) suggests a HI mass of $1.3 \pm 0.2 \times 10^9 M_\odot$, with a thick HI disc inside the optical radius as well,

\[ \text{Note that the input } M_d \text{ parameter is the exponential disc mass, and is distinguished from the true disc mass because the exponential disc must be truncated in the outer regions, and because the inner truncation in scenarios K and KG also removes a large chunk of mass from the disc.} \]
Figure 8.7: The 2-D marginal probability distribution of the bulge mass versus the bulge mass-to-light ratio \((M/L_R)_b\). Colours are as in Fig. 7.7.
Figure 8.8: The 2-D marginal probability distribution of the disc mass versus the disc mass-to-light ratio \((M/L_R)_d\). Colours are as in Fig. 7.7.
suggesting that a significant fraction of the disc mass that we find is actually in the form of gas.

$M/L_R$ ratios for the bulge and disc are found in Figs. 8.7 and 8.8. In order to reproduce the velocity dispersion drop, the bulge $M/L_R$ ratios must be small, and this is in fact what we find in scenarios K, KL and KG, but not E. The first three scenarios show that the bulge $M/L_R$ is less than 1, while the disc $M/L_R$ is roughly twice that (in the case of scenario KG, it is an order of magnitude larger). We find the disc $M/L_R$ is $1.86 \pm 0.33$ and the bulge $M/L_R$ is $0.87 \pm 0.12$ under scenario K. In scenario E, we fixed the disc $M/L_R$ to be $0.53M_d$, yielding a disc $M/L_R$ of $1.60 \pm 0.21$ while the bulge $M/L_R$ is $1.23 \pm 0.11$, which, although larger than 1, is still lower than the disc $M/L_R$.

We also show the ratio of bulge light to total light $B/T$ in Table 8.1 for each scenario. This quantity is a proxy for Hubble type, and low $B/T$ values correspond to later galaxy types. The values we find are very low; in all cases, $B/T \lesssim 0.06$. This is consistent with observations of $B/T$ in late type galaxies, where this ratio frequently falls below 0.1 (Laurikainen et al., 2007; Weinzirl et al., 2009).

<table>
<thead>
<tr>
<th></th>
<th>K</th>
<th>E</th>
<th>KL</th>
<th>KG</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B/T$</td>
<td>0.045</td>
<td>±0.015</td>
<td>0.061</td>
<td>±0.014</td>
</tr>
</tbody>
</table>

Table 8.1: A table of $B/T$, the ratio of bulge light to total light, for each scenario with 1σ error bars.
Chapter 9

Discussion

We now discuss some of the more interesting results from our modelling.

9.1 Comparison with earlier work

Visible on Fig. 7.7 are the points corresponding to the models used by BG97, and they are very clearly located well outside the $2\sigma$ confidence interval. The reason for this is that $Q$ is obtained from the radial dispersion $\sigma_R$, which, in BG97’s models, is directly coupled to the surface density (cf. equations 4.5 and 4.6). Thus, BG97’s $Q$—values are determined by the disc surface density, whereas this is not the case for the WPD models. After settling, we expect their models to be roughly consistent with ours, and from their Fig. 9 the minimum $Q$ in the disc after settling declines so that it falls in line with our best-fit region on Fig. 7.7. Thus, their disc mass—$Q$ data points are consistent with our disc mass—$Q$ PDF. However, their discs do not display bar formation, while ours in this region of parameter space do. There are two possible reasons for this. First, from Fig. 7.9, we see that most of our models only develop a bar
after $\sim 2$ Gyr, while BG97 evolved their simulations for $\sim 1.3$ Gyr. Second, and more importantly, it is now known that live halos can trigger bar formation (Athanassoula, 2002); BG97 used rigid halo potentials for their simulations, so this bar formation trigger is not in play. We evolved one of our bar unstable models in exactly the way described in §7.2.1, except the number of disc particles was reduced to 40K. We found that the bar formation in this model was delayed by $\sim 1$-2 Gyr, suggesting that sufficient numerical resolution is needed to capture the correct evolution. This is consistent with the live halo triggering bar formation, because inadequate numerical resolution will fail to resolve the resonant interactions required for this mechanism to occur.

9.2 The nature of the bulge

All our results suggest that NGC 6503’s bulge is a pseudobulge. As mentioned, this suggests that the bulge was formed via secular processes in the absence of mergers. However, NGC 6503’s bulge also displays certain characteristics typical of nuclear bulges. Nuclear bulges are distinguished from normal bulges in that they are often offset from the dynamical centre of the galaxy (Matthews and Gallagher, 2002). NGC 6503’s bulge displays this phenomenon – the offset of the photometric centre lies around 5″. However, the bulge is also much larger than typical nuclear bulges by about two orders of magnitude (Walcher et al., 2005). The bulge radii found here are actually larger than found by studies of bulge-to-disc scale length correlation by roughly a factor of 2 (Courteau et al., 1996). However, the mass we find is only slightly larger than found by Walcher et al., so this is a reasonably large bulge with a surprisingly small mass. In addition, Walcher et al. (2005) find that their nuclear
clusters' velocity dispersions are similar to the $\sim 25 \text{ km s}^{-1}$ found in NGC 6503. Their nuclear $I-$band $M/L$ ratios are also consistent with the $R-$band values we find (§9.4), in that we expect $I-$band $M/L$ values to be slightly lower than our $R-$band values, which is what occurs. Thus, NGC 6503’s bulge has properties consistent with both nuclear clusters and ordinary pseudobulges.

The origin of nuclear bulges is not clear, but it is plausible that they are formed by similar processes as normal pseudobulges – i.e., secular evolution, especially in an isolated late-type spiral such as this galaxy. Walcher et al. (2006) find that repeated episodes of gas infall may contribute to nuclear cluster star formation, which is required to maintain their high luminosity and make such a scenario work. In the case of scenario E, bar formation may induce gas inflow that would generate a nuclear cluster that could in turn generate the $\sigma-$drop. It would be interesting to learn whether or not an exponential density profile is common among nuclear bulges, and therefore whether or not NGC 6503’s bulge could be categorized as being both a pseudobulge and a nuclear cluster. The most likely scenario may be a pseudobulge that has a luminous nuclear cluster at the centre dominating the observed light.

### 9.3 The cusp value

The issue of halo cuspiness has been a controversial one – it was initially thought that galaxies demonstrated evidence of cored halos (e.g., Blais-Ouellette et al. 2001), but many complications present themselves. The observational uncertainties alone are an issue, as van den Bosch and Swaters (2001) find that the difference between cusp and core rotation curves is less than the uncertainties, and early studies suffered from beam smearing. Some researchers find that the rotation curves used to probe halo
central structure are consistent with both cored and cuspy profiles (Swaters et al., 2003; Spekkens et al., 2005). More serious caveats concern the correct interpretation of the rotation curve; the presence of noncircular gas motions, as might be caused by a triaxial halo, and the fact that gas has its own asymmetric drift (which may itself be the origin of some noncircular motions) point to the complexity of the problem (Hayashi et al., 2004; Rhee et al., 2004; Simon et al., 2005; Spekkens et al., 2005; Valenzuela et al., 2007). Note that Oh et al. (2008) find that the non-circular motions in their sample cannot rectify the discrepancy.

Moreover, the interpretation of simulations and the applicability of the fitting formulae used for observed rotation curves is not well-established, because (1) the most recent simulations suggest that halos do not have a constant cusp value, but rather one that becomes shallower as $r \to 0$ (Navarro et al., 2008); and (2) both cosmological simulations and rotation curve fitting suggests that halos are not universal and that ‘hybrid’ profiles between cusped and cored may provide a better fit (Simon et al., 2005). Given the last point, it should be emphasized that the question of cusps versus cores is critically dependent on the choice of model, and the possibility that a cored profile fits the data better is not necessarily an indication of the existence of a core, especially if a shallow cusp also fits the data well.

We find that the derived cusp value depends on the density profile of the disc and the choice of rotation curve. Unsurprisingly, a stronger cusp is required to compensate for the lack of central mass in an inner truncated disc. In addition, we find a statistically significant difference in cusp values depending on which rotation curve is fit, but both cases are consistent with or cuspier than cosmological simulations. If the assumption that gas traces the circular velocity is erroneous, then the slope of the
circular velocity curve must be steeper at the centre (as seen in Fig. 7.4), and hence the derived cusp value will be larger. Thus, modelling asymmetric drift correctly is essential to obtaining the correct halo density profile; we discuss possible issues with the asymmetric drift in §9.7.

Although many cosmological simulations have suggested that galaxy halos are universal and cusped with $\gamma = 1 - 1.5$, Navarro et al. (2008) find, using very high resolution simulations, that halo cusps are shallower than previously found, with $\gamma = 0.9 \pm 0.1$ at the steepest. Our cusp values are larger than found by Navarro et al. (2008), and, like their halos, the halo of NGC 6503 is relatively isolated. The discrepancy likely can be attributed to modifications to the halo profile that would occur with the presence of baryons (Navarro et al.’s halos were pure dark matter). The effect of adiabatic contraction would be to increase the central density, but that is a simplified scenario and the full effect of baryonic physics during galaxy formation is still poorly understood. Nonetheless, Abadi et al. (2009) find that the central density still increases when more detailed baryonic effects are included. Another modification to the halo profile relates to bar formation and its effect on the central density cusp, and there has been some debate in the literature about this. Weinberg and Katz (2002) argue that bars are capable of washing out central cusps, an effect seen in Holley-Bockelmann et al. (2005) and Weinberg and Katz (2007). Meanwhile, Sellwood (2003, 2008) argues that bar formation tends to draw mass inward, increasing the strength of the cusp; Dubinski et al. (2009) find with their simulations that a live bar maintains the cusp. The discrepancy is unresolved but may be due to differences in the codes used. If NGC 6503 did once possess a bar, our results argue against bar-induced cusp flattening in this galaxy.
9.4 Mass-to-light ratios

The models we use do not account for changes in the stellar $M/L$ ratios due to stellar population changes or colour variations. In particular, Bell and de Jong (2001) and de Jong and Bell (2007) find that, in galaxy centres where redder colours prevail, stellar $M/L$ ratios tend to increase. Here, we have simply assumed constant $M/L_R$ throughout the galaxy disc in the absence of a stellar population model. From B89, $B - R$ across two disc scale lengths for NGC 6503 varies from 1.10 to 1.41, correlating with $M/L_R$ ratios of $\sim 1.3 - 2.7$ using Table 3 of Bell and de Jong (2001), which is consistent with our findings. Upper limits to the dynamical $M/L_R$ values were obtained by Broeils and Courteau (1997), who find $M/L_R \lesssim 4$, also consistent with our values.

The $M/L$ values for the bulge do not, by contrast, coincide with other bulge values found in the literature for nearby spiral galaxies. Yoshino and Ichikawa (2008) generally find that their bulge $M/L$ ratios are larger than their disc values with few exceptions; $M/L_V$ ratios are $\sim 4.5 \pm 2.4$ and $M/L_I$ ratios are $\sim 2.7 \pm 1.8$. In no case is the bulge $M/L_V$ value lower than 1, and only in a few cases does $M/L_I$ fall below 1. NGC 6503’s bulge $M/L_R$ lies significantly below nearly all the bulges in this sample. However, $M/L$ ratios are highly sensitive to the formation of massive, luminous stars, and our small $M/L_R$ values, together with the work of Wozniak et al. (2003) and Wozniak and Champavert (2006) showing that $\sigma-$drops can be caused by massive star formation, strongly suggest that the bulge of NGC 6503 has a star-forming component. Thus, the better comparison may be to the nuclear clusters of Walcher et al. (2005), who find $M/L_I$ ratios that generally lie below 1.
Figure 9.1: Ratio of velocity dispersions for NGC 6503. Black refers to scenario K, green refers to scenario KL, red refers to scenario KG, and blue refers to scenario E. The thick lines use the best fit values found in Table 7.1 for scenario K and the best fit values for scenario KG, while the thin lines correspond to the specific models identified by stars on Fig. 7.7. There is no thick line for scenario KL because the means for scenario KL typically fall between two separate likelihood peaks.

9.5 The ratio of velocity dispersions

To examine the velocity dispersion ratio for NGC 6503, we took the best fit model for NGC 6503 (that is, the means from Table 7.1) and all models used to test bar stability and calculated this ratio as a function of radius. The results are found in Fig. 9.1, showing that NGC 6503’s dispersion ratio depends critically on the assumed model – scenario K exhibits a nearly linear decline with radius in which $\sigma_z/\sigma_R > 0.6$ within two disc scale lengths, with no indication of flattening that would justify adopting a single value over the entire disc, and very little scatter. Scenarios KL and KG
display considerably more scatter, although only scenario KG displays anything close to a relatively constant ratio of 0.6 over a significant range; models from scenario KG appear to converge to $\sigma_z/\sigma_R \sim 0.6$ at small radii, but the scatter increases at larger radii. The shape and slope of the curve is critically dependent on the choice of $R_d$ and $R_\sigma$; in scenario K, $R_d \ll R_\sigma$, which tends to drag down the dispersion ratio at outer radii. Note that when $R_d = R_\sigma$ the ratio is constant by Eqs. 4.5 and 4.6, although this may not hold in an embedded disc. We find that there is little tendency toward a constant dispersion ratio in this galaxy.

### 9.6 Asymmetric drift, stability and the cusp value

The issues of cusps, asymmetric drift and stability are intertwined because all hinge on the reliability of rotation curve data and all depend on a full accounting of effects that alter the interpretation of the rotation curve. As already indicated, it is inappropriate to assume that the gas rotation traces the gravitational potential, implying that: (1) the difference between the gas and stellar rotation curves is not a measure of the asymmetric drift; (2) the gas rotation cannot be used to assess the cusp value; and (3) the proper choice of rotation curve to model is critical to finding the correct stability region. Applying Eq. 4.2 to the observed rotation curve yields theoretical asymmetric drift values of $\sim 5 - 10$ km s$^{-1}$ over the inner two scale lengths depending on the value chosen for the central radial dispersion, provided an exponential radial distribution is assumed. This is much larger than the difference between the gas and stellar rotation curves for NGC 6503, which hovers around 2 km s$^{-1}$. We have conducted separate MCMC runs in which we fit the gas rotation curve as a tracer of the circular velocity and the stellar velocity simultaneously. This only yields models
whose $Q$-values lie below 1, a function of the very low asymmetric drift implied by the nearly coincident rotation curves. Such models, when simulated, display sharp instabilities including disc fragmentation and cannot reproduce the observed properties of the galaxy ($\S 7.3$). Our preference for fitting the stellar rotation curve and finding that gas cannot be assumed to trace to circular velocity is corroborated by Pizzella et al. (2008). They analyse rotation curves of LSB galaxies and find that stellar rotation curves are much more regular and amenable to modelling than gas curves, which suffer from numerous issues including noncircular and vertical motions that affect their speeds relative to the circular velocity.

The impact on the cusp value is notable because, if gas does not trace the gravitational potential, the circular velocity is steeper in the centre than the slope of the gas rotation. Assuming that the gas trace the gravitational potential will therefore lower the inferred cusp value. This is a possible source of error in early work assessing the observed cusp value in LSB galaxies, as small changes in central slope can lead to large changes in the cusp value or imply a cored halo ($\gamma = 0$). Using the stellar rotation curve evades this problem because modelling the stellar motions already accounts for the asymmetric drift of the stars using the epicycle equations (the GalactICS models are designed to include stellar dispersions in all three spatial directions and hence account for the asymmetric drift naturally). However, accounting for the gas asymmetric drift is more difficult; it would require accounting for noncircular gas motions and dissipative processes (such as spiral formation) and is beyond the scope of the GalactICS models, which are axisymmetric and collisionless by design. A proper study of asymmetric drift requires evolving a dissipative model with a gas disc, preferably with the effect of star formation and supernova feedback.
and is beyond the scope of this work.

9.7 Alignment of the velocity ellipsoid

Eq. 4.2 is derived under the assumption that the velocity ellipsoid is aligned along cylindrical axes and that the disc is purely exponential. The former is a reasonable assumption, but there is little evidence that it holds true; the latter is, as we have seen, questionable for this galaxy. A more general version of the equation is given by Eq. 2.15. What happens when the velocity is purely spherically aligned and an inner truncated disc is used? In the case of cylindrical alignment, we obtain

$$v_a = \frac{\sigma_R^2}{2v_c} \left[ \frac{2R}{R_d} + \frac{1}{2} \left( \frac{R}{v_s} \frac{\partial v_s}{\partial R} + 1 \right) - 1 - 2\alpha \left( \frac{R_h}{R} \right)^\alpha \right]. \quad (9.1)$$

This expression differs from Eq. 4.2 by the subtraction of a positive term on the right in the brackets. Therefore, we expect the asymmetric drift to decline in an inner truncated disc.

With a spherically aligned velocity ellipsoid, we have

$$v_a = \frac{\sigma_R^2}{2v_c} \left[ \frac{2R}{R_d} + \frac{1}{2} \left( \frac{R}{v_s} \frac{\partial v_s}{\partial R} + 1 \right) - 2\alpha \left( \frac{R_h}{R} \right)^\alpha - 2 + \frac{\sigma_z^2}{\sigma_R^2} \right]. \quad (9.2)$$

This expression will match Eq. 9.1 only if $\sigma_z = \sigma_R$. As we saw in §5.2, $\sigma_z \lesssim \sigma_R$, the ratio hovering around $0.5 \sim 1$. If $\sigma_z/\sigma_R = 0.7$ then $\sigma_z^2/\sigma_R^2 \simeq 0.5$ and, as before, the asymmetric drift will decline relative to what Eq. 4.2 predicts.

One might suspect that the change is quite small, but even a few km s$^{-1}$ can have a significant impact on fitting algorithms. It would also have an impact on the interpretation of stellar rotation curves. Both expressions reduce the asymmetric drift relative to what Eq. 4.2 predicts; in particular, the addition of a Kormendy hole
reduces the expected drift significantly in the inner parts. The size of the change is only a few km s$^{-1}$, but it is still a large fraction of the predicted drift. This is seen in Fig. 9.2, in which the effects of adding a Kormendy hole and spherical alignment of the velocity ellipsoid are examined. It is noteworthy that the inner truncation can effectively wipe out the asymmetric drift over a large range in radius; a similar effect in Fig. 7.4 is seen. For this figure, we use the fiducial values found in Table 7.1. Note that the increased asymmetric drifts at larger radii are due to the large $R_\sigma$ of this
run, and the asymmetric drift expression breaks down inside \( \sim 0.5R_d \). We test cases in which \( R_\sigma \) is reduced to 1/2 of the fiducial value, which also produce good fits (cf. §7.1), and find that the asymmetric drift is reduced in the outer part. Thus, as is evident from Fig. 9.2, a combination of factors can serve to reduce the asymmetric drift nearly to the observed values in this galaxy. One should keep in mind that the sensitivity of the asymmetric drift expression to such changes means that the predicted drifts may fluctuate wildly depending on the assumptions.\(^1\)

### 9.8 The surface brightness profile

Surface brightness profiles for NGC 6503 are available in the \( B- \) and \( R- \) bands from B89 and in the \( V- \) and \( I- \) bands from Héraudeau and Simien (1996). In addition, \( K- \) band data have been made available to us before publication from Arsenault et al. All these data are plotted in Fig. 9.3, and several discrepancies are immediately apparent. The \( V- \) and \( I- \) band profiles display a Type II hump, but it extends farther than in the \( R- \) band profile. In the \( K- \) band, the Type II hump extends even farther. All bands except the \( B- \) band display an external shallower exponential, but the \( R- \) band outer exponential appears farther in than in the other bands, and in fact it is brighter than in the \( I- \) band. It appears that the \( R- \) band profile is the odd one out, but the \( B- \) band data is also discrepant.

However, a cursory check of the structural parameters implied by the other profiles suggests that our results are largely consistent with the other surface brightness profiles.

---

\(^1\)Note that imposing a spherical velocity ellipsoid would also affect the dispersions slightly, including the dependence of the dispersion ratios on radius. Thus, eq. 9.2 is an approximation, but these effects are expected to be quite small. Larger deviations are introduced by the fact that the dispersion ratios break down in the centre. In the absence of a way to quantify the change in \( \sigma_z/\sigma_R \) with radius, we assume here that this ratio is linear with radius, as suggested by Fig. 9.1.
Figure 9.3: Several different surface brightness profiles for NGC 6503. The $B$– and $R$–band profiles in blue and red, respectively, are from Bottema (1989). The $V$– and $I$–band profiles in purple and orange, respectively, are from Héraudeau and Simien (1996). The $K$–band profile is in green, from Arsenault et al. and the black profile is 2MASS data. For clarity, error bars on the $K$–band data are not shown.

profiles. The slope of the outer exponentials in the $V$– and $I$–bands is very close to the 1.18 kpc obtained in the $R$–band, and the slope of the declining part of the Type II hump for the $V$–, $I$– and $K$–band profiles is consistent with the value of $R_d = 0.71$ kpc that we find.

9.9 Evidence for a bar

New research from Freeland et al. (2009) published after initial submission of this thesis suggests that NGC 6503 may possess a nuclear disc of radius $\sim 100$ pc and an
end-on bar. This scenario is consistent with a ring of star formation at the location of
the Type II hump, and it would explain the $\sigma$—drop, because of both bar kinematics
and star formation at the centre caused by bar induced gas inflow. These authors note
that the position angle and ellipticity profiles suggest the presence of a nuclear disc,
while deprojection of their $H$—band image of NGC 6503 shows a bar like structure
perpendicular to the major axis with clear spiral arms emanating from the ends.
Combined with the Type II surface brightness profile and diagnostics from Bureau
and Athanassoula (2005) suggesting that end-on bars may generate $\sigma$—drops, this
argues compellingly for a bar in NGC 6503. Notably, they estimate the age of the
star formation ring at $\sim 0.5$ Gyr, which is well below the timescale for bar destruction
determined by Bournaud et al. (2005) of $\sim 1$ Gyr.

The presence of a bar in NGC 6503 is unlikely to change our main conclusions
because it is end on. There is no clear kinematic indicator of bar structure in the
stellar or ionized gas rotation curve; we expect that the measured values are largely
unaffected by the bar since only a small region of the slit used for observation would
have intersected the bar. The $\sigma$—drop would be partly due to the bar, and hence the
bulge would not need be as large to generate the $\sigma$—drop. Thus, the bulge may not
be as large and have as low a density as found in chapter 8. Other bulge parameters
are less likely to be affected; in particular, the conclusion that there is a star forming
component in the centre is unaffected.
The Bayesian/Markov chain Monte Carlo technique appears to be well suited to modelling disc galaxies. We find models for NGC 6503 that satisfy the observational constraints using this technique, and are able to constrain the input parameters. We do this for four separate cases: an inner truncated disc, an inner truncated dust model, a model in which gas traces the gravitational potential, and a model in which the outer exponential represents the true disc. We find that the first of these provides the best fit to the data and is most likely to offer the bar stability required to correctly model the galaxy, and we also find that the last of these provides a realistic model for the galaxy. Using an exponential disc leads to highly unstable models, ruling out dust extinction as the cause of the surface brightness profile. We also find that the gas rotation cannot trace the circular velocity, as that makes it impossible to correctly reproduce the stellar rotation. Further properties of the galaxy can be discerned; the bulge is a pseudobulge, and the bulge $M/L_R$ is lower than the disc $M/L_R$, suggesting a star forming component that is probably responsible for the $\sigma$–drop. We also find that the halo must be cusped, a result that is robust to all
fitting methods. The Bayesian/MCMC technique used to discover these results is robust and flexible; although it suffers from high computational cost, we believe it is an effective tool for fitting galaxy models in complex parameter spaces, and its versatility allows investigations of a wide range of galaxy properties.

10.1 Improving the MCMC technique

A great deal of research is currently being undertaken to determine how to improve the classic MCMC approach to improve convergence, efficiency and parameter estimation. In particular, a long standing obstacle to employing MCMC for many problems lies in the computational cost, as we have found; a typical MCMC run takes roughly 2 months on a desktop computer. Another problem is to verify that the resulting PDFs reflect the true posterior distribution, complete with possible substructure and multiple minima. Some of these have been addressed by the stepwise simulated annealing/covariance approach described in §6.6.

Several possible improvements present themselves. An alternative to pure MCMC techniques is Hamiltonian Monte Carlo (HMC), in which the chain moves along trajectories of constant energy and momentum. HMC has already been shown to efficiently determine cosmological parameters by Hajian (2007). Currently, the best option for automatically determining the joint posterior distribution may be a parallel annealing technique (Gregory, 2005). In this method, multiple MCMC chains are run simultaneously from different starting points, and at specific intervals the concurrent chains ‘swap’ trajectories with a random probability. This method appears to be the best at capturing substructure that is otherwise missed by standard MCMC techniques, and also prevents a single chain from getting stuck in a local minimum. It is not clear
that this method can be combined with the method described in §6.6, however.

10.2 Future work

There are several projects that stem from this work.

1. As noted, stellar kinematics provides the best way to glean the halo density profile. MCMC lends itself very well to modelling halo profiles using GalactICS, because it is difficult to obtain the halo density profile from a stellar rotation curve while simultaneously accounting for asymmetric drift and line of sight integration effects. Our method does this automatically. Given the apparent discrepancy in simulated and observed halo profiles, an excellent line of enquiry is to use MCMC to derive cusp values for a sample of disc galaxies for which the stellar rotation is available. Such data are available from Héraudeau and Simien (1998), Vega Beltrán et al. (2001) and Pizzella et al. (2008), and other papers by R. Bottema (Bottema, 1988, 1992). A sample of about 10 galaxies are immediately amenable to the technique described in this thesis, as extended rotation curves and surface brightness profiles are readily available in the literature.

2. There are as yet no studies of the stability properties of inner truncated discs. It would be interesting to examine two things: the stability of inner truncated discs to bars, and whether the inner truncation is stable in nonbarred discs. We have verified the latter in certain cases, but only for the purpose of ensuring that the modifications we made to GalactICS are sound. A systematic study of this issue would be more illuminating, but will require careful attention to the $N$–body parameters in the centre.
3. A study of dispersion ratios is warranted. Section 9.5 demonstrates that constant dispersion ratios are not implied by the observations, contrary to what is often assumed. Thus, it would be useful to know of such patterns in other galaxies for which dispersions have been measured.

The first of these is currently underway.
Bibliography


BIBLIOGRAPHY


BIBLIOGRAPHY


Appendix A

A Proof that MCMC Yields the Correct PDFs

Here is a proof that the Metropolis-Hastings algorithm converges to the correct posterior probability distribution, taken from Gregory (2005). The Markov chain that results from the Metropolis-Hastings algorithm will hit a stationary distribution, provided three conditions are met. The first is that the chain be irreducible, that is, the chain must be able to hit all states with probability > 0, given enough time. The second requirement is that the chain be aperiodic. The final requirement is that the chain be positive recurrent.

We first show that, if a Markov chain and a posterior distribution $P(p_i|D)$ satisfy the detailed balance equation, $P(p_i|D)$ is the Markov chain’s stationary distribution. The detailed balance equation states that

$$P(p_i|D)P(p_{i+1}|p_i) = P(p_{i+1}|D)P(p_i|p_{i+1}).$$  \hspace{1cm} (A.1)

The joint probability of $p_i$ and $p_{i+1}$ is $P(p_i, p_{i+1}) = P(p_i|D)P(p_{i+1}|p_i)$. Integrating
APPENDIX A. A PROOF THAT MCMC YIELDS THE CORRECT PDFS

this joint probability, we obtain

\[ \int P(p_i|D)P(p_{i+1}|p_i)dp_i = \int P(p_{i+1}|D)P(p_i|p_{i+1})dp_i \]  \hspace{1cm} (A.2)

\[ = P(p_{i+1}|D) \int P(p_i|p_{i+1})dp_i \]  \hspace{1cm} (A.3)

\[ = P(p_{i+1}|D) \]  \hspace{1cm} (A.4)

which is simply the posterior distribution for \( p_i \).

We now show that detailed balance holds for the case of the Metropolis-Hastings algorithm. The probability of drawing a sample \( p_i \) is given by \( P(p_i|D) \), and the probability of accepting the sample \( p_{i+1} \) is given by the transition kernel, \( P(p_{i+1}|p_i) = Q(p_{i+1}|p_i) \min(1, r) \), where \( r \) is given by Eq. 6.4. We then have that the joint probability \( P(p_i, p_{i+1}) \) is given by

\[ P(p_i|D)P(p_{i+1}|p_i) = P(p_i|D)Q(p_{i+1}|p_i) \min(1, r) \]  \hspace{1cm} (A.5)

\[ = \min(P(p_i|D)Q(p_{i+1}|p_i), P(p_{i+1}|D)Q(p_i|p_{i+1})) \]  \hspace{1cm} (A.6)

\[ = P(p_{i+1}|D)Q(p_i|p_{i+1}) \min(1, r^{-1}) \]  \hspace{1cm} (A.7)

\[ = P(p_{i+1}|D)P(p_i|p_{i+1}) \]  \hspace{1cm} (A.8)

and therefore

\[ P(p_i|D)P(p_{i+1}|p_i) = P(p_{i+1}|D)P(p_i|p_{i+1}) \]  \hspace{1cm} (A.9)

which is the detailed balance equation again. Thus, \( P(p_i|D) \) is the stationary distribution for the Markov chain generated by the Metropolis-Hastings algorithm, and the proof is complete.