# Modelling the Milky Way stellar halo


Francesco Fermani
Merton College
University of Oxford

A thesis submitted for the degree of
*Doctor of Philosophy*
Trinity Term 2013

# Modelling the Milky Way stellar halo

Francesco Fermani
Rudolf Peierls Centre for Theoretical Physics
Merton College, University of Oxford

**Abstract**

We motivate the importance of understanding the kinematics and dynamics of the Milky Way stellar halo both in unravelling the formation history and evolution of our host Galaxy and in the more general context of galaxy dynamics.

We present a cleaned picture of the kinematics of the smooth component of the stellar halo: we develop a method to quantify the average distance error on a sample of stars based on the idea of Schönrich et al. (2012), but adapted so that it uses velocity information only on average. We use this scheme to construct an analytic distance calibration for Blue Horizontal Branch (BHB) field halo stars in Sloan colours and demonstrate that our calibration is a) more accurate than the ones available and b) unbiased w.r.t. metallicity and colour. We measure the rotation of the smooth component of the stellar halo with a tool-set of four estimators that use either only the l.o.s. velocities or the full 3D motion. From two samples of BHB stars from the Sloan Digital Sky Survey, we favour a non-rotating single halo. We critique conflicting results in the literature based on similar samples and trace back the disagreement (either in the sign of rotation or in the morphology of the halo) to sample contaminations and/or neglect account of the halo geometry.

We propose a scheme that generalizes any isotropic spherical model to a model where the potential is axisymmetric and the distribution function is a function of the three actions. The idea is to approximate the Hamiltonian as a function of the actions with a library of quadratic fits to surfaces of constant energy in action space and then make explicit the dependence of the energy on the three actions in the ergodic distribution function. The transparency of the physics implied by the model we achieve, should make it possible to combine our spheroidal models to the $f(\mathbf{J})$-models of Binney (2010) for the disks and of Pontzen & Governato (2013) for the dark-matter halo, and obtain a complete actions-defined dynamical model of the Milky Way Galaxy.

A thesis submitted for the degree of Doctor of Philosophy
at the University of Oxford
Trinity Term 2013

# Acknowledgements

I thank Prof. James Binney for stimulating discussions during the preparation of this thesis. It is a pleasure to acknowledge the care he has always invested in providing detailed feedback on my written pieces of work: if any clarity has been achieved in this work, I owe it to his helpful comments and constructive suggestions. My collaborator, Ralph A. Schönrich, has been an incomparable mentor: his hard training has been invaluable in obtaining most of the results presented in this thesis. John Magorrian, Heather Morrison and the members of the Oxford Dynamics group generously made available their time to discuss early results and provided helpful critiques.

In particular, I am grateful to the following people for technical support. My calculations have greatly benefited from the codes shared by Prof. James Binney for the computation of multi-dimensional integrals, the flattened Isochrone potential and angle-action coordinates in the isochrone and Stäckel potentials. Paul J. McMillan provided the code for the Torus machine and support throughout its use. Xiangxiang Xue kindly provided her SDSS catalogue of blue horizontal branch stars. Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is http://www.sdss3.org/. Alis J. Deason made available on request the polygonal to mask out the Sagittarius stream. Mason Alexander Porter and Antonis Papachristodoulou gave up their time to discuss a numerically convenient implementation of Floquet analysis to extract orbital frequencies.

I acknowledge financial support from the Science and Technology Facility Council (UK) and from Merton College, Oxford during the preparation of this work.

# Derivative Publications

The following publications have arisen out of the work of this thesis:

Parts of Chapter 3 and 4 have appeared in:
Fermani F. & Schönrich R. A., 2013, Mon. Not. R. Astron. Soc., 430, 1294

Parts of Chapter 3 and 5 have appeared in:
Fermani F. & Schönrich R. A., 2013, Mon. Not. R. Astron. Soc., 432, 2402

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

The understanding of our host Galaxy's morphology, dynamics and evolutionary history is a key task in Astronomy. Both for being the only cosmological structure that will practically ever influence human beings and since it is the closest laboratory to understand astronomical phenomena. The highly representative nature of the Milky Way by virtue of its position in the Schechter function plot moreover, makes it worthy of study in all details.

Within a survey of our Galaxy, we can resolve single stars: with a single epoch observation we identify the angular position on the sky and the radial velocity via redshift measurement and with repeated observations of the same object we track the angular displacement and hence the tangential motion. The stellar component of a galaxy though, is a mixture of stars of different types (due to their mass) and of a different formation history (due to the environment they formed in). Due to the phase-mixed status of the system and the lack of informations such as mass and age, chemistry and dynamics are therefore the only way to resolves the different components.

Wide observational programs have released in the past two decades enormous amounts of data (astrometric, Hippacros; photometric, DENIS, 2MASS, SDSS; and spectral GCS, RAVE, SEGUE) with the ambitious goals of mapping the Galaxy structure, verifying its putative dark matter content and unravelling its formation history. Three critical issues then arise in this attempt: on one hand the instrumentations available have technical limits and are capable of spanning only a fraction of the observable space. Therefore we need to understand how much this portion is representative for the whole sample and how confident we can be in inferring any general conclusions on these bases. Secondly, different surveys map a different observable space and provide different *types* of information. A theoretical scheme that glues together complementary constraints is then essential to assemble a single coherent

picture. Finally, one has to find a way to represent the results that make them easy to analyse, compare and interpret.

## 1.1 The Milky Way Galaxy

The fact that the Milky Way is a discrete system of stars dates back to the 1610, when Galileo invented the telescope and first resolved distinct objects within the extended luminous patch in the Sky referred as the *Via Lactea*.

Inferences on its shape antedated technological capabilities: in the $18^{\text{th}}$ century, Kant, inspired by the work of Wright, concluded with pure reasoning that the Milky Way has a disk-like structure.[1] Observational confirmation came in 1785, when Herschel produced the first map of the Milky Way by counting stars at successive limits of apparent magnitude. The result was only qualitative though, due to the lack of any intrinsic luminosity information. Further, neglecting the existence of absorption in the disk, made him misestimate relative distances and place the Sun near the centre of the Galaxy. In the $20^{\text{th}}$ century, multi-epoch observations made possible to measure the tangential motion of stars on the Sky (proper-motion) and hence the definition of an absolute distance scale. All the same, misestimates of the extinction within the disk, lead to conflicting conclusions on the scale of the Galaxy and our position within it (underestimated absorption, Kapteyn 1922, Milk Way size: $\sim 10\,\text{kpc}$; distance of the Sun from the Galactic Centre: $\sim 0$), (overestimated absorption Shapely 1918, 100 kpc and 15 kpc respectively). A further technological advancement was needed to discern between the possible models: Bahcall & Soneira (1980) pioneered the assembling of computer programs to quantify the effect of extinction on star counts and first built a quantitative coherent picture of the Milky Way.

By no means our Galaxy is special: in structure and kinematics, it is qualitatively indistinguishable from external spiral galaxies. This realisation required historically two fundamental steps: we first had to identify our Galaxy as not the entire Universe, allowing for the existence of comparable but separate stellar systems, and then discover the similarity among them. In 1923, with a 100-inch telescope, Hubble observed variable stars in the Andromeda Nebula (M31) associable for their oscillatory brightness with Cepheid variable stars: he estimated a distance three times what the size of

[1] Kant had shown that it is the balance between the gravitational attraction between the Sun and the Planets and the ordered revolution of the latter to cause the Solar System planar shape. By analogy, he then argued that the same mechanisms of inward gravitational pull and systematic rotation apply within the stars in the Galaxy and balance out as in the Solar System to produce a disk-shape.

the Milky Way was believed to be at the time ($\sim 100$ kpc Shapely 1918, which is $\sim 3$ times the most recent measurements) and demonstrated that the two stellar systems were in fact distinct. Lindblad (1927) had then the revolutionary idea to infer the structure of the Galaxy from the motion of its constituents, in essence quantifying the ideas of Wright and Kant. Oort (1927) followed up immediately after compiling a comprehensive study of the kinematics of the Milky Way: the unflattened central spheroidal as well as the flattened extended disk he concluded for our Galaxy, were the same features observed in edge-on external spiral galaxies; ultimately demonstrating that our Galaxy is qualitatively identical to many other comparable galaxies.

Finally, the Milky Way is not only made of stars of course: the first confirmation of the presence of matter in the galaxy not just in the form of stars is due to Ewan, Purcell, Christiansen, Muller and Oort, who in 1951 measured independently the 21 cm Hydrogen emission line. The presence of and the possibility to observe atomic Hydrogen resulted immediately in a major advancement: the 21 cm line is unaffected by dust, hence suffers no extinction and it is therefore ideal to map the large-scale structure of the galaxy. The spiral structure within the Milky Way for example, was first evident in the Hydrogen map of Oort, Kerr & Westerhout (1958).[2] Further Hydrogen maps started the quest for dark matter: the mass required to explain the gas observed orbital velocity does not match the one imputable to visible objects; it is in clear excess and the discrepancy rises smoothly with distance from the centre of the galaxy.

### 1.1.1 The current picture

The Milky Way morphology is only to zero-th order described by a disk and a spheroid (*bulge*) at its centre. In the $[\alpha/\mathrm{Fe}] - [\mathrm{Fe/H}]$ plane, the disk is bimodal (Lee et al. 2011), indicating two components of different age: a *thin* and a *thick* disk. The idea of a double disk dates back to Gilmore & Reid (1983), who noted that the space density as a function of distance z from the plane for main-sequence stars shows a double exponential trend. The age-metallicity distribution revealed indeed that disk stars have two different sets of physical properties: a component younger than 10 Gyr,

[2]Understanding the dynamics of spirals proves to be a very challenging task. The characteristics to take into account are several: spirals can be seen as (i) global modes of smooth disks dynamically cool outside and hot in the inner region (Bertin & Lin 1996); (ii) recurrent transients (Sellwood & Carlberg 1984; Sellwood 2010a); (iii) responses to density fluctuations (Zibetti, Charlot & Rix 2009) or (iv) arising from non-linear spiral dynamics (Tsoutsis et al. 2009; Athanassoula, Romero-Gómez & Masdemont 2009). All the simulations agree anyway in the need of a constant supply of new stars on near-circular orbits to maintain spiral activity, hence the necessity of a gas layer.

with metallicity greater than -0.4, a scale height of 0.3 kpc and star formation still ongoing (*Thin Disk*) and a component older than 10 Gyr, with metallicity peaking at -0.6 and a scale height of 1.4 kpc (*Thick Disk*). Kinematics then reinforced this dichotomy showing a general trend of higher velocity dispersion for the supposed thick disk stars compared to the candidates for the thin disk (Binney & Merrifield 1998). Even though a growing body of data coming from different surveys support the division in thin and thick disk (see Steinmetz 2012; Ivezić, Beers & Jurić 2012, for a review), the debate does not seem to be settled yet (Bovy et al. 2012).

21 cm line observations (Liszt & Burton 1980), near-IR photometry (Blitz & Spergel 1991) and red clump stars (Stanek et al. 1994), all provide direct evidence of triaxial peanut-shape structure (*bar*) within the bulge of our Galaxy. The latter is almost completely made of stars that are at least 5 Gyr old with a wide range of metal abundances (two peaks at [Fe/H] = 0.3 and -0.25, Barbuy et al. 2008), characteristics which open to a debate on its formation history. A scenario of hierarchical formation (Kauffmann, Guiderdoni & White 1994; Baugh, Cole & Frenk 1996; Freeman & Bland-Hawthorn 2002) and a formation from growth of disk instabilities (Combes & Sanders 1981; Kormendy 1993; Courteau, de Jong & Broeils 1996; Shen et al. 2010) appear both viable.

A fourth component, the *stellar halo*, has been introduced after discovering that the metallicity distribution of the Galaxy is bimodal. Kinman (1959) found that metal-rich globular clusters formed in more flattened and more rapidly rotating environments in comparison with the more metal-poor globular clusters. This dichotomy was later confirmed by Zinn (1985) and Armandroff (1989), who observed a clear bimodality in the globular clusters metallicity distribution, advocating the need for a new component different from the disk in both shape and formation history.

### 1.1.2 The stellar halo

The stellar halo is the component of our Galaxy spatially coincident with the postulated dark matter halo[3] and includes both a smooth distribution of stars and a significant fraction of spatially organised substructures (Newberg et al. 2002)[4]. It

[3]The presence of dark matter in our Galaxy is inferred by analogy from other external galaxies, where HI rotation curves advocate the presence of more mass than the one visible, with this discrepancy growing smoothly with radius (Faber & Gallagher 1979).

[4]Before Newberg et al. (2002), the stellar halo was modelled as a pure smooth component via power-law fits (Oort & Plaut 1975) or de Vancouleurs profiles (de Vaucouleurs 1977); with the visualization of the *field of streams* by Belokurov et al. (2006), this option has been definitely ruled out.

is the only one of the four components to have a negligible mass w.r.t. the local potential. Stars belonging to the stellar halo are therefore tracer objects and given that they are even more sparse than in the disk, a collisionless system. None the less, this "light component" can provide profound insights on the formation history of our Galaxy and constitute a direct tracer of its dark matter content.

The presence of numerous streams in the halo (Belokurov et al. 2006; Grillmair 2006; Newberg, Yanny & Willett 2009) is itself evidence for a history of formation by infall of protogalactic fragments: close parabolic encounters of two galaxies generate brief and violent tidal forces which strip a tidal tail from the less massive galaxy. These distortions evolve in narrow elongated features made of stars and gas forming a near-side spiral arm extending toward the satellite galaxy (Toomre & Toomre 1972). Moreover, embedded in the streams is knowledge about the potential of the Milky Way as a whole so that they can be used to constrain the mass distribution within it and trace the global gravitational field of the Galaxy. In particular, streams can probe the distribution of dark matter: the idea is that if we subtract from the total potential the contribution from the stellar component, we isolate the dark matter. Poisson's equation then provides straightforwardly the density distribution and Eddington (1916) inversion provides a distribution in phase space of dark matter.

Preliminary analyses produce conflicting pictures of the dark halo, but ultimately prove we have yet to scratch the complexity of its structure: the kinematics of the Sagittarius dwarf tidal tails for example, seems to be consistent at the same time with a spherical dark halo (Ibata et al. 2001; Majewski et al. 2003), a prolate distribution (Law et al. 2004; Helmi 2004), an oblate one (Johnston, Law & Majewski 2005) and with a triaxial halo potential (Law, Majewski & Johnston 2009; Law & Majewski 2010). A growing body of techniques is being manufactured to distinguish between the different possibilities (e.g. Johnston et al. 1999; Varghese, Ibata & Lewis 2011; Eyre & Binney 2011), giving hope that the dark matter profile will be clinched in the near future.

The phase-space distribution of the smooth component of the stellar halo encodes information on the formation history of the Galaxy: a transition into the degree of flattening with radius (from a flattened inner to a more spherical outer halo) or any duality in the morphology can suggest for example, a halo formation in different stages and via different physical processes (Hartwick 1987; Preston, Shectman & Beers 1991). The mean-streaming motion as a function of galactocentric radius can inform us on the average infall direction and rate of mergers: dynamical friction is proportional to the relative velocity between the infalling merger and the main galaxy

at low velocities and then decays monotonically for larger relative velocities (Quinn & Goodman 1986; Byrd, Saarinen & Valtonen 1986). The disk has a strong and relatively narrow distribution in azimuthal velocity: thus, if the merger has a streaming motion aligned with the one of the disk, dynamical friction acts more strongly upon it than if it was rotating in the azimuthal direction oppositely to the disk. In particular if the mergers are even between pro and counter-rotating, one predicts a counter-rotating stellar halo (Norris & Ryan 1989). Given also that dynamical friction is inversely proportional to the distance between the two encounters, if there is a significant number of mergers, one expects a gradient in the mean streaming motion as a function of radius (Murante et al. 2010). Within the picture of hierarchical formation, there is still space for formation by dissipation (Eggen, Lynden-Bell & Sandage 1962; Freeman & Bland-Hawthorn 2002). In this scenario, the inner part of the halo is populated by stars from evaporated clusters, once associated with the disk. Then, given that very little gas and dust is found in the halo[5] and almost no star formation takes place in situ, the chemical profile of the stars in the inner halo can be compared with the one in the disk and provide a unique test for stellar evolution models and a measure of the contribution of the interstellar medium in the chemical enrichment of the disk stars.

The relaxed and unrelaxed component of the stellar halo, namely a smooth stellar halo and the substructures within it, have yet to be disentangled though. In particular the quest for substructures is strongly limited by our ignorance on the smooth background to subtract and hence what is an overdensity and what is noise arising from the discreteness of the data. The unavoidable first step is therefore a model for the smooth component of the stellar halo, which for simplicity we will refer hereafter just as *stellar halo*. Unfortunately, none of the profiles associated with the smooth component of the stellar halo have been firmly pinned down either. The best attempt of a density distribution is a broken power-law with conflicting steepness and flattening parameters though (e.g. Carollo et al. 2007; Jurić et al. 2008; Smith et al. 2009). The rotational profile is an even more controversial measurement with one analysis suggesting the opposite thing from the other and some inferring the existence of a double halo (Chiba & Beers 2000; Sirko et al. 2004; Carollo et al. 2010; Deason, Belokurov & Evans 2011a). Also, the degree of velocity anisotropy of the stellar halo

[5]The mapping of gas clouds in the halo could serve as a further probe of formation history given the different possible origins of those clouds (Maller 2005). The problem is degenerate if the phase-space distribution of the stellar component is unknown though.

appears to be a parameter with a short time-scale of evolution in the literature (e.g. Sirko et al. 2004; Smith et al. 2009; Kafle et al. 2012).

The debate on the morphology and kinematics of the stellar halo has its roots in several factors: on one side obtaining a dataset that spans the required volume of real space and is a fair sampling of the overall distribution is a problem at the edge of the current technological capabilities. The signal depletes with distance, due to both structural limits of the instruments and mere geometric effects. The chemical identity is intrinsically harder to track down due to the halo being extremely metal-poor ($\langle[\mathrm{Fe/H}]\rangle \sim -1.9 \pm 0.5$). Finally, observations are spatially entangled making virtually any halo samples contaminated by disk stars[6] and by streams to a certain extent. In essence we are missing a coherent framework to handle and interpret the available data, i.e. a model.

## 1.2 Galaxy modelling

The first steps in modern galaxy modelling were taken by Bahcall & Soneira (1980), who assumed that the Galaxy is built up of components like those seen in external galaxies, and used this similarity to discriminate among the copious statistically equivalent solutions to the star-count equations from which their models were derived. Caldwell & Ostriker (1981) improved the modelling strategy by assuming no mass-to-light ratios (and therefore avoiding the problem of deciding whether and where there is any dark component) and then choosing among different mass distribution models by minimizing $\chi^2$ given known kinematic measurements (Oort constants, tangent-velocity curve, escape velocity at the Sun) with their estimated associated errors. The consequent step forward was then to make large-scale structures consistent with Newton's laws of motion. A good, though not definative, attempt has been the Besançon model by Robin et al. (2003): they assumed four main populations (thin disk, thick disk, stellar halo and outer bulge) and a model of extinction. Then the scale height of an isothermal relaxed population was constrained by its velocity dispersion and Galactic potential through the Boltzmann equation. Dynamical self-consistency was therefore approached but not fully reached as the constraints were only in the vertical direction and dark matter was not taken into account, ruling out the possibility of a coherent picture at large-scales.

[6]Vice versa, Binney (2012b) made clear that a mature model of the disk requires knowledge on the level of contamination from halo stars, which entangle in observations with disk ones.

The final goal is therefore a fully dynamical model: this means a model where the dynamics of the observed objects in it is fully justified by the components the model contains. The key assumption to be made is that the baryonic component of the Galaxy is in *equilibrium*. Otherwise any distribution of dark matter would be consistent with any phase-space distribution of the baryonic matter. Bar and spiral arms in the disk and streams in the halo are a good reminder that a steady state is only an approximation but one can definitely rule out large-scale contractions or expansion. The secular evolution of the Galaxy's disk and bulge through scattering of stars and dark matter particles is driven by these non-equilibrium features[7] and it is taken into account in models by perturbing the Hamiltonian of an equilibrium model. Finally, disentangling the formation history in the halo translates into determining the collisional term in the Boltzmann equation via the Fokker-Plank diffusion theory, which ultimately requires a sensible initial guess: the only intellectually defendable one is an equilibrium model.

### 1.2.1 From models to physics

When conceiving a dynamical model we require it to achieve a realistic description of the system, but we also want an immediate intuition about the physics implied by the model. It is therefore essential to identify the projection which makes analysis easy and allows one to identify the distinctive features of a phenomenon. Jeans' theorem (Jeans 1915) states that for a well-mixed stellar population, the phase-space density $f(\mathbf{x}, \mathbf{v})$ is a function of the isolating integrals of motion only. The advantage of the integrals is that they encode information on the whole structure of the galaxy other than a mere snap-shot of the local distribution. May & Binney (1986) for example, showed that assuming a potential consistent with chosen galactic rotation curve, Oort limit and local escape speed, the phase-space density of a spherical isotropic $R^{\frac{1}{4}}$ Galaxy could be determined for up to 75% of the system using only observations of stars passing through the Sun.

Effectively, there is not much confusion on which set of integrals one can use: Binney & Spergel (1984) showed that action integrals are a natural choice of isolating integrals for regular orbits and given that the fraction of phase-space that is occupied by either irregular or resonant orbit is small (Binney & May 1986), they are a natural choice for the whole system. They constitute a complete and well defined set of reference (Jeans's strong theorem) and are specific to a star's orbit: the radial

[7]For an updated review on dynamics and evolution of disk and bar, see Sellwood (2010b).

action $J_r$ measures the radial oscillation, the vertical one $J_z$ quantifies the oscillation perpendicular to the equatorial plane and $J_\phi$ is the angular momentum. On the contrary there is still debate on the classical integrals such as $L$ or $E_z$ as third integral of motion (Contopoulos 1963; Ollongren 1965; Gebhardt et al. 2000) so that the only viable scheme to obtain $I_3$ is indeed action computation. This and the fact that one of the first two integrals must be the angular momentum anyway, makes it a conceptual muddle to choose anything other than the triplet of actions as full set of integrals. Moreover, one practical advantage in using actions is that they are adiabatically conserved: invariance of actions under slow variations of the potential, implies for example that sub-structures are more easily identifiable in action space: stars from a single progenitor preserve similar values for their actions even when their spatial distribution has turned featureless due to phase-mixing[8] (Helmi et al. 1999; McMillan & Binney 2008).

The effectiveness of action based models is steadily consolidating for both the disk (Binney 2010, 2012b) and the dark matter halo (Pontzen & Governato 2013); the virgin ground in these extends, seems to be the stellar halo.

## 1.3 Structure of this work

This thesis divides into three main parts: we first construct what in essence is a reproducible laboratory, and outline techniques to make consistent measurements of the kinematics of the stellar halo. While in principle these methods are suited to study any galactic stellar halo, current status of data implies that we will always refer to the Milky Way stellar halo. Within this framework we trace the stellar halo rotational profile and critique the consistency of all the comparable results in the literature on this topic. In the last part, we present a new class of dynamical models for spheroidal components.

### 1.3.1 Part I: Methodology & Techniques

**Synthetic catalogues** Synthetic catalogues are the essential laboratory in which to test models and speculative theories, but also quantify how much bias an uncertainty

[8]By *phase mixing* one means the process for which the local density around any comoving point in phase-space remains the same due to incompressibility of phase-space flow (Boltzmann equation), but the coarse grained density evolves towards a stationary state. Again, the assumption that galaxies are phase-mixed is only an approximation as they are at most few dynamical-times old in their outer parts and because mergers and star formation continuously add new stars (Tremaine 1999).

in the data introduces in a certain measurement and to what extend the selection function of a survey distorts the recovered picture. We motivate this case and focus on what, in the industry of constructing synthetic catalogues, can be considered in abstract of observational input: the sampling of a model.

Before getting into the details of model-sampling, we use this chapter to recall the fundamental notion of distribution functions, which we will use throughout this thesis. Then, we recall the three main techniques to sample a probability distribution, namely inverse sampling, acceptance/rejection method and Markov Chain Monte Carlo sampling, and focus our discussion to the essential technical requirements these scheme have to satisfy for use in Astronomy. In particular, we illustrate some properties of the Monte Carlo sampling method in order to provide a complete, yet clean and quantitative, picture of which refinement of the main algorithm is capable of achieving which performance.

We speculate that synthetic catalogues can be used to define the optimal survey size, given a specific scientific question, providing appealing application for budget control. Some very rudimentary schemes are proposed.

**The phase-space of a stellar system** The position and kinematics of stars can provide substantial informations on the galaxy morphology and formation history. In this chapter we exemplify which kinematic measurements provide insight in which aspect of the galaxy and in particular of the halo. The main complications for stellar halo measurements are the noisiness of the data (signal-to-noise decreases with distance), the unfavourable geometry of the halo (the tangential motion of halo stars requires systematically longer periods of surveying in order to be detailed above the noise) and the metal-poor chemical identity of its tracers, which complicates the estimate of the radial velocity of a star. Therefore, the data most often available for stars in the deep halo (distances greater than 10 kpc from the centre of the Galaxy) are angular position, inaccurate distance, radial velocity and tentative proper-motion of the size of the associated error.

We start our work by compiling a library of examples to show/understand how distance errors propagate in kinematics and in which way they can distort the picture we infer from the data. Then, inspired from an idea of Schönrich, Binney & Asplund (2012), we present a general method to quantify the fractional average distance error of a distance calibration on a particular sample. On a set of control cases, the method retrieves the correct distance misestimate within 2%. Other than a useful test, this

method makes it possible to correct iteratively the distance calibration until the average distance error converges to zero.

In the second part of this chapter we study four different ways to extract kinematics descriptors such as the rotational signature of the stellar halo from incomplete/noisy phase-space information. Two methods relys on l.o.s. velocity only, while the other two use proper-motions in an average way, assuming their error are purely random. We explore the dangers implied with the use of noisy proper-motions and present consistency tests that can assess whether the systematic floor actually associated with the tangential motion was sufficiently low that the assumption of purely random error is fair. We conclude by outlining the three different scenarios that are a priori consistent with a discontinuous kinematic profile: this is particularly of interest in the study of the halo given the partial relaxed status it is in.

### 1.3.2 Part II: Results for the Milky Way stellar halo

The results in the chapters of this section arise from a collaboration with Ralph Schönrich and were first presented in Fermani & Schönrich (2013a,b) respectively.

**Blue Horizontal Branch stars as tracers** We select two samples of Blue Horizontal Branch stars from the Sloan Digital Sky Survey to study the stellar halo of the Milky Way. Using the method outlined in the previous chapter, we construct and validate a new distance calibration for these stars that returns absolute magnitude as function of both colour and metallicity. Such calibration proves to hit the accuracy threshold set by the sample size and is superior to all available calibrations of the Blue Horizontal Branch in Sloan colours. In particular, because it covers the main dependences on colour and metallicity, it avoids the systematics associated with the other calibrations, which can distort the picture we infer of the stellar halo as exemplified in the previous chapter. The close connection of our calibration with the BASTI stellar evolutionary models allow us a basic assessment of the capability of these models to describe the Horizontal Branch, both absolutely and in comparison with other theoretical models, such as the Dartmouth stellar models.

**Rotational Signature** We present four different measures of the Milky Way stellar halo mean-streaming motion using the model and estimators presented in the third chapter and tracing the halo with the samples selected in the fourth chapter. We challenge our results to all known sources of uncertainty both in the sample and in fundamental parameters such as the Sun's position and velocity. We assess the

consistency of our measurements according the a posteriori tests introduced in the third chapter. We then focus on the work of Deason, Belokurov & Evans (2011a), who with a similar sample finds contradicting results, but also critique the analysis of Carollo et al. (2007, 2010), Beers et al. (2012), Kafle et al. (2012) and Hattori et al. (2013) who most recently contributed to the study of the stellar halo. We conclude that the Milky Way stellar halo is consistent with a non-rotating single component picture.

### 1.3.3 Part III: Dynamics

**f(J)-models for spheroids** We construct a new class of dynamical models defined in terms of actions for spheroidal components. Within an axisymmetric potential, we present a scheme to approximate the Hamiltonian as a function of the actions, which then enables us to write any isotropic spherical distribution function as a function of the three integrals of motion. The dependence on the actions both guarantees the flexibility of the model to describe realistic features such as flattening, anisotropy and rotation, but also offers a clean and intuitive understanding of the physics implied by the model. We further investigate the correlations between the two components of the model (i.e. potential and distribution function) and how they couple in shaping a stellar system. While these models might have been conceived for the Milky Way stellar halo, they are in fact suited to describe the dynamics of any galactic halo of spheroidal shape.

### 1.3.4 Conclusions and Future Work

In the last chapter we summarise the contribution of this work to the understanding and modelling of the smooth component of the Milky Way stellar halo and bring our results into context by discussing the next steps to take to unravel the Galaxy formation history and evolution.

# Part I

# Methodology & Techniques

# Chapter 2

# Synthetic catalogues

## 2.1 Introduction

Since the $17^{\text{th}}$ century the scientific method has consisted of an investigation protocol: observation followed by the formulation of a model followed by testing it with further experiments. In Astronomy, the experiments are the surveys (telescopes and satellites) and the first run of observations inspire the model. Successive surveys with different selection criteria provide then a test for the model: if it is able to fit also the new data, then is validated; otherwise the modelling exercise starts again, but on the larger data set constituted by both sets of observations.

Selection criteria are the knobs of our experimental apparatus: by adjusting them, we study how the phenomenon under investigation depends on the different variables. At each twist we make a prediction of what we will observe and challenge our understanding. Having control over the selection criteria and how they act on the data is mandatory: we cannot select directly on the quality we are trying to probe, otherwise we are automatically biasing our conclusions. In practice, if we are fitting a model to a set of observations, which inevitably will not cover the entire sky nor every object in the field of view, we need to know how incomplete or distorted is the picture we infer from them. Unfortunately selection criteria of surveys are anything, but transparent: they are affected by both technological limits of the apparatus (systematics) and environment variability (random). To test how these selections affect the collection of data, we can construct a catalogue of mock observations, a *synthetic catalogue*, and filter it in different ways until we obtain the picture acquired with the survey. Once the selection function is know, we just need to invert it in our analysis to obtain an unbiased picture. Vice versa, if the selection function is known, we can take a synthetic catalogue and apply it to the data. We then fit what survives with our model: comparing the inferred distributions with the original ones that generated the mock

catalogue, gives a sense of how much data we can loose before any inference from the data becomes meaningless.

The advantage is that for these mock observations, we know the true generating distribution and every data entry to arbitrary precision. This is the framework where we can test the consistency of our procedures: the comparison between the original distribution, which one supposes unknown, and the fitted model provides both a quantitative and qualitative prediction of the internal accuracy of our analysis and reveals whether the latter was designed in such a way as to introduce any significance bias in the measurements. For example, assume the data distributed, in whatever projection, according to a double-peaked probability density: the mean of the values then, returns a measure in between the two most probable ones, which in fact is in turn very little representative of the whole distribution as it is associated with a lower probability. Here our analysis consisted merely in extracting the mean of a dataset, but equally trivial is the inadequateness of the descriptor we chose to represent the observations.

To a greater extend we can use synthetic catalogues for budget control: along the same lines of the selection function study outlined above, we can answer the question, "what is the minimal amount of information required to answer a scientific question?" I.e., "how much data do I effectively need to infer the feature I am trying to probe?" Answering this question means a tremendous economical improvement efficiency-wise.

Even though the exercise is simple in principle, mapping a model in observable space, the practicalities involved in its implementation are an industry in continuous development. In this chapter we detail the general algorithm to map a model into observable space and critique the sampling phase of this process according to the available techniques and methods. In particular we enlarge on the difficulties specific to astronomy-related catalogues such as size and high multi-dimensionality. We conclude with a proposal to use synthetic catalogues to determine the optimal size of a survey.

## 2.2 From model to catalogue

A catalogue is a list of objects with measurables and errors, labelled by some physical properties such as apparent magnitude, colour, surface gravity, temperature and metallicity. Most of these properties depend on more fundamental variables such as mass and age. The information we want to input in our model to produce a catalogue, are the ones that allow a clear understanding of the physics implied by the model.

For example, the same magnitude or temperature are associated with a variety of masses, ages and evolutionary phases of the star.

On the other hand, stellar evolutionary models (e.g. §6.1.3 Binney & Merrifield 1998) show that there is no degeneracy in the physical profile of a star once the initial mass, the metallicity and an age have been chosen.[1]

Generating a synthetic catalogue therefore requires a phase-space distribution or distribution function (hereafter, DF) and the triplet of the population specific inputs: initial mass function (MF), metallicity distribution (MDF) and age distribution (ADF).[2]

### 2.2.1 A model seed: the Distribution Function

A distribution function is the probability to find at a certain time, a randomly chosen star in a given phase-space volume. Formally it is a solution of the collisionless Boltzmann equation, which describes the conservation of probability in phase-space:

$$\frac{\partial f}{\partial t} + \dot{\mathbf{q}}\frac{\partial f}{\partial \mathbf{q}} + \dot{\mathbf{p}}\frac{\partial f}{\partial \mathbf{p}} = 0 \tag{2.1}$$

where $t$ is time and $(\mathbf{q}, \mathbf{p})$ the generic canonical coordinates in phase-space. The potential of the system appears within the $\dot{\mathbf{p}}$ term of (2.1): a priori, any potential can be consistent with any DF, solution of (2.1). The only intellectually defendable assumption to break the degeneracy is to assume steady-state. Even though locally a galaxy is not in this regime (see spirals and bars in disks and streams in halos), overall it proves to be a good approximation and for any non-equilibrium study, a steady-state solution is the unavoidable starting point.

The relationship between a distribution function and observables is straightforward: the zero moment of the DF is the probability to find a star in a given volume $\mathrm{d}^3\mathbf{q}$ and multiplying this by the total number of stars $N$ gives the real-space density $\rho(\mathbf{q})$; the first moments describe the average motion $\bar{\mathbf{p}}$ and the second moments constitute the velocity dispersion tensor $\sigma_{i,j}$:

$$\nu(\mathbf{q}) = \int \mathrm{d}^3\mathbf{p} f(\mathbf{q}, \mathbf{p}) \tag{2.2}$$

$$\rho(\mathbf{q}) = N\nu(\mathbf{q}) \tag{2.3}$$

[1] In practice there is some degeneracy arising from the discreteness in the realization of the evolutionary tracks. Even though in same cases this discreteness is non-negligible (e.g. Schönrich, Asplund & Casagrande 2011), it still proves to be a cleaner descriptor than magnitude, temperature or else.

[2] In fact, observational evidence suggests that the initial mass function is universal for masses less than the solar mass (e.g. §5.1.9 Binney & Merrifield 1998).

$$\bar{p_i}(\mathbf{q}) = \frac{1}{\nu(\mathbf{q})} \int \mathrm{d}^3\mathbf{p}\, p_i f(\mathbf{q},\mathbf{p}) \tag{2.4}$$

$$\sigma^2_{i,j} = \frac{1}{\nu(\mathbf{q})} \int \mathrm{d}^3\mathbf{p}\, (p_i - \bar{p_i})(p_j - \bar{p_j}) f(\mathbf{q},\mathbf{p}). \tag{2.5}$$

In particular, the distribution function and the potential are linked by virtue of the Jeans' Theorem: the DF is a function of the isolating integrals of motion only (Jeans 1915), which ultimately are determined by the underlying potential. If we assume self-consistency (i.e. the gravitational field is entirely generated by the phase-space described by $f$), the distribution function and the potential are explicitly linked via the real-space mass density. From the DF we can compute $\rho(\mathbf{q})$ using (2.2) and determine the potential $\Phi(\mathbf{q})$ by integrating Poisson's equation:

$$\Delta\Phi(\mathbf{q}) = 4\pi G\rho(\mathbf{q}). \tag{2.6}$$

In reality, we are ignorant on the distribution of dark matter and can only trace the visible matter. A priori, we therefore renounce the requirement of self-consistency and describe a stellar system via the independent pair: underlying potential and distribution function.

The inverse problem, namely recovering the DF from a given potential and density profile, does not allow a general strategy for its resolution. In the case of spherical symmetry, Eddington showed that a DF of the form $f(E)$ can be determined uniquely from the pair $(\Phi(r), \rho(r))$ via the integral equation (Eddington 1916):

$$f(\mathcal{E}) = \frac{\sqrt{2}}{4\pi^2} \frac{\mathrm{d}}{\mathrm{d}\mathcal{E}} \int_0^{\mathcal{E}} \frac{\mathrm{d}\Phi}{(\Phi - \mathcal{E})^{\frac{1}{2}+\beta}} \frac{\mathrm{d}\rho}{\mathrm{d}\Phi} \tag{2.7}$$

where $\mathcal{E}(r,v) = -\Phi(r) - \frac{1}{2}v^2$ is the relative energy of the system and $\beta = 1 - (\sigma^2_\theta + \sigma^2_\phi)/(2\sigma^2_r)$ is the anisotropy parameter. Equation (2.7) contains a fictitious singularity, which can be extracted from the kernel leading to a generalized Abel integral equation:

$$f(\mathcal{E}) = \frac{\sin(\pi(\frac{1}{2}-\beta))}{\pi(\frac{1}{2}-\beta)} \int_0^{\mathcal{E}} \frac{\frac{\mathrm{d}}{\mathrm{d}\Phi} g(\Phi)}{(\mathcal{E}-t)^{\frac{1}{2}-\beta}} \mathrm{d}\Phi \tag{2.8}$$

where $g(\Phi)$ contains the first derivative of the density and therefore depends on the profile chosen.

For axisymmetric profiles, Lynden-Bell (1962) showed that the part of a distribution function $f(\mathcal{E}, L_z)$ that is even in the angular momentum $L_z$, $f_+(\mathcal{E}, L_z)$, can be deduced from the density $\rho(R,z)$ and its confining potential $\Phi(R,z)$. To determine the odd part we require an extra information though, namely the streaming motion

profile. The practical implementation of this idea is due to Hunter & Qian (1993), who alleviate some numerical difficulties in Lynden-Bell's method (having the density analytically continuable in the complex plane and making explicit the vertical coordinate dependence on the potential) by obtaining the DF from a contour integral.

The provision of a distribution function can be either analytical or as a realization of an N-body model. A N-body code is a software designed to follow the trajectory of a set of particles that evolve according to their mutual gravitational interaction. Every particle is representative of a set of stars or of dark matter and is integrated until the global picture resembles the observations one wants to model. The computational cost makes them not very appealing though: the burn-in time to produce a realization can be of the order of months, while is effectively zero for analytic DFs. Further, the discreteness of the realization is yet another (see following Section) source of noise and the dependence on initial conditions is less than ideal as it is the opaque understanding of which physics other than gravity (e.g. star formation rate, supernova explosions, etc.) needs to be considered to produce a convincing output. On the other side with analytic distribution functions, we can test directly the consequences of the assumptions on which the model is based simply by fixing all except one parameter and studying its effect on the associated phase-space distribution. Therefore, throughout this work, we will focus on analytic distribution functions.

Even though we postpone the main discussion on analytical DFs to chapter 6, it is worth mentioning that analytical distribution functions have already provided satisfactory account for solar neighbourhood observations (Binney 2012b) and of simulated dark matter halos (Pontzen & Governato 2013). Hence, our choice to focus on well defined DFs to create synthetic catalogues is fully justified and actually encouraged.

### 2.2.2 The Catalogue

The actual catalogue of mock observations is a sampling of the distribution function, with the dimensionality of each point sampled enlarged by the addition of a physical profile. The non phase-space coordinates of each point, namely the mass $M$, the metallicity [Fe/H] and the age $\tau$ are themselves sampled from the triplet of inputs that define a single stellar population, i.e. the initial mass function, the metallicity distribution and the age distribution (ADF)[3].

The final catalogue is then the projection into observable space of this sampling; depending on which use we want to make of the catalogue, we might filter it with a

[3]Given the time-scales involved in stellar evolution, the age distribution of a single population is approximated with a delta function.

given selection function. Position and velocity are converted in heliocentric coordinates by a simple frame transformation,

$$(d, l, b, v_{\rm los}, \mu_l, \mu_b) = \mathcal{T}_{\mathbf{x}_\odot, \mathbf{v}_\odot}(\mathbf{x}, \mathbf{v}) \tag{2.9}$$

where the subscript on the frame-transformation symbol implies its dependence on the velocity of the Sun and its position w.r.t. the model frame. The mapping of the physical profile into observables such as magnitude, colour, surface gravity and temperature (respectively $M_V, J-H, \log g, T_{\rm eff}$)[4] is by interpolation between stellar evolutionary models. In particular, the age, identifies the appropriate isochrone[5] and within it one finds the points in mass, metallicity-space closest to the sampled $(M, [\mathrm{Fe/H}])$; the observables associated with those points are then interpolated and the output taken to be the observables associated with $(M, [\mathrm{Fe/H}])$.

The creation of a synthetic catalogue is summarized in the following scheme:

$$\left.\begin{array}{c}\text{sampled } f(\mathbf{x},\mathbf{v}),\text{ gives } (\mathbf{x},\mathbf{v})\\ +\\ \text{sampled (IMF,MDF,ADF),}\\ \text{gives } (M,[\mathrm{Fe/H}],\tau)\end{array}\right\} \xrightarrow[\text{selection function}]{\mathcal{T}_{\mathbf{x}_\odot,\mathbf{v}_\odot},\text{ isochrone}} \begin{array}{c}\text{catalogue in obs. space}\\ (d,l,b,v_{\rm los},\mu_l,\mu_b,\\ m_V, J-H, \log g, T_{\rm eff}).\end{array}$$

where $m_V$ is the apparent magnitude obtained from $M_V$ and $d$.

The set of inputs to generate a synthetic catalogue, namely $(\Phi, f(\mathbf{x},\mathbf{v}), \mathrm{IMF}, \mathrm{MDF}, \mathrm{ADF})$, is fully experiment-dependent. Even though the initial mass function hints at some sort of universality for masses lower than the solar mass, all the other profiles are specific to the component to probe and to which tests need performing. For example if we are interested in validating the ability of a velocity estimator to measure the kinematics of a sample, we can disregard the information associated with the physical profile of the stars and generate only phase-space coordinates in the assumption we are looking at a clean tracer sample. On the other hand, if we work with a double component, introducing two relevant MDFs might be an essential discriminator to distinguish members of a different population a posteriori.

Sampling is instead a general problem and can be studied in abstract of observations. It also represents the one technical phase that needs to be optimal: if we over-sample, we waste computational resources, but if we under-sample, we obtain a biased realization of the model that is not fit for purpose.

[4]The $V$ subscript to magnitudes and the $J-H$ colour refer to the Johnson system (§2.3 Binney & Merrifield 1998), which is chosen for illustrative purposes only.

[5]An isochrone is a library of the physical parameters of a same-age population obtained by integrating a stellar evolutionary model.

## 2.3 Sampling a model

Sampling a model translates into sampling the probability densities that define the distribution of its parameters (the potential is assumed fixed). Hence the interest in the general problem of sampling a probability distribution function (PDF): creating a sample of points whose distribution reflects the original PDF. We review here three of the mainstream techniques in the field, namely: inverse sampling, the acceptance/rejection (A/R) von Neumann method and Markov Chain Monte Carlo (MCMC) sampling.

**Inverse sampling** Given a PDF $f(x)$, we compute the cumulative distribution

$$F(x) := \int_{x_{\min}}^{x} \mathrm{d}x' f(x'). \tag{2.10}$$

We then select a random point $\xi \in [0,1]$[6] and find the root of

$$F(x) = \xi. \tag{2.11}$$

This algorithm is clean and easy, but is basically restricted to one dimension. If $x$ is a vector in a space of dimension $d > 1$, (2.11) is the equation of an hyper-space and the solutions are $\infty^{d-1}$. We could extend the algorithm simply by applying it in each dimension separately, but this requires $f$ to be separable in its variables, which is a rather strong requirement.

**Acceptance/Rejection method** The essence of the A/R method is to take a point $x$ at random in the domain to sample and a value $\zeta \in [0, f_{\max}]$, where $f_{\max} \geq f(x)\,\forall x$ in the domain, also at random. The point is accepted if $\zeta \leq f(x)$. The implementation is computationally cheap (one random drawing and one function evaluation per step), but suffers of a high waste of energy due to the elevated number of rejections. Intuitively:

$$\text{rejections} \propto \int_{\mathcal{D}} \mathrm{d}x f_{\max} - \int_{\mathcal{D}} \mathrm{d}x f(x) = f_{\max}\,\mu(\mathcal{D}) - 1 \tag{2.12}$$

where $\mathcal{D}$ is the domain of $f$ and $\mu$ is a measure. There is a number of refinements that can improve the efficiency of the sampling. For example, given an *envelope* function $g(x)$ such that $g(x) \sim f(x)\,\forall x \in \mathcal{D}$, and $g(x)$ is analytically integrable, its cumulative

[6]Here we are assuming that $f(x)$ is a well-defined probability, i.e. it is everywhere positive and it sums up to one over its definition domain. If for whatever reason $f(x)$ does not obey the second requirement, we need to normalise the cumulative distribution: $\tilde{F}(x) := F(x)/F(x_{\max})$.

distribution being $G(x)$, we obtain an informed first guess for $x$ by inverse-sampling $g$. We take a random value $\xi \in [0,1]$ and determine $x$ as root of $G(x) = \xi$. Then we draw a second random number $\zeta \in [0, f_{\max}]$ and accept $x$ if

$$\zeta \leq k \, \frac{f(x)}{g(x)} \tag{2.13}$$

where $k$ is a parameter that ensures that the ratio on the RHS of (2.13) is less than one. The points we select now, have a priori a distribution that resembles the one of the true PDF because by construction $g$ is similar to $f$. Given that $f(x)/g(x) \sim 1$, the rejection rate is going to drop and we expect:

$$\text{rejections} \propto 1 - \int_{\mathcal{D}} \mathrm{d}x \left( k \frac{f(x)}{g(x)} \right) g(x) = 1 - k. \tag{2.14}$$

The closer $g$ is to $f$ (and the closer $k$ is to 1), the lower the number of rejections. The success of the scheme then depends on how easy is to find a suitable envelope function.

A convenient aspect of the A/R method is that it scales naturally to dimensions higher than one: the only modification is that we do not sample a random point $x$, but rather a random vector $\mathbf{x}$. For the version that uses the envelope function to control the number of rejections, one has just to make sure that $g$ is separable in its variables so that the inverse sampling can be applied on each component separately and overall the inversion is non-degenerate.

**MCMC sampling** Markov chain Monte Carlo sampling is a non-hysteric (memoryless) random walk: from a random starting point $x_0$ a new position $x_1$ is drawn also at random and is added to the chain with probability $\min(1, f(x_1)/f(x_0))$. This means that if the new position is equally or more probable than the initial one, we are certain to move to it; otherwise we make the step with probability $f(x_1)/f(x_0) < 1$. In practice, this translates into drawing a random value $\xi \in [0,1]$ and add $x_1$ to the chain if $\xi < f(x_1)/f(x_0)$. If the new position is rejected, the old one must be re-added to the chain:

$$\begin{aligned} &\text{if } [0,1] \ni \xi < \min(1, f(x_1)/f(x_0)) \quad \Rightarrow \quad x_0 \mapsto x_1 \\ &\text{else} \qquad\qquad\qquad\qquad\qquad\quad \Rightarrow \quad x_0 \mapsto x_0 \end{aligned}$$

The process is non-hysteric (or Markovian) because it does not depend on the previous steps, but only on the current state. As for the A/R method, the scaling to dimensions

greater than one is straightforward: instead of taking a point on a line, we randomly select in space, but the criterion to move to the new step stays exactly the same.

Regardless of the dimension, the computational cost is therefore very low: one random drawing and one function evaluation per step. In order not to be biased by where we start the chain though, we need to allow a certain burn-in time, which unfortunately is problem-specific. Also, the chain must be long enough for the sample to be a fair realization of the PDF and the convergence time unfortunately is also problem-specific. Finally as any random process, we are concerned in keeping the rejection rate low to control the discretization noise: at each step one point must be added to the chain, even if this is equal to the previous one, therefore consecutive rejections cause the oversampling of one particular point, leaving less time (steps) to explore its neighbourhood. This exacerbates the discreetness and hence the noise of the final distribution.

## 2.4 Sampling for astronomical catalogues

The high dimensionality of astronomical observations is one of the big challenges in the industry of building synthetic catalogues for this field. We have 6 phase-space coordinates and at least 3 more parameters to define the physical profile of the "observed" star (e.g. $M, [\mathrm{Fe/H}], \tau$) so that the domain $\mathcal{D}$ to sample is effectively isomorphic to $\mathbb{R}^n$, with $n \geq 6$. This automatically rules out the possibility to inverse-sampling and challenges the effectiveness of the two other methods (A/R and MCMC) to generate a suitable realization of the observations.

The random nature of both A/R and MCMC introduces the need to reject certain points so that the final distribution is not uniform, but resembles the input PDF. The higher the number of rejections, the higher the steps wasted (for A/R) and the higher the discreteness noise (MCMC). Further when MCMC sampling we are likely to have points in $\mathcal{D}$ that have been sampled more than once: while globally this might not be too much of a problem as the input PDF is on average recovered anyway, it is physically impossible to have more than one star at exactly the same position in real space, let alone them having the same speed and physical profile. Therefore, on top of having to control the discreteness noise, one has to design a way to somehow redistribute this clump of observations. Finally, for the sampling to be fair we need to wait until convergence to the input PDF: this implies that if one is interested in creating a realization of one portion of $\mathcal{D}$, has in fact to generate a catalogue over the entire domain and then throw away the points that are not included in the region of

interest. An easy fix to the this problem is to set the distribution function to zero in the regions not of interest: sharp changes in the PDF and edge effects require none the less careful handling (e.g. Cole et al. 1998; Girolami & Calderhead 2011).

In this section we motivate the thesis that even though refinements of general sampling algorithms do exist, the trade-off between efficient sampling and low computational cost/fast convergence requires tuning of the sampling strategy according to the performance goal. Through a series of considerations/examples we build up evidence that current schemes cannot be optimal under both aspects simultaneously.

### 2.4.1 An *ad hoc* refinement of A/R: *GALAXIA*

Sharma et al. (2011) put forward a series of refinements to the A/R method in the specific case of a synthetic catalogue for the Milky Way: they refer to the output code as $GALAXIA$. To reduce the number of rejected points, they apply the A/R method locally: within a given cell, the variation of the PDF is certainly less than what it would be over the whole domain. If for every cell, one recomputes $f_{\max}$ (see §2.3), the rejection rate drops because the points within the cell have more similar probability, without the need of determining a separable envelope function. The ideal cell grid would have the least density variation within each cell, but this criterion is likely to split the domain in a large number of almost empty cells; Sharma et al. (2011) choose instead to sub-divide the space into cells containing a fixed probability. A final burst of efficiency comes from folding a priori the selection function with the PDF: the probability of observing a low mass star decreases with distance from the observer for example or certain regions of the Sky might have just not been mapped by the survey one wants to test. It is therefore more efficient to avoid these blind patches in either real space or parameter space and sample the visible regions only.

The scheme still presents space for improvement though: the sub-division in cells containing a fixed probability requires the computation of density integrals over the whole domain, iteratively until convergence. Sharma et al. (2011) propose to estimate the probability enclosed in a cell by $p = f_{\max} \cdot V$, where $V$ is the cell's volume. This avoids the need for an integral and for slowly varying distributions is a fair zeroth-order approximation; it does not change the fact that several iterations are required to tune the grid of cells until each has roughly the same number-density though. Finally, if one wants to simulate the entire galaxy, how should the grid of cells be sampled? I.e. given cells containing equal probability, what is more efficient, MCMC or A/R?

### 2.4.2 Optimizing MCMC sampling

MCMC sampling is a very computationally cheap algorithm for it requires only drawing a random number and one function evaluation per sampled point. Two of its major pitfalls though are the fact that the time for convergence is problem-specific and the more non-uniform a distribution is, the higher the rejection rate and hence more inefficient the sampling. All the same, minor refinements exists that make the general algorithm both general and effective.

#### 2.4.2.1 Convergence

Verde (2007) proposed a simple, yet effective automated criterion to establish convergence of an MCMC chain: given $T$ chains of length $2N$, each with a different starting point, one estimates the chain-to-chain variance $B$ and the average variance of each chain $W$, which is an underestimate of the true variance:

$$B = \frac{1}{T-1}\sum_{j=1}^{T}(\bar{\theta}^j - \bar{\theta})^2 \tag{2.15}$$

$$W = \frac{1}{T(N-1)}\sum_{j=1}^{T}\sum_{i=N+1}^{2N}(\theta_i^j - \bar{\theta}^j)^2, \tag{2.16}$$

where $\theta_i^j$ is the point in parameter space in position $i$ of chain $j$, $\bar{\theta}^j$ is the mean of the chain $j$, $\bar{\theta}$ is the mean of all the chains and the burn-in time $N$ is half the length of the chains.

An overestimate of the variance is

$$V = \sigma^2 + \frac{B}{NT} = \frac{1}{N}\left[(N-1)W + B\left(1+\frac{1}{T}\right)\right], \tag{2.17}$$

where $\sigma^2$ is the true variance. Heavens (2009) suggests that for convergence one should require $\frac{V}{W} \sim 1.01 - 1.1$: the precision reported is empirical and does not follow from a rigorous proof. One should therefore opt for a much more conservative threshold.

To avoid numerical issues when dividing quantities that are potentially null, we modify Verde (2007) criterion to the following equivalent one:

$$\frac{V}{W} = (1+\varepsilon) \Rightarrow V - W = \varepsilon W, \tag{2.18}$$

where $\varepsilon > 0$ and

$$V - W = \frac{B}{N}\left(1+\frac{1}{T}\right) - \frac{W}{N}. \tag{2.19}$$

$T$ chains are therefore started from $T$ different locations (ideally such that the PDF differs the most across them) and are then run separately but simultaneously: every time a new step is added to the chains, $W$ and $B$ are re-computed, and if the convergence criterion above is not yet satisfied, the chains keep growing. This avoids the need for an a priori fixed chain length and burn-in time.

#### 2.4.2.2 Efficency

When we run an MCMC chain, to control the noise arising from the discreteness of the sampled distribution we want to keep the probability of transitioning from one position to a new one fairly uniform, that is to say, maintain the rejection rate low. A possible strategy is to "walk" with *small steps*: if the new position is drawn within a small distance from the current one, we are more likely to compare points with similar probability than if the new position is drawn at random over the whole domain (assuming the PDF is smooth), thus having fewer rejections. Even though intuitively promising, the autocorrelation between steps is inevitably higher and the burn-in time longer with small steps sampling. We shall briefly exemplify these pitfalls and quantify the effective gain efficiency-wise.

Small steps MCMC-sampling is effectively a diffusion process and for this we need to make sure we have reached equilibrium otherwise the domain is not fairly sampled. Take for example a domain $\mathcal{D}$ divided in one region of high probability, $\mathcal{R}_H$, and one of low probability, $\mathcal{R}_L$, but in such a proportion that

$$\int_{\mathcal{R}_L} \mathrm{d}x\, p(x) = \int_{\mathcal{R}_H} \mathrm{d}x\, p(x) \qquad \text{with } p(x) = \begin{cases} p_L & \text{for } x \in \mathcal{R}_L \\ p_H & \text{for } x \in \mathcal{R}_H \end{cases} \tag{2.20}$$

and $p_L, p_H \in \mathbb{R}$ ($p_L < p_H$). A priori we expect to have the same probability to move from one region to the other; in fact, if we start sampling in $\mathcal{R}_L$, the diffusion across the whole domain is held up by the "sticky" region $\mathcal{R}_H$ due to the sharp density gradient $p_L/p_H$, and before equilibrium is reached, the transition $\mathcal{R}_H \to \mathcal{R}_L$ results disfavoured. At the same time it is easy to show that sampling according to the original algorithm is fair at every step:

$$P(\mathcal{R}_L \to \mathcal{R}_H) = \frac{V_H}{V_\mathcal{D}} \min\left\{1, \frac{p_H}{p_L}\right\} = \frac{V_H}{V_\mathcal{D}} \tag{2.21}$$

$$P(\mathcal{R}_H \to \mathcal{R}_L) = \frac{V_L}{V_\mathcal{D}} \min\left\{1, \frac{p_L}{p_H}\right\} = \frac{V_L}{V_\mathcal{D}} \frac{p_L}{p_H} = \frac{V_H}{V_\mathcal{D}} \tag{2.22}$$

$$\tag{2.23}$$

where $V_X := \int_{\mathcal{R}_X} \mathrm{d}x$ is the volume of the region $X$ (i.e. of $\mathcal{D}, H, L$) and in the last equality, we have used the fact that $p_L V_L = p_H V_H$ from Eq. 2.20.

## 2.5 Quest for the ideal catalogue size

An appealing application of synthetic catalogues is to use them to define the minimal set of information required to answer a specific question. Beside the heuristic interest, such a study can lead to substantial economical savings in the market of large surveys by determining a priori how long it is worth running a survey for w.r.t. a particular question.

Let us assume that a given scientific question is associated with the systematic uncertainty $\sigma_0$, or equivalently that we want to decide which one between two models $f_1$ and $f_2$ is correct, with the discrepancy in their prediction being $\sigma_0$. We want to know how large a catalogue needs to be to answer the question, or equivalently decide between model $f_1$ and $f_2$. There will be a survey size $N_{\min}$ for which the sampling error is approximately $\sigma_0$. The principle is that if we collect a data set larger than $N_{\min}$ the sampling error shrinks below the systematic level, but with no gain in information because our knowledge is anyway limited by $\sigma_0$. The problem then becomes: given a specific question (and therefore a model, a dataset and its systematic uncertainty $\sigma_0$), how does the sampling error $\sigma(N)$ vary with the sample size $N$ and how can we determine $N_{\min}$ such that $\sigma(N_{\min}) \simeq \sigma_0$?

We have seen that when MCMC sampling, the speed of diffusion depends on the functional type of the PDF: at a fixed number of steps, the realization of an exponential model is closer to convergence than the one of a power-law model for example. Let us say that the functional form of our model is a fair approximation of the one describing the underlying physical law that governs the observed system; a fixed number of observations provides a better accuracy (and a lower sampling error) if this functional form is an exponential than if it is a power-law, which ultimately means that $\sigma(N)$ depends on the functional type of the model. The sampling error will therefore decay faster for an exponential model than for a power-law one. If the main dependence of $\sigma(N)$ is the functional type of the PDF that defines the model, we can use synthetic catalogues to construct a library of sampling error trends for the different functional-types of the PDF. Note that we are not assuming that if the PDF has a certain functional form, then $\sigma(N)$ is of the same functional-type. Rather that for a fixed functional-type of the PDF, the sampling error has a universal functional dependence on $N$, regardless of any specification of the physical problem. If this holds, then we estimate the sampling error for a few low values of $N$ on the real dataset and fit the appropriate pattern $\sigma(N)$ from the library to those. The optimal survey size is finally the root of $\sigma(N) = \sigma_0$.

In reality a survey is not an MCMC sampling of observable space, but it is fair to say it resembles a small steps MCMC sampling, at least in real space. Future work will experimentally confirm or disprove the speculations above and for realistic applications, it will be crucial to generalize to account for the survey selection function.

## 2.6 Conclusions

Synthetic catalogues provides a reproduceable laboratory frame in which to test the consistency of a model in retrieving a sensible answer and appear as an appealing way to investigate the optimal size of a survey. The construction of mock observations amounts to the sampling of an appropriate model, expressed in the form of a distribution function. While the definition of the model itself is highly problem-specific, the problem of sampling can be considered in abstract. Plenty of possibilities are available for this task (e.g. inverse sampling, acceptance/rejection method, Markov Chain Monte Carlo sampling), but not all are suited for astronomical-like catalogues: the high dimensionality-large sizes involved with astronomical observations challenge the effectiveness of a sampling scheme in light of the limited computational resources. Refinements in this sense of algorithms, such as the A/R method and the MCMC sampling, are possible: all exploit the idea of localized sampling. Considering a subdomain of the space to sample where there is a smaller density variation, causes a drop in the number of rejected points and hence in the discreteness noise. At the same time we either need to pre sub-divide the space in cells containing a known probability (local A/R method Sharma et al. 2011), which is expensive or we experience an increase in the correlation within the sampled points, which depending on the specific problem can be a negligible loss or highly undesirable.

Even though some basic principles are general across different problems, when designing a synthetic catalogue an a priori assessment of the performance required is needed, so that the machine that generates them can be optimized accordingly.

# Chapter 3

# The phase-space of a stellar system

## 3.1 Introduction

The kinematics of stars is a direct probe of the galaxy morphology and can provide indirect information on its underlying potential and formation history. The ground principle is that different formation scenarios leave different kinematic traces in the stars' phase-space distribution and by tracking them down we can identify the fundamental pieces that compose the system and constrain their formation history.

The kinematics of the Milky Way bulge for example, can reveal whether the observed barrishness was generated by mergers or it has grown from disk instabilities: stars in a merger-generated bulge are expected to rotate like in a spheroid with rotation at high latitudes being weaker than at low ones, while the streaming motion of instability bulges is thought to be coherent across different heights (Abadi et al. 2003; Athanassoula 2002; Howard et al. 2009)[1]. In the halo, substructures such as recent mergers or satellite galaxies of our own, might be "invisible" in real space having undergone significant tidal stripping at pericenter, but they are still resolved as separate objects in phase-space (e.g. Helmi & White 2000; McMillan & Binney 2008). As we have seen in the Introduction, the sole rotation information can pose constrains on the accretion history of a galaxy (see §1.1.2). Considering dynamical friction-type of arguments only, a halo that rotates with the disk indicates that it mainly did not form from mergers, while non-rotation or counter-rotation are consistent with a rich accretion history.

Intriguingly, Smith, Wyn Evans & An (2009) proposed that the alignment of the velocity dispersion tensor with a particular coordinate system can provide direct

[1]Most recently, Ness et al. (2013) concluded that the most likely origin of the Milky Way bulge is actually from disk instabilities, studying the kinematics of $\sim 17,400$ stars from the Abundances and Radial velocity Galactic Origins Survey carried out with the Anglo-Australian Telescope.

insight on the symmetry of the overall potential: they proved, in particular, that if the velocity dispersion tensor has one nondegenerate eigenvector pointing in the radial direction everywhere, then the underlying potential must be spherical. More recently McMillan & Binney (2013) showed how to determine the potential also quantitatively, starting from astrometric measurements with varying precision.

### 3.1.1 The phase-space for the halo

Phase-space measurements in the halo are structurally difficult: because we are dealing with distant objects, the tangential motion appears weakly on the sky and often of the order of the measurements uncertainty, so that the only kinematic information we rely on might be the l.o.s. velocity $v_{\rm los}$. The spheroidal geometry of the halo makes the streaming motion vary with latitude and in general observations at high distances have a lower signal-to-noise than closer ones. Finally spectra of halo stars have very weak metal lines due to the intrinsically metal-poor nature of the halo, which makes it harder to fit an appropriate template to the observed spectra and estimate the star's radial velocity.

In this chapter we exemplify some cases when errors in distance or inconsistent kinematical analysis lead to a distorted picture of the halo. We then focus on general techniques that assess the accuracy of a distance calibration and achieve robust measures of a sample's kinematics, in particular of its rotation. All the methods presented are conceived to make use of only l.o.s. measurements and tentative proper motions, which are the information effectively available for halo stars.

The considerations and techniques presented in this chapter are the result of a collaboration with Ralph A. Schönrich and have previously appeared in Fermani & Schönrich (2013a,b).

## 3.2 Effect of distance errors

The literature offers a geneours reservoir of examples, where distance errors significantly distorted the inferred picture of the Milky Way stellar halo: Ryan (1992) demonstrated that errors in proper motions due to overestimates in distance can bias the kinematics towards a counter-rotating trend. Schönrich, Asplund & Casagrande (2011) prove that the counter-rotating signature in the kinematics of the Galactic halo found by Carollo et al. (2007, 2010) is due to an inappropriate treatment of

observational uncertainties, and fundamentally is largely enhanced by distance misestimates of order 50%. Deason, Belokurov & Evans (2011a) find a rotation gradient across metallicity in a sample of field halo stars, but Fermani & Schönrich (2013a) showed that the error on the distance scale adopted by Deason, Belokurov & Evans (2011a) correlated with metallicity, concluding that the discrepancy in the kinematics of metal-poor and metal-rich stars presented had also been inflated by distance errors. Hattori et al. (2013) have recently argued for a tangentially biased metal-poor halo, based on a study where they deproject the l.o.s. velocity onto the three axis of the velocity ellipsoid to estimate the dispersion tensor. However, their conclusions are strongly undermined by failing to account for the spheroidal geometry of the halo (see §3.4.2.1).

In general, we do not know the real density profile of a sample of tracers with certainty due to selection effects (see chapter 4). In a maximum likelihood analysis, if the steepness of the profile is overestimated, a model sees distant stars as less probable than they really are, and vice-versa for close ones. Consequently, if we overestimate a star's distance, we give it more weight in our analysis than its closer companions and vice-versa if we underestimate its distance.

Further, distance errors propagate across all velocity components, introducing unphysical correlations. Schönrich, Binney & Asplund (2012) show that these correlations can actually be used to detect a calibration failure, when both proper-motion and l.o.s. velocity are available. In the case of the halo though, most often we have only tentative tangential velocities and we are likely to marginalize over them and use only the radial velocity as kinematic information. Overestimates in distance cause the portion of velocity space that the model expects to be populated by stars to shrink so that we marginalize over a smaller domain; oppositely for underestimates. In particular, as we overestimate the distance of a star, we assign it more and more weight, up to the distance at which the velocity space allowed by the model shrinks to zero: at that point the stars goes from being increasingly more probable to be "physically" impossible and is therefore assigned zero probability and neglected in the analysis.

Moreover, misestimating the distance from the Sun, moves the star along its l.o.s. varying the angle of deprojection with the galactocentric (GC) reference frame and hence its contribution to the principal directions of motion: Fig. 3.1 shows for example how distance errors influence the rotation estimated from $v_{\rm los}$ (see §3.4.2.1). As the star is pushed away from us, $v_{\rm los}$ becomes more perpendicular to the rotation direction: overestimates in distance requires therefore more rotation to produce the observed $v_{\rm los}$, while underestimates shrink the inferred rotation. The same geometrical effects

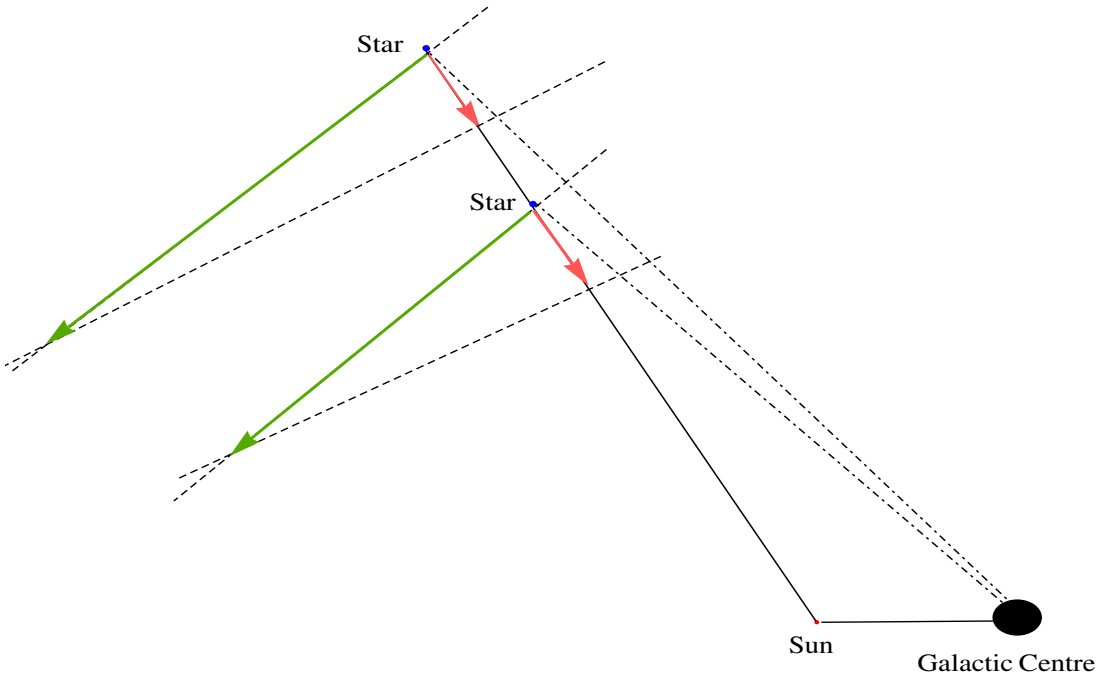


Figure 3.1: Velocity diagram (dashed lines) overplotted in real space (solid black lines) of a star for which only $v_{\rm los}$ is available (red line): even though not true for a single star, on average across the Sky we can assume the observed $v_{\rm los}$ to be a projection of $v_\phi$ only (green line; see §3.4.2.1). As we slide the star (due to distance misestimates) along the l.o.s., the $v_\phi$ required to produce the observed $v_{\rm los}$ changes.

apply for the anisotropy of the velocity ellipsoid: as distance gets overestimated, the angle between the line of sight and the line GC-star becomes smaller. Therefore to produce the same observed $v_{\rm los}$ we require a bigger tangential velocity, while the radial velocity diminishes. The sample appears then more tangentially biased and vice versa underestimates in distance reflect into a more radially biased velocity ellipsoid.

Finally, correlation of distance error with chemical parameters can distort the morphological view we infer of the halo: if distance errors correlate with metallicity, stars distances at one end of the metallicity range will be overestimated and underestimated at the other end. The structure and kinematics of metal-rich and metal-poor stars of a single component can then appear to differ.

## 3.3 Assessing a distance calibration

Schönrich, Binney & Asplund (2012), hereafter SBA12, investigated how distance errors introduce correlations between velocity components. They consider a distance

fractional error $f$ such that the measured distance is $s' = (1+f)s$, where $s$ is the true distance. They take $(U, V, W)$ to be the velocity components of the Cartesian reference frame where the Sun is at the origin. $U$ points from the Sun towards the Galactic centre, $V$ is in the direction of Galactic rotation and $W$ points towards the North Galactic Pole. The LSR velocity vector is the one determined by Schönrich, Binney & Dehnen (2010): $(U, V, W)_\odot$=(11.1,12.24,7.25) km s$^{-1}$. We shall adopt the same conventions throughout this thesis.

The observed $(U, V, W)$ velocity components can be then expressed as functions of both the real velocities $(U_0, V_0, W_0)$ and of $f$, which introduces correlations between the velocity components:

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = (1 + f\mathbf{MPM}^{\mathrm{T}}) \begin{pmatrix} U_0 \\ V_0 \\ W_0 \end{pmatrix} \tag{3.1}$$

where $\mathbf{M}$ is the coordinates transformation $\mathbf{M} : (s\mu_b, s\mu_l, v_{\mathrm{los}}) \mapsto (U_0, V_0, W_0)$, $(\vec{\mu}, v_{\mathrm{los}})$ being the proper motions and l.o.s. velocity respectively and $\mathbf{P}$ accounts for the fact that distance errors affect the tangential velocity only.[2]

The main method of SBA12 is not applicable to stars with noisy proper-motion (which is the case for most of the tracers of the deep halo available), since the (uncertain) proper-motion errors enter the correlation multiplied with the square of the distance. In practice, with current proper motions the method is applicable to samples of stars with $s < 5\,\mathrm{kpc}$.

However, we can directly use the linear estimator given in their equations 42-45. This estimator uses the fact that halo stars have a large mean heliocentric azimuthal $(V)$ velocity because the halo does not have any significant rotation so we see the full reflex motion of the Sun. Any mean distance error then translates this mean motion into a mean motion in the radial $U$ and vertical $W$ components of heliocentric velocity depending on galactic angles according to the matrix elements in $\mathbf{T} = \mathbf{MPM}^{\mathrm{T}}$. Eq. 3.1 expanded (eq. 7 in SBA12), reads

$$V = (1 + fT_{VV})V_0 + fT_{VU}U_0 + fT_{VW}W_0 \tag{3.3}$$

$$W = (1 + fT_{WW})W_0 + fT_{WU}U_0 + fT_{WV}V_0 \tag{3.4}$$

[2]From equations 4, 6 of SBA12:

$$\mathbf{M} = \begin{pmatrix} -\sin b\cos l & -\sin l & \cos b\cos l \\ -\sin b\sin l & \cos l & \cos b\sin l \\ \cos b & 0 & \sin b \end{pmatrix} \text{ and } \mathbf{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.2}$$

where $T_{ij}$ indicate the matrix elements of $\mathbf{T}$, with $i, j = U, V, W$. Under the assumption that $\langle U_0 \rangle = -U_\odot$ and $\langle W_0 \rangle = -W_\odot$, they take the sample mean of the second equation above (eq. 42 in SBA12):

$$\langle W \rangle + (1 + fT_{WW})W_\odot + fT_{WU}U_\odot = fT_{WV}\langle V_0 \rangle \tag{3.5}$$

and estimate the size of the bias on $\langle V \rangle$ from assuming the wrong distance scale with the first term on the RHS of eq. (3.3):

$$\langle V_\odot \rangle \simeq \frac{\langle V \rangle}{1 + f\langle T_{VV} \rangle}. \tag{3.6}$$

Finally, with the further assumption that $\langle W \rangle + (1 + fT_{WW})W_\odot + fT_{WU}U_\odot \simeq \langle W \rangle + W_\odot$, they combine the two equations above and provide a handy expression for the correlation between the mean rotational motion and the vertical heliocentric mean velocity (equation 45 of SBA12):

$$\langle W \rangle + W_\odot = c \cdot T_{WV} + \varepsilon \tag{3.7}$$

where $c = f\langle V \rangle / (1 + f\langle T_{VV} \rangle)$ is the slope of the linear best fit between $\langle W \rangle + W_\odot$ and $T_{WV}$, while $T_{VV} = 1 - \cos^2 b \sin^2 l$ and $T_{WV} = -\frac{1}{2} \sin 2b \sin l$ are the angular terms through which distance errors introduce correlations between the apparent components of velocity (see Table 1 of SBA12 or eq. 3.1 in this thesis). In particular a correlation between $W$ and $T_{WV}$ implies that the stellar halo is moving towards the disk on one side of the Galaxy (w.r.t. the Sun-GC line) and dispersing away from the disk on the other side (the "falling sky " effect, in the nomenclature of SBA12).

In practice, we perform a linear least squares fit of $\langle W \rangle + W_\odot$ versus $T_{WV}$ to estimate the correlation term $c$. Then, we iteratively multiply all distances by correction factors until this correlation on the sky disappears. The fractional average distance error is then just given by the correction factor that makes this correlation zero and the uncertainty on this estimate directly translates into the error bars of $c$. As the method is linear, random proper-motion errors are irrelevant, apart from increasing the error bars by expanding the measured dispersion in $W$. We are, however vulnerable to systematics in the proper motions and one should correct for them beforehand[3]. The technique is applicable to any distribution of stars with no prominent substructure within it: as long as we have a comparable amount of kinematically unbiased

[3]For example, in using data from the Sloan Digital Sky Survey (SDSS Eisenstein et al. 2011), one should use the proper-motion correction worked out by Schönrich (2012) on the Schneider et al. (2010) quasar sample in Galactic angles $(l, b)$: this corrects for the frame-dragging effect (see §5.2.1.1), which is the most significant systematics known afflicting the sample.

stars on the "left" and "right" side of the Galaxy ($l \in [0, 180]$ and $l \in [180, 360]$ respectively) we can assess whether the sky is falling or not. In particular we are not bound to have a symmetric distribution between north and south hemisphere as the effect of the galaxy aggregating towards the plane on one side and evaporating away on the other would still be noticeable even if the stars in our catalogue sample only one of the two hemisphere (the only plausible phenomenon that could account for this behaviour is indeed the presence of prominent substructures with independent kinematics).

### 3.3.1 Internal accuracy of the method

If the sky is "falling", in the sense of SBA12, then we know there is a distance bias. We want to assess how accurately can the method above quantify the mean fractional distance error. For this we take a sample of Blue Horizontal Branch stars by Xue et al. (2011) (see chapter 4 for details) with the distance calibration of Fermani & Schönrich (2013a): with this setting we observe no correlation between $W$ and $T_{WV}$ and can therefore consider the sample with its distance calibration, having zero average distance error.

We generate a control case by systematically biasing the stars distances and tangential motions so that the average distance offset amounts to $f_{\rm in}$:

$$\left\langle \frac{d_i - \tilde{d}_i}{d_i} \right\rangle = f_{\rm in} \quad , \tag{3.8}$$

where $\tilde{d}_i = d_i c_n \mathrm{ran}(\zeta) g(l_i, b_i, d_i)$ is the perturbed distance of the $i$-th star, $g(l_i, b_i, d_i)$ is the dot product between the l.o.s. and the rotation direction and accounts for the angular position of the star in the sky, $\mathrm{ran}(\zeta)$ is a random number uniformly distributed in $(0, 1)$, and $c_n$ makes sure that (3.8) is satisfied.

We uniformly select eight different values of $f_{\rm in}$ in the range $[-0.2, 0.2]$ (i.e. $\pm 20\%$). Using the $V$-$W$ estimator to estimate the mean fractional distance error on the sample, we recover the true bias within one standard deviation:

$$\langle f_{\rm in} - f_{V-W} \rangle = 0.0162 \pm 0.0187. \tag{3.9}$$

## 3.4 Kinematics

One way to extract morphological and kinematic information from a sample is to fit all the data to a sufficiently general theoretical model. Commonly studied parameters of

the Milky Way stellar halo are the rotation and the deviation of the velocity ellipsoid from a spherical shape: the ellipsoid has usually principal axes taken parallel to $\hat{\mathbf{r}}$, $\hat{\theta}$ and $\hat{\phi}$ and the radial velocity dispersion is varied against the two other directions. A generally accepted method is to assume a density profile for the stellar component and an overall potential for the Galaxy. Then one builds up a phase-space distribution function (DF) in the steady-state approximation, which will depend on parameters identifiable with anisotropy and rotation. One either can marginalise over observables that are considered too noisy or, more properly, convolve the distribution function with an error function replicating the observation process. To estimate a value for the physically meaningful parameters, one can look at the likelihood of the data given the model, varying the free parameters. The results must be robust to the uncertainty in the data and give consistent results for different subsets, and the model needs to be flexible enough to fit various behaviours and trends.

Here, we critique the procedure that, in a spherical potential, takes the popular double power-law model by Evans, Hafner & de Zeeuw (1997) and marginalises over proper-motion so that rotation and anisotropy are estimated from $v_{\rm los}$ information only. This procedure has been recently used by Deason, Belokurov & Evans (2011a) on a sample of Blue Horizontal Branch stars drawn from SDSS (DR7): we recall the model definition following their notation closely, to facilitate the comparison of our results later (see chapter 5). As an alternative, we assess the possibility of directly estimating rotation using either $v_{\rm los}$ only and full 3D motion. The latter estimators extract the average rotation and therefore are suitable even at large proper motion errors, as long as the systematic bias is small. We conclude this section by recalling general principles of consistency for the rotational signature of a smooth component: these will later be useful to test the results of our analysis.

### 3.4.1 Likelihood analysis

Deason, Belokurov & Evans (2011a) assume a radial potential of the type $\Phi \sim r^{-\gamma}$, $\gamma = 0.5$, normalized so that the escape velocity at the Sun is the one estimated by Smith et al. (2007)[4]. The density profile of the tracer stars is assumed to be also of power-law type: $\rho \sim r^{-\alpha}$ with $\alpha = 3.5$.

To estimate the rotational signature and anisotropy of the sample, they adopt a double power-law distribution function depending on two parameters: $\beta$ (anisotropy

[4]The Sun is taken to be at a distance of 8.5 kpc from the Galactic Centre. The escape velocity reported by Smith et al. (2007) is $v_{\rm esc} = 498 - 608$ km s$^{-1}$ with 90% confidence and an average of 544 km s$^{-1}$. We adopt the average value.

parameter as defined in Binney & Tremaine 2008, §4.3.2) and $\eta$ (rotation), such that $\eta = 0$ corresponds to complete pro-rotation and $\eta = 2$ to complete counter-rotation:

$$F_{\eta,\beta}(E, L, L_z) = c_\beta \left(1 + (1-\eta)\tanh\left(\frac{L_z}{\Delta}\right)\right) L^{-2\beta} E^{\zeta}, \tag{3.10}$$

where $E$ is the binding energy, $\Delta$ is a smoothing parameter, $c_\beta$ is a constant that enables us to make the real-space density independent of $\beta$ and

$$\zeta = \frac{\beta(\gamma-2)}{\gamma} + \frac{\alpha}{\gamma} - \frac{3}{2}. \tag{3.11}$$

Assuming dominant errors on the proper motions, the DF is marginalized over transverse velocities.[5] The free parameters $(\eta, \beta)$ are then estimated via MCMC sampling on the parameter space with the probability density being the likelihood of the data[6] given the model, so the probability density in parameter space is the exponential of:

$$L(\eta, \beta) = \sum_{i=0}^{N} \log F_{\eta,\beta}(l_i, b_i, d_i, v_{\mathrm{los},i}), \tag{3.12}$$

where

$$F_{\eta,\beta}(l, b, d, v_{\mathrm{los}}) = \int\int \mathrm{d}v_l \mathrm{d}v_b F_{\eta,\beta}(l, b, d, v_l, v_b, v_{\mathrm{los}}). \tag{3.13}$$

#### 3.4.1.1 Notes on the method

The choice of an index $\alpha \geq 3$ for the density profile makes it impossible to compute a normalization constant for the DF in (3.10) as it contains an infinite mass at the centre: $\int \mathrm{d}^3x\ \mathrm{d}^3v\ F_{\eta,\beta}(E, L, L_z) = \infty$. Therefore, in order to be able to compare different models we must fix the density in the sampled region of real space. Thus

$$\rho(r) = \int \mathrm{d}^3v\ F_{\eta,\beta}(E, L, L_z) \tag{3.14}$$

is independent of $\beta$.

[5]The errors associated with $v_{\mathrm{los}}$ measurements can be taken into account by convolving the DF with Gaussian errors of variance the size of the observational uncertainty. Applying the model (3.12) with or without the convoluted $v_{\mathrm{los}}$ errors, makes no appreciable difference in our results though.

[6]Heliocentric galactic polar coordinates are converted into galactocentric cylindrical coordinates reference frame, by assuming $(x, y, z)_\odot = (8.5, 0, 0)$ kpc, $v_{c\odot} = 220$ km s$^{-1}$. We are aware that this specific parameter combination is at odds with the measured proper motion of Sgr $A^*$, but keep those values consistent with previous studies. We further assume a solar Local Standard of Rest velocity of $(U, V, W)_\odot$=(11.1,12.24,7.25) km s$^{-1}$, as updated by Schönrich, Binney & Dehnen (2010).

A convenient consequence of fixing the density is that we can neglect selection effects which would otherwise apply (McMillan & Binney 2012). The uncertainty on the density profile of the tracers remains a matter of concern though: as we remarked in §3.2, if the steepness of the profile is overestimated, for example, the model considers distant stars less probable than they really are. This results in these objects having more weight than their closer companions. An unfortunate aggravation of this bias is that fainter/more remote stars are known to be noisier and are intrinsically a bad tracer of the rotational signature. The only safe strategy would be therefore to marginalize also over the profile steepness $\alpha$.[7]

When marginalizing over proper motions with $\beta > 0$, we compute an integral that involves a singularity in the DF (3.10) at $L = 0$ and the treatment of this singularity is crucial to retrieve $\beta$ correctly. Fortunately, the singularity is integrable for $\beta < \frac{1}{2}$.

The estimate of the anisotropy is still a matter of concern though: $\beta$ appears as $L^{-2\beta}$. For $\beta < 0$, this term gives rise to a bimodality in velocity space, which cannot adequately fit the single-peaked velocity distribution observed for local halo members. Depopulating the velocity distribution around $V \sim 0$ induces an azimuthal bias, which can then explain why the anisotropy is at odds with observations of a strongly radially dominated velocity ellipsoid for local stars (see e.g. Smith et al. 2007; Schönrich, Asplund & Casagrande 2011).

Equation (6) of Deason, Belokurov & Evans (2011a) states that the mean rotational velocity for model (3.10) depends on radius only. By computing

$$\langle v_\phi \rangle = \frac{\int \mathrm{d}^3 v \; v_\phi F_{\eta,\beta}(E, L, L_z)}{\int \mathrm{d}^3 v F_{\eta,\beta}(E, L, L_z)} \tag{3.15}$$

we see that this is true only in the limit $\Delta \to 0$. Indeed, one can show that if $\tanh(L_z/\Delta) \neq \mathrm{sgn}(L_z)$, then $\langle v_\phi \rangle$ depends on the full 3D position. We will therefore adopt this assumption even though it was not stated in Deason, Belokurov & Evans (2011a). We plot in Fig. 3.2 the mean streaming motion as a function of the rotational parameter for three values of $\beta$ at two different radii, $r = 1\,\mathrm{kpc}$ (dotted lines) and $r = 50\,\mathrm{kpc}$ (full lines).

#### 3.4.1.2 Test of implementation

We test the implementation of likelihood analysis just introduced against 12 mock samples: we generate a sample by MCMC sampling phase-space ($4 \times 10^4$ points)

[7] The first guess for the prior of $\alpha$ is a flat distribution; a sensible refinement would be to consider a Gaussian distribution centred at $\alpha = 3.5$, with a variance such that discrepant values in literature (also relative to other tracer samples, whether available) are consistent within one sigma.

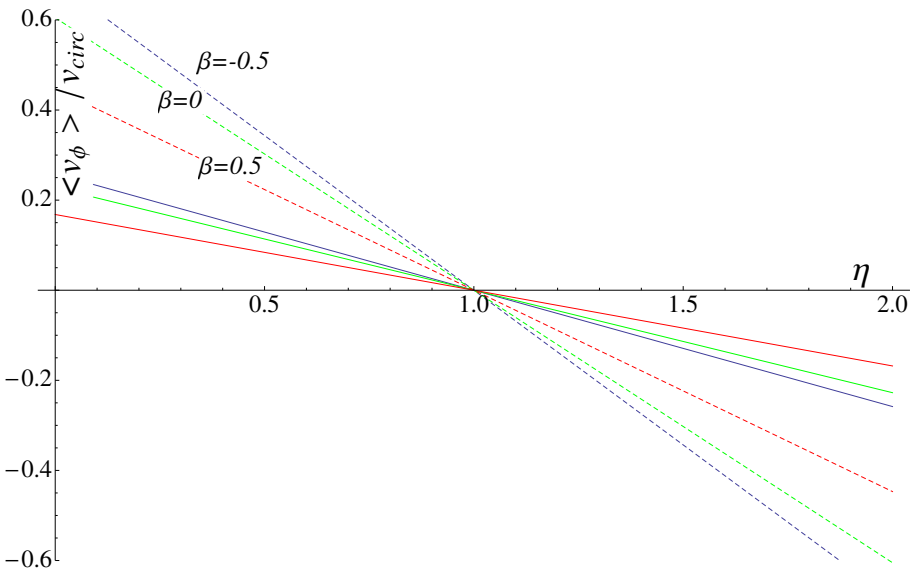


Figure 3.2: Mean streaming motion as a function of the rotational parameter $\eta$ for three values of $\beta$ and at two different radii: $r = 1\,\mathrm{kpc}$ (dotted lines) and $r = 50\,\mathrm{kpc}$ (full lines). The units for $\langle v_\phi \rangle$ are in $v_{\mathrm{circ}}$.

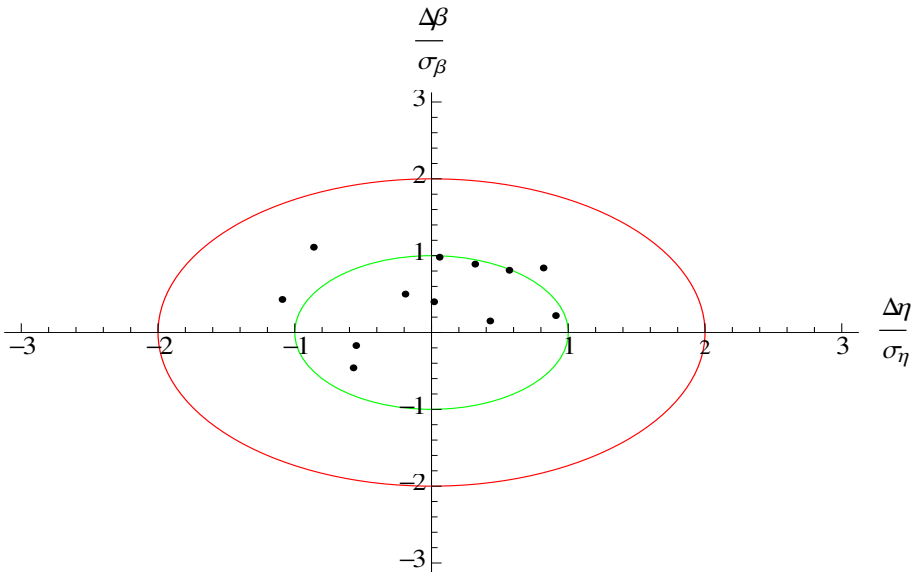


Figure 3.3: Performance of the MCMC fitting technique: the plot shows the discrepancy between the actual parameters and the fitted ones, weighted by the formal error. The fit is performed by MCMC sampling parameter space using the likelihood of the pseudo-data given the model. The green circumference represents the one $\sigma$ contour, the red one the two $\sigma$ contour.

according to the distribution function in equation (3.10) and random parameters $(\eta, \beta) \in [0, 2] \times [-0.5, 0.5]$. The reliability of the sample obtained is checked by comparing the density profile and rotational signature derived from star-counts of the synthetic catalogue produced with the theoretical profiles, revealing perfect agreement.

Subsequently, we determine the probability density of the model by MCMC sampling $(\eta, \beta)$-space using the likelihood of the pseudo-data given the model. The convergence of the MCMC chain is secured by simutaneously running two separate chains on each sample, starting from different starting points. Each pair of chains terminate when the variance between them is of order $10^{-3}$, which requires typically $10^2 - 10^3$ steps (see §2.4.2.1 for a detailed description of this MCMC-convergence criteria). The original parameters are recovered within one sigma as demonstrated in Fig. 3.3. From the distribution of the deviations we also infer that the error estimates are reasonable.

### 3.4.2 Model-independent estimators of rotation

We present here three model-independent estimators of rotation: one relies on $v_{\rm los}$ information only (a similar approach has been previously adopted by Frenk & White 1980) and the other two on the full 3D motion. SDSS proper-motions for example, at these distances are of the size of the noise associated with them. All the same, given that the estimators rely on averages on the sky, valuable information can still be extracted from them provided that their errors are random and not systematic. To set ourselves in such a condition we correct for the known systematic of astrometric frame-dragging (see §5.2.1.1 or discussion in Schönrich 2012); if there is any other significant systematic left, the estimators will be biased differently and will therefore contradict each other (see SBA12). Further, one of them outputs the U-component of Sun velocity and the latter information can be directly compared to the values in literature to assess the presence of proper-motion systematics significant enough to bias the estimator.

#### 3.4.2.1 $v_{\rm los}$ estimator

Inferring the rotational signature using only $v_{\rm los}$ presupposes knowledge of the global velocity distribution: i.e. how the velocity ellipsoid projects into $v_{\rm los}$. In fact, it is sufficient to assume that the azimuthal velocity $v_\phi$ is not correlated with the other components and that the means of the other two components are $\sim 0$. This minimal set of assumption allows us to measure rotation from the de-projected $v_{\rm los}$ on the azimuthal direction. In general we have:

$$v_{\rm los} = (\mathbf{v} - \mathbf{v}_\odot) \cdot \hat{\mathbf{s}} \Rightarrow \mathbf{v} \cdot \hat{\mathbf{s}} = v_{\rm los} + \mathbf{v}_\odot \cdot \hat{\mathbf{s}}, \tag{3.16}$$

where $\hat{\mathbf{s}}$ is the l.o.s. unit vector, $\hat{\mathbf{e}}_\phi$ is the unit vector in the azimuthal direction of Galactic cylindrical coordinates and $\mathbf{v}_\odot$ is the Sun's velocity w.r.t. the LSR. Then if $\langle v_R \rangle$ and $\langle v_z \rangle \sim 0$, we can estimate the average rotation by

$$\langle v_\phi \rangle \simeq \Big\langle \frac{v_{\rm los} + \mathbf{v}_\odot \cdot \hat{\mathbf{s}}}{\hat{\mathbf{s}} \cdot \hat{\mathbf{e}}_\phi} \Big\rangle. \tag{3.17}$$

Hereafter, we shall refer to the de-projected l.o.s. velocity as

$$\pi^{-1}(v_{\rm los}) := \frac{v_{\rm los} + \mathbf{v}_\odot \cdot \hat{\mathbf{s}}}{\hat{\mathbf{s}} \cdot \hat{\mathbf{e}}_\phi}.$$

If $v_\phi$ is independent of position we expect that $\nabla_{\mathbf{x}} \langle v_\phi \rangle \sim 0$, where $\mathbf{x} = \{x_j\}_{j=1,2,3}$ is the 3D position vector. The best estimate of the mean streaming motion is then obtained by averaging the deprojected measurements scaled by their associated variances following a maximum likelihood argument. In practice, we fit a straight line in the $x_j$, $\pi^{-1}(v_{\rm los})$ plane with the "weighted" least square routine of `gnuplot` having marginalized over the other two position variables. If $\pi^{-1}(v_{\rm los})$ depends on position, the estimator still gives a best estimate of the position-dependent rotation velocity. The variances must account for the projection of the velocity ellipsoid onto the l.o.s. other than the measurement error on the radial velocity. If the velocity ellipsoid has direction vectors $\hat{\mathbf{e}}_j$ and dispersions $\sigma_j$, the variance to consider on the deprojected $v_{\rm los}$ of the $\zeta$-th star (and therefore the inverse of the weight to assign it in the weighted least square fit) is:

$${\sigma_\zeta}^2 = \frac{\left(\sigma^2_{v_{{\rm los},\zeta}} + \sum_{j=1}^3 (\sigma_j \hat{\mathbf{e}}_j \cdot \hat{\mathbf{s}})^2\right)}{(\hat{\mathbf{s}} \cdot \hat{\mathbf{e}}_\phi)^2}, \tag{3.18}$$

where $\sigma_{v_{{\rm los},\zeta}}$ is the radial velocity measurement error for the $\zeta$-th star. The deprojection factor has a singularity, where $\hat{\mathbf{s}}$ is perpendicular to $\hat{\mathbf{e}}_\phi$. In Fig. 3.4 we show $\hat{\mathbf{s}} \cdot \hat{\mathbf{e}}_\phi$ in the $(x, y)$-galactocentric Cartesian plane at altitudes of $|z| = 5$ and $|z| = 20\,$kpc. A deprojection factor of 0.1 means that a rotation velocity of $20\,{\rm km}s^{-1}$ leaves a signal of just $2\,{\rm km}s^{-1}$ in the mean l.o.s. velocity. Hence the region where l.o.s. velocities can be reliably used is confined to small galactocentric radii and low altitudes with the best regions along the y-axis. The variation also implies that weighting the data points according to their uncertainty is mandatory to obtain any meaningful results.

Neglect of $\sigma_\zeta$ will lead to large errors (see §3.4.3.1) and fluctuations of the results.[8] Fortunately we do not require exact knowledge of the uncertainty. Since variations in the numerator of (3.18) are small and since weighted least squares only improve the statistics but do not introduce any bias, we can safely set $\sum_{j=1}^{3}(\sigma_j \hat{\mathbf{e}}_j \cdot \hat{s})^2$ constant, as long as we recover the approximate shape around the singularity.

#### 3.4.2.2 3D estimators

We have extracted rotational information out of $v_{\rm los}$ alone both via a model and via a direct estimator. $v_{\rm los}$ does not hold the full kinematics though and we could be missing stream-like structures in our sample with peculiar motions. A stream-like structure is of concern in the kinematic analysis of the smooth component when it emerges as an overdensity in velocity space: such an anomaly w.r.t. the smooth component might be more or less prominent depending on which projection of velocity space we consider. Say we have a sub-structure in the sample which appears faint in $v_{\rm los}$, but is prominent in the tangential motion: the streaming motion estimated via $v_{\rm los}$ will be mildly shifted by the stream and possibly even within the error bars, which are anyway generous given the limited sample size. Then we are incapable of deciding whether we are observing a statistical fluctuation or the imprint of a structure. Of course enlarging the sample size "enough", can discriminate between the two possibilities; however, if the dataset is already as large as the technological limitations allow, we are left with the one choice of checking whether the shift is matched or not in the other two components of velocity. To disentangle the possible bias by streams, we therefore *need* to make use of the full 3D velocity.

This implies that the major part of velocity information actually comes from proper motions. While velocities of individual stars are not reliable - the error is approximately as large as the signal - large number statistics solve that problem for aggregate quantities like the average rotation.[9] The error in the mean quantities is anyway best estimated from the as-observed scatter and does not require a-priori knowledge of the proper motion (random) errors. More of a concern are systematic errors on the proper motions like astrometric "frame-dragging" (§5.2.1.1). One must quantify how much an eventual correction to this bias affect the estimates as well as test whether data releases from different pipelines lead to inconsistent conclusions.

[8] We suspect this is the case in the study of Hattori et al. (2013), who do not mention any care towards this issue.

[9] Even if the noise is actually bigger than the signal, which is the case for SDSS measurements of stars beyond 20 kpc from the Sun, this introduces no preferred direction of motion and hence no bias.

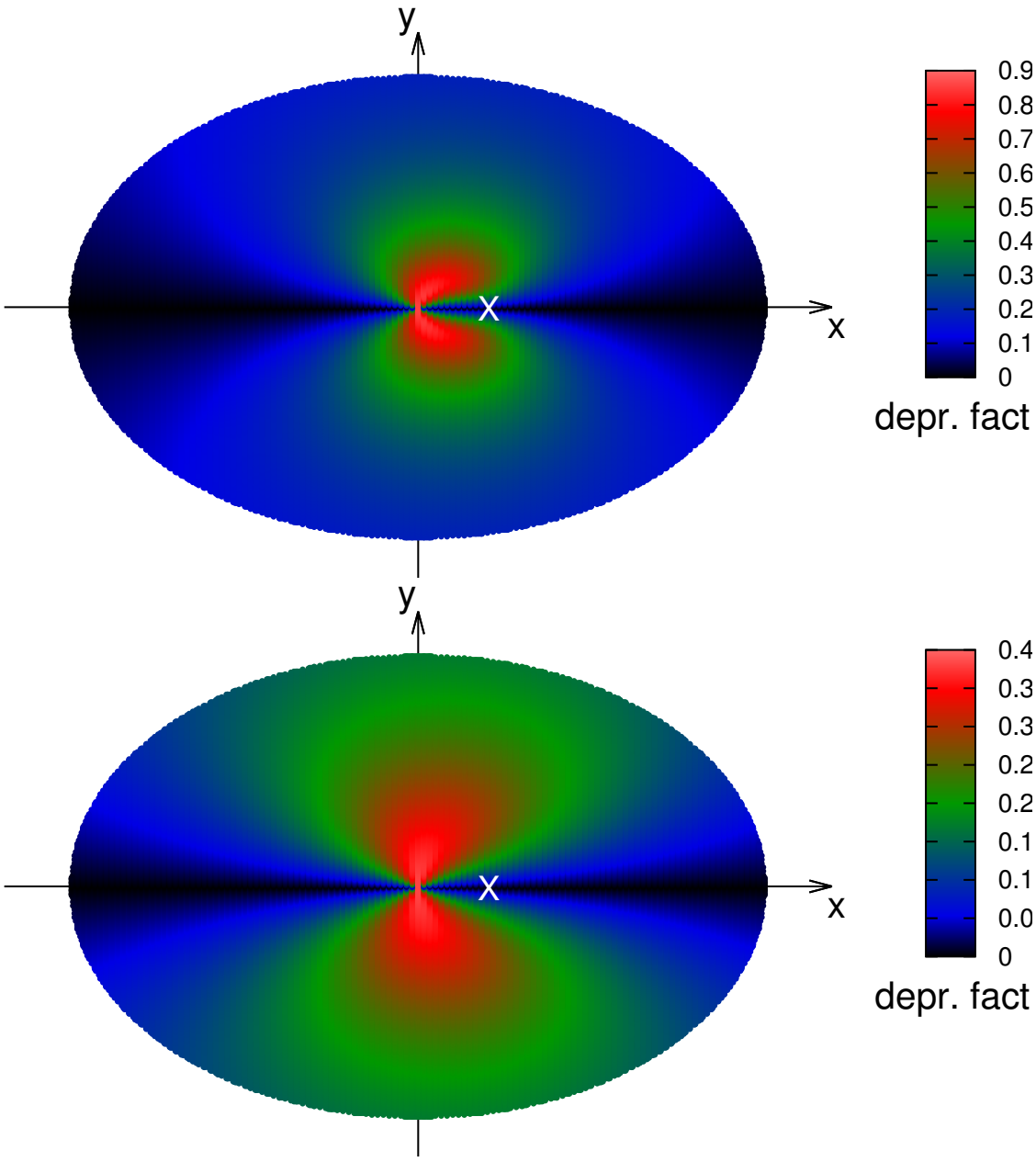


Figure 3.4: Colour coded deprojection factor in the $(x, y)$-galactocentric Cartesian plane at altitudes of $|z| = 5\,\mathrm{kpc}$ (top panel) and $|z| = 20\,\mathrm{kpc}$ (bottom panel), limiting $\sqrt{x^2 + y^2} < 45\,\mathrm{kpc}$. The white cross denotes the Sun's position; note the different colour scales of the plots.

SBA12 show that the mean streaming motion can be recovered from the position-dependent component of velocity ($U$) in the local frame[10]:

$$U + U_0 = \mathbf{v} \cdot \hat{\mathbf{e}}_x = v_R \cos\alpha + v_\phi \sin\alpha, \tag{3.19}$$

where the velocity vector of the star $\mathbf{v}$ is projected by $\hat{\mathbf{e}}_x$ onto the local radial direction, $U_0$ is the $U$-component of Sun's velocity, $v_\phi$ is the effective rotation and

$$\alpha = \arctan\left(\frac{d \sin l \cos b}{R_0 - d \cos l \cos b}\right) \tag{3.20}$$

is the angle between the Sun-GC line and the projection onto the plane of the long axis of the velocity ellipsoid (see equations 26-28 in SBA12). Taking the average of (3.19), one gets:

$$\langle U \rangle = \langle v_\phi \rangle \sin\alpha - U_0. \tag{3.21}$$

By fitting a straight line in the $U$-sin($\alpha$) plane, one can recover then both the mean rotation and the $U$-component of Sun's velocity.

Finally the direct estimate of the mean streaming velocity of the sample can serve as consistency check that the $U$-estimator is not biased from proper-motion systematics: if it were, contradictory results would emerge between the two estimators given their different bias.

### 3.4.3 Interpretation of discontinuous kinematics

A broken kinematic signature across different sub-divisions of a sample is the consequence of either a distance error, a pipeline error or a stream-like structure. Provided that tracers of all prominent unrelaxed sub-structures have been removed from the sample, no matter how we sub-divide our sample, each sub-set must have the same kinematics. Whatever portion of the sky we look at, we have to obtain the same rotation estimate, within the uncertainty. In particular, splitting the sample in random sub-samples decreases the statistical signal-to-noise and consequently enhances the allowance for discrepancy in the signature between two bins. Therefore, an anomaly among different sub-divisions of the sample is even more concerning and serves as a strong indication for a bias in the analysis.

[10]We maintain the reference-frame notation of SBA12, as declared also in §3.3.

#### 3.4.3.1 Distance errors

When we recover the rotational signature of a population of stars where only $v_{\rm los}$ is available or reliable, an error in distance induces changes in both the angle between the line of sight and the vector from the star to the Galactic Centre, and the projection of $v_{\rm los}$ onto the plane tangential to the direction to the Galactic centre. Thus a distance error influences directly the rotational signature. This will not be a constant systematic, but a function of the position on the sky. Distance errors shift stars en masse in velocity space. Distance over-estimates tend to shift velocities upwards towards the escape velocity, thus narrowing the range of physically accessible velocities to be considered. An increase in velocity will nearly always reduce the likelihood of a star when a model is being used to calculate the probability of the data; by carrying a star beyond the escape velocity they can even reduce the star's probability to zero. When a direct estimator is used, a distance over-estimate tends to increase the computational weight of a given star.

This effect does not change the sign of the rotation contribution of individual stars, but it distorts the values with a significance that scales with distance from the observer. In Fig. 3.5 we see that the factor by which the star's velocity is multiplied increases almost monotonically with heliocentric distance. Indeed, distant stars are intrinsically a bad tracer if proper-motion is not available: at large radii the streaming motion $v_\phi$ contributes very little to $v_{\rm los}$; in fact $v_{\rm los} + \mathbf{v}_\odot \cdot \hat{\mathbf{s}} \approx v_r$ and $v_\phi$ is dominated by the tangential component of velocity (hereafter $\mathbf{v}_\mu$). In order to change $v_{\rm los}$ by a small amount, $v_\phi$ needs to change significantly: the signal-to-noise scales inversely with radius making rotation estimated via $v_{\rm los}$ of distant stars, dominated by noise.

#### 3.4.3.2 Stream-like structures

A discontinuous rotational signature can hint at the presence of a sub-structure in the data: a stream for example can lead to inconsistencies between different rotation estimators, esp. for the $v_{los}$ based estimator and model. Even though we do not hunt for substructures here, we must allow for unidentified accretion remnants mimicking rotation. To break the degeneracy between biased data and a real stream, we need to identify what is characteristic of a substructure and what is not. An unrelaxed substructure is identified as such by its peculiar signature in parameter space (both physical and kinematic). Correlations between kinematics and metallicity/colour are examples of signatures of a possible accretion event. By contrast binning in luminosity type acts like a random selection w.r.t. kinematics and hence a coherent signature

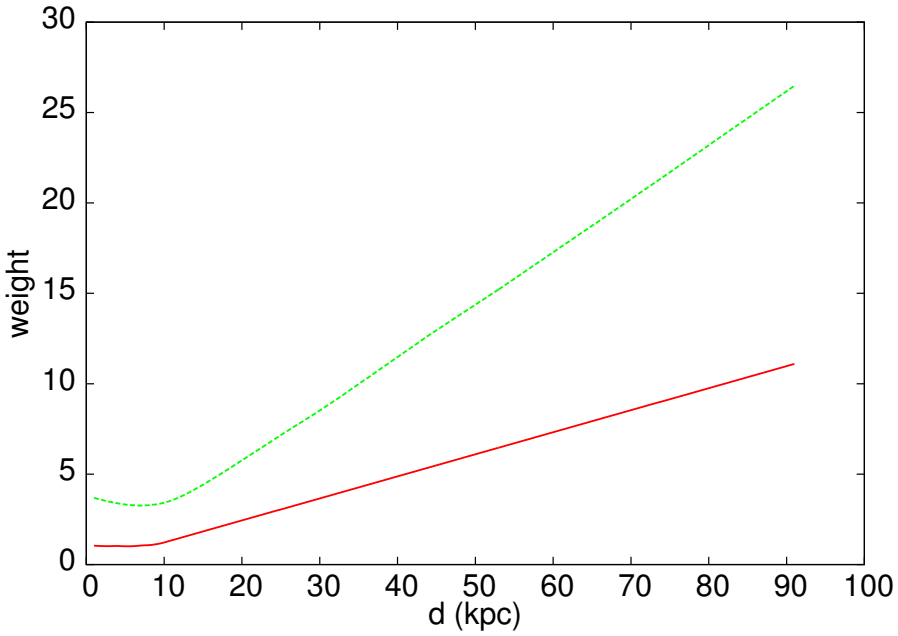


Figure 3.5: We slide each star in the FS12 sample (see §4.2) along the l.o.s. to put it at distance $d$, then calculate $\langle 1/(\hat{\mathbf{s}} \cdot \hat{\mathbf{e}}_\phi) \rangle$ and plot it versus $d$ (green line). The red line represents the minimum $1/(\hat{\mathbf{s}} \cdot \hat{\mathbf{e}}_\phi)$: both are almost perfectly linear functions of distance beyond 10 kpc.

across the bins must be observed. Further, provided that all detectable over-densities in observable space have been removed, binning the sample spatially implies that even if there are remnants of accretion events left they are washed out by the smooth component and thus a coherent signature is expected across all the sub-samples.

#### 3.4.3.3 Pipeline Systematics

Hot metal-poor stars, have spectra almost free of metal-lines: when we determine the line-of-sight velocity from their very broad Balmer lines, the estimate is highly vulnerable to template misfits. A correlation between a template/line index and kinematics for example, is consistent with a pipeline systematic misfit.

## 3.5 Conclusions

Stars' kinematics and position can provide a great deal of information on the morphology and formation history of the Galaxy. It is hard to obtain accurate astrometric measurements in the halo due to both its geometry and intrinsically metal-poor nature. It is not uncommon that the only kinematic information available for a single

star is its radial velocity, and the proper-motions are only useful when considered on average. In this setting it is imperative to search for the best accuracy achievable and essential to challenge the robustness of every result against plausible errors in the distance and velocity determination.

We have studied some ways in which distance errors propagate in a kinematic or even morphological study and proposed a method to detect inconsistencies in the distance calibration and quantify the fractional average distance error introduced. The method uses the idea from Schönrich, Binney & Asplund (2012) to track the average distance error through the fake correlation it produces between the velocity components, but is conceived to use the proper-motions only on average. We test the technique on eight control cases and conclude that the fractional average distance error is retrieved with an accuracy of $\sim 1.6 \pm 1.9\%$. The very same method can also be used to improve the distance calibration of a sample: one can estimate the average distance error for a given calibration and correct the latter iteratively until the average error is of the order of the formal error imposed by the sample size.

In the second part of this chapter, we critiqued the popular double-power law DF, spherical potential model Deason, Belokurov & Evans (2011a) recently used to estimate the rotation and ansiotropy of the stellar halo. We find that even though locally informative as rotation is concerned, the anisotropy measurement is biased by the $L^{-2\beta}$ term, which for $\beta < 0$ forces an intrinsic bimodality in velocity space. Further the model depends on strong assumptions on the potential and the density distribution of the tracers, which are both poorly constrained. Therefore, we explored the possibility to estimate the rotation of the stellar halo directly. Each estimator extracts rotation using a different manipulation of the phase-space coordinates so that if there is any significant systematics present in the sample, the results are inconsistent with each other, thus signalling a problem with the data. This is the principle that allows us to use proper-motions, even though they have uncertainty of the order of the signal. The errors on the tangential motion are assumed random so that they can still provide useful information on average: if on the contrary the sample suffers any significant systematic, the estimators are biased in a different way, so output conflicting results, and we do not accept their estimates. Finally, besides offering a robust estimate of rotation, model-free estimators can also quantify how much model assumptions bias the recovered picture.

# Part II

# Results for the Milky Way stellar halo

# Chapter 4

# Blue Horizontal Branch stars as tracers

## 4.1 Introduction

Stars in the solar neighbourhood have well-determined kinematics (e.g. GCS, RAVE: Nordström et al. 2004; Steinmetz 2003), however, they contain small numbers of halo stars prohibiting an uncontaminated and unbiased halo sample (Kinman 1995). To study the Galactic halo in situ[1] we need tracers that are bright enough to allow for distances $> 5$ kpc. The studied stars should not be easily mixed up with other types of stars, e.g. foreground dwarfs from the Galactic disk(s) and have a relatively well-determined absolute magnitudes, which should be also confined to a narrow range. The danger is otherwise to aggravate the selection function of the survey: given a certain magnitude limit, for every spectral type there will be a specific threshold distance further than which stars become effectively invisible. For a population with a narrow distribution of absolute magnitudes, this threshold is roughly constant for all stars in the sample and it is trivial to determine to which spatial extend our inferences hold. On the other hand, if we enlarge the sample including more luminous objects, they contribute unevenly to the overall distribution: at large distances it is then unclear whether the number-density drops as the star-count suggests or if we are in the region where only one of the different spectral types considered is actually visible. For example with an instrumental observational limit of 18 mag, at 4 kpc we can see stars with absolute magnitude down to 5 mag, at 10 kpc down to 3 mag (main-sequence stars with effective temperature 4,000-10,000 K and white dwarfs become invisible) and at 50 kpc we can only see stars brighter than 0.5 mag (Blue Horizontal

[1]By in situ we refer to stars that are actually observed there, rather than stars which are extrapolated to be there via orbit-integration.

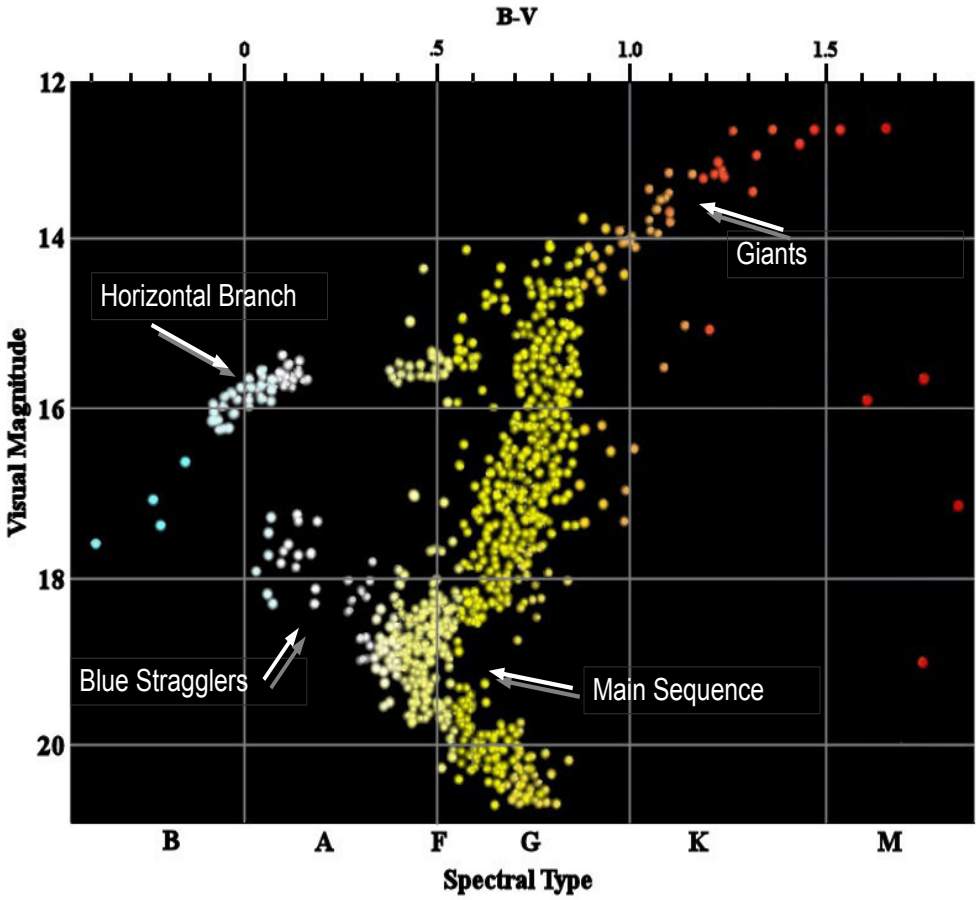


Figure 4.1: Colour-magnitude diagram for Messier 3 showing the spectral type and magnitude difference between main-sequence, giants, blue stragglers and horizontal branch stars. Adapted from _M3_color_magnitude_diagram.jpg (R. J. Hall) under the Creative Commons ShareAlike 1.0 License.

Branch stars, RR Lyrae and giants). See for reference the colour-magnitude diagram of M 3 in Fig. 4.1.

Further, distances for halo stars are obtained spectro-photometrically: this implies that the uncertainty on the distance correlates with how noisy is the spectrum of a star. To assign an absolute magnitude to the star one needs to identify its luminosity type (Beers et al. 2000) and therefore to have an accurate assessment of its surface gravity $\log g$. The danger of misclassification is especially an issue for the noisy spectra of remote and hence faint objects deeper in the halo (e.g. Yanny et al. 2000; Sirko et al. 2004). For example, parameters in automated pipelines can be drawn into unphysical values or into points that are a local minimum by averaging over the most likely points, in case the pdf is not unimodal (see §4.4.3, Lee et al. 2008a).

Late type giants are problematic by their easy confusion with dwarf stars and by

their steep relation between colours and magnitude, which also amplifies reddening uncertainties. The optimal objects would be RR Lyrae stars (cf. Klement 2010), but in addition to spectroscopy they require a good coverage in time to resolve their oscillations and at this stage are not yet globally available (Sesar et al. 2010).

Blue horizontal branch (BHB) stars have the advantage of a relatively narrow absolute visible magnitude distribution that shows only weak dependence on metallicity. It should be kept in mind, however, that there is the blue hook (see Whitney et al. 1998; D'Cruz et al. 2000; Brown et al. 2001, for a definition) with less luminous objects on the hot side of the BHB, and that, despite most Galactic stars being redder in colour, they mingle in observations with the fainter blue stragglers and main-sequence objects. Three techniques have been developed to identify BHB stars by looking at hydrogen lines: measuring the size (Philip & Adelman 1993) or the steepness (Chalonge & Divan 1973) of the Balmer Jump and using the width of the Balmer lines (Searle & Rodgers 1966). After identification BHB stars are treated as standard candles. It should be kept in mind, however, that the use of BHB stars biases any sample to old, metal-poor populations. Moreover, theoretical models show a metallicity dependence of order $\sim 0.3\,\mathrm{mag}$ in the V-band magnitude (and similarly in the Johnson g-magnitude), emphasizing the need to control any result not only for contamination, but also for absolute magnitude uncertainties (see §4.4).

By far the largest sample of BHB stars comes from the Sloan Digital Sky Survey (SDSS, Eisenstein et al. 2011). SDSS and its sequel Sloan Extension for Galactic Understanding and Exploration (SEGUE) obtained an unprecedented photometric and spectroscopic mapping of the Milky Way galaxy, allowing the study of nearby dwarf populations as well as giant star populations up to 100 kpc from the Sun (Yanny et al. 2009). The SEGUE Stellar Parameter Pipeline provides line index measurements for the spectra and the three primary stellar parameters: effective temperature ($T_{\mathrm{eff}}$), surface gravity ($\log g$) and metallicity ([Fe/H]) (cfr. Lee et al. 2008a,b; Allende Prieto et al. 2008). Alternative methods have been developed to estimate the above parameters: we will adopt the estimates produced by the method of Wilhelm, Beers & Gray (1999) as the latter was particularly designed for hot stars.

## 4.2 Sample selection

We construct two samples from the 9th Data Release (DR9, Ahn et al. 2012) of SDSS: one selected via filters in colour, surface gravity and effective temperature and another selected by Xue et al. (2011) via direct use of Balmer lines. We will refer to

the first sample as photometrically selected calling it FS12 and to the second sample as spectroscopically selected, calling it X11. Even though the FS12 sample is drawn from data for which spectra are in fact available, the chosen labels aim at emphasizing the difference between using and not-using Balmer-line shape measurements in the selection process (see §4.2.2). To both samples we apply geometric cuts to exclude the disk and Sagittarius.

### 4.2.1 Selection via photometry

We query for BHB candidates in the SDSS database[2] using an SQL query we report in Appendix A. A second-stage selection aims at minimizing contamination from other spectral types within the generous colour box which SDSS uses to classify stars as likely BHB. We benefit from the extensive work done by Yanny et al. (2000) and Sirko et al. (2004), who revised the BHB identification procedure working with SDSS data.

Both agree on an initial colour filter such as: $g-r \in [-0.4, 0]$ and $u-g \in [0.8, 1.4]$. The comparison with ultra-violet colour $u-g$ provides indirect information on the Balmer jump and hence is an indirect gravity estimator. Still in this colour regime (which we will denote $\mathcal{R}$) there are contaminations from other A-type stars (mostly blue stragglers) and from F stars due to "intrinsic variations and photometric errors" (Sirko et al. 2004). While F stars can be filtered out by their lower temperature, A-type stars mostly differ by their surface gravity, so that both groups consequently use the Balmer line width to separate them. Their understanding of the problem can be summarized as follows: main-sequence A stars can be distinguished from lower surface-gravity BHB stars (which lie in the region $\log g \in [2.5, 3.2]$) by looking at the width of the $H\delta$ lines alone (Yanny et al. 2000) or $H\delta$ and $H\gamma$ lines (Sirko et al. 2004). The discrimination is not sharp and one must still expect contamination of order $\sim 10\%$ (Sirko et al. 2004). The situation deteriorates towards fainter magnitudes and hence noisier spectra: at $g > 18$Sirko et al. (2004) estimate contamination of order $\sim 25\%$. Indeed Lee et al. (2008a) show that the low spectral resolution in SDSS and limited signal to noise lead to larger uncertainties in stellar parameters including $\log(g)$ up to a level where subgiants and dwarfs get mixed up. This represents a serious challenge for the proposition that stars can be cleanly sorted in type using only the $\log(g)$ estimates from the SEGUE stellar pipeline (e.g. Carollo et al. 2007, 2010; Deason, Belokurov & Evans 2011a).

[2]http://www.sdss3.org

According to the dangers reviewed in the previous paragraph, we minimise the risk of contamination from other luminosity types, by selecting:

$$\begin{cases} g < 18 \\ 2 < \log g < 3.3 \\ 0.8 < (u-g)_0 < 1.4 \\ -0.4 < (g-r)_0 < 0 \\ 7250 < T_{\rm eff}/{\rm K} < 9700 \\ |z| > 4\,{\rm kpc}\ \&\ r > 10\,{\rm kpc} \\ (\alpha,\delta) \notin {\rm Sgr} \end{cases} \tag{4.1}$$

where the last two criteria are geometric cuts to exclude disk stars and members of the Sagittarius stream (e.g. Deason, Belokurov & Evans 2011a)[3]. We remark, that the estimates of surface gravity, effective temperature and metallicity are from the analysis of Wilhelm, Beers & Gray (1999) as it was specifically designed for hot stars. This leads to a sample of 1585 stars. In Fig. 4.2 we plot the geometry of the sample: the top panel shows the distribution in the $(x, y)$-plane colour-coded according to the z-coordinate (Cartesian Galactocentric reference frame) and in the bottom panels we show the heliocentric (left) and galactocentric (right) distance distributions.

We further investigate the purity of our sample in the colour-colour plane and find grounds to suspect the presence of the *blue hook* in the stripe $-0.3 < (g-r)_0 < -0.2$ (see Fig. 4.3), which suggests that a $u-g$ dependent cut in the $g-r$ colour might more cleanly identify the horizontal branch. To a zero-th order approximation, a constant cut at $(u-g)_0 > 1.15$ (rather than $(u-g)_0 > 0.8$) is cost-effective to mask it out and will therefore be our initial choice to calibrate the reliability of (4.1). The tighter cut in $(u-g)_0$ shrinks the sample to 1297 objects.

Even though our stringent selection criteria cause a drop in the sample size w.r.t. previous works (e.g. the sample of Deason, Belokurov & Evans (2011a) has $\sim 3500$ stars due to more generous filters in surface gravity and apparent magnitude), none of our filters correlate with kinematics and thus they do not prejudice our analysis. They also do not bias the distributions of physical parameters such as for example metallicity: the distribution peaks at $[{\rm Fe/H}] = -1.9$, in good agreement with the estimates from the samples of Xue et al. (2008, 2011), who do not impose any cut in surface gravity or apparent magnitude for example. Further, we judge the increase in the statistical noise in the model fits a convenient price to pay in light of the reduced risk of contamination. The fact that our selection criteria in $\log g$ and $g$-band magnitude

[3]The masking of Sgr is performed in right ascension-declination space $(\alpha, \delta)$ according to the polygon kindly provided by the authors of Deason, Belokurov & Evans (2011a) on request (see Appendix B).

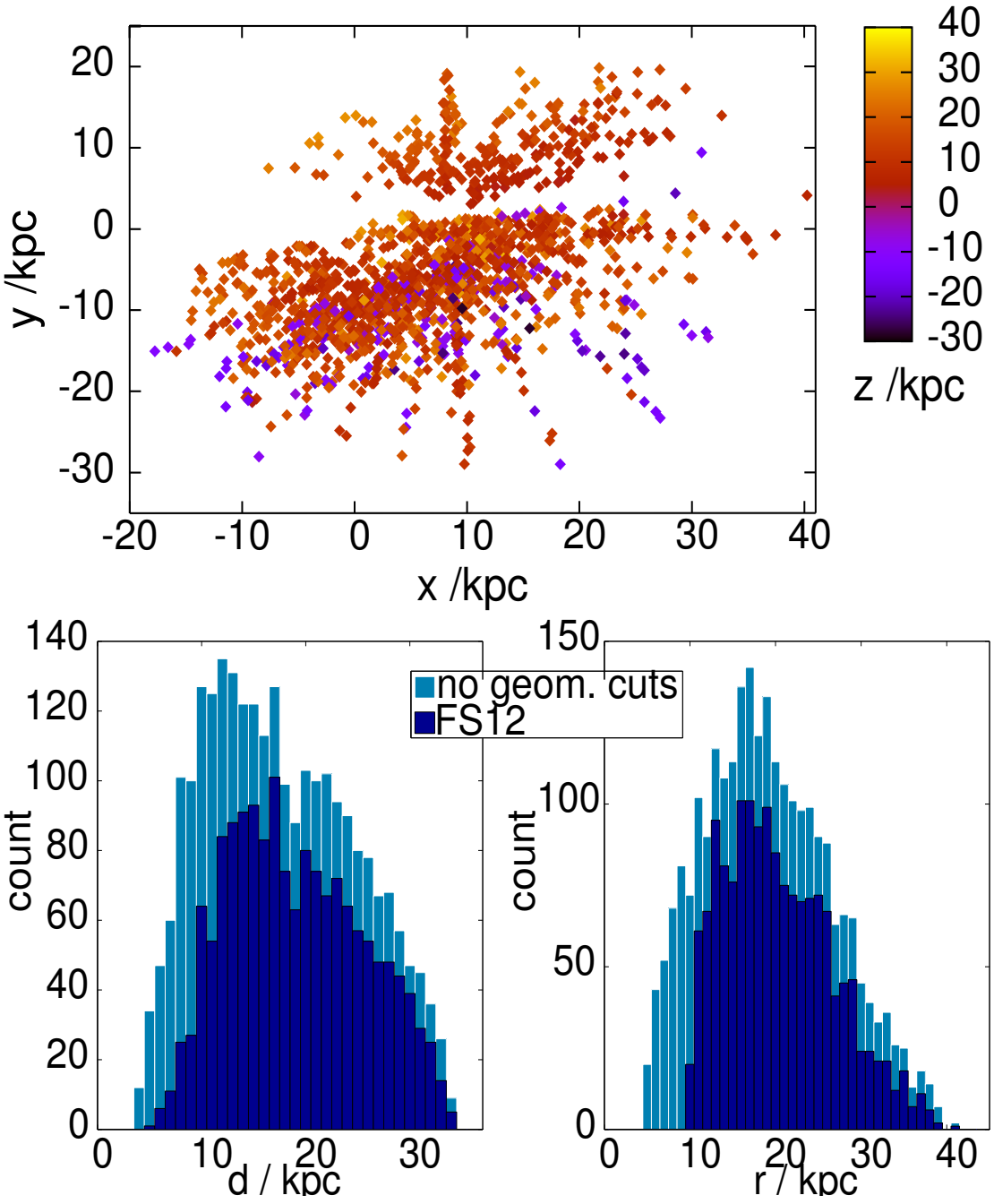


Figure 4.2: Geometry of the FS12 sample (top panel): distribution in the $(x, y)$-plane with the origin at the Galactic Centre, altitude z is colour-coded (Cartesian Galactocentric reference frame). 207 out of 1585 stars (13%) are in the southern galactic hemisphere. The bottom panels show the heliocentric (left) and Galactocentric (right) distance distribution. The dark blue histograms refer to the FS12 selection (4.1), while the light blue ones show the sample without the geometric cuts.

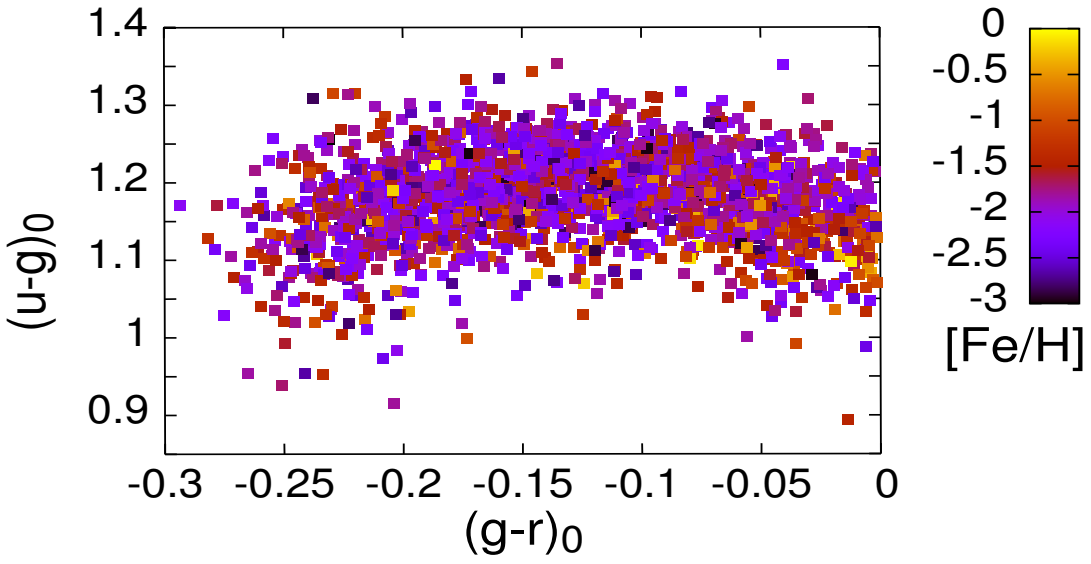


Figure 4.3: BHB stars drawn from SDSS DR9 included by the selection (4.1) in the $(u-g)_0$, $(g-r)_0$ plane. Every point is coloured according to the metallicity of each star. The commonly adopted colour cuts appear to include some part of the *blue hook*.

do not bias the inferred kinematics is confirmed by a second sample selected via more generous photometric filters, but combined with Balmer-line shape measurements, which presents an homologous rotational signature (see following Section).

### 4.2.2 Spectroscopic selection

Xue et al. (2011) presented a spectroscopically selected sample of $\sim 5000$ BHB stars drawn from SDSS DR8 and which has been kindly provided to us. The advantage w.r.t. a selection via photometric information is that the three independent methods of identification separate BHB stars from blue stragglers and main-sequence stars more cleanly by combining colour cuts with Balmer-line shape measurements: the scale width-shape method fits Sersic profiles to the $H\gamma$ lines (cf. Clewley et al. 2002; Sirko et al. 2004; Xue et al. 2008, 2011). The parameters that describe such a profile are essentially the amplitude ($a_\gamma$), the dispersion ($b_\gamma$) and the steepness ($c_\gamma$) of the modified Gaussian that is fitted to each $H_\gamma$ line:

$$y(\lambda) = 1.0 - a_\gamma \exp\left[-\left(\frac{|\lambda - \lambda_0|}{b_\gamma}\right)^{c_\gamma}\right]. \tag{4.2}$$

We retrieve Xue et al. (2011)'s stars in DR9 and use the astrometry of the latest data release; we exclude contributions from the disk and the region of Sagittarius

stream by applying the same geometric criteria of the previous section: this leads to a sample of 2563 stars (2455 with [Fe/H] $\in [-3, 0]$). We show its geometry in Fig. 4.4. Among these $\sim$ 1400 have $c_\gamma \geq 0.95$, which is the limit value for which the method of Clewley et al. (2002) has been calibrated. Given that neither Sirko et al. (2004) nor Xue et al. (2011) have re-calibrated the method in the region $c_\gamma \geq 0.95$ we cautiously flag these stars and first consider them in our analysis but later analyse them separately.

We retrieve 1400 stars of our photometrically selected sample in the sample of X11, which implies that almost 90% of the FS12 sample is included in X11: the latter sample was selected using a classification based on both Balmer-lines and colour cuts and therefore constitutes the more robust identification of BHB stars to the current understanding. Consequently, cross-matching our FS12 sample with the sample of Xue et al. (2011) provides a rough estimate for the level of contamination from non-BHB stars that we can expect: i.e. $\sim$ 10%, which very well matches the independent prediction of Sirko et al. (2004). It is interesting to see that larger samples selecting with more generous filters in surface gravity and apparent magnitude suffer a higher degree of impurity. When we compare the sample of Deason, Belokurov & Evans (2011a), hereafter D11, with X11, only 47% of their stars have Balmer-line shapes identifiable with BHB type stars, even though D11 drew their sample from DR7 and X11 is based on DR8, which is significantly larger. If we exclude high surface gravity and faint stars from the sample of D11 in DR7, the proportion of stars matched by Xue et al. (2011) increases to 71%, meaning that the majority of the stars we rejected are non-BHB stars according to Xue et al. (2011).[4]

## 4.3 Distance calibration

Estimating the absolute magnitude of BHB stars is also a non trivial problem. Sirko et al. (2004) show that the assumption of a constant luminosity and mass, temperature-independent for $T_{\rm eff} \in [8000, 12000]$K, (Baev, Markov & Spassova 2001) would lead to an average overestimate in distance modulus of 0.18 mag with respect to a temperature-dependent relationship for the luminosity (Dorman, Rood & O'Connell

[4]Note that the value reported in D11 at the beginning of their §3.1 in these regards can be somehow misleading. The authors report that 88% of the stars of Xue et al. (2008) (based on DR6) is found in their larger sample. But this actually means that only 41% of the stars in D11 sample is found in the one of Xue et al. (2008) and we do not have any reassurance on the level of contamination for the remaining 59% of the stars in their sample of ($\sim$ 3274 objects).

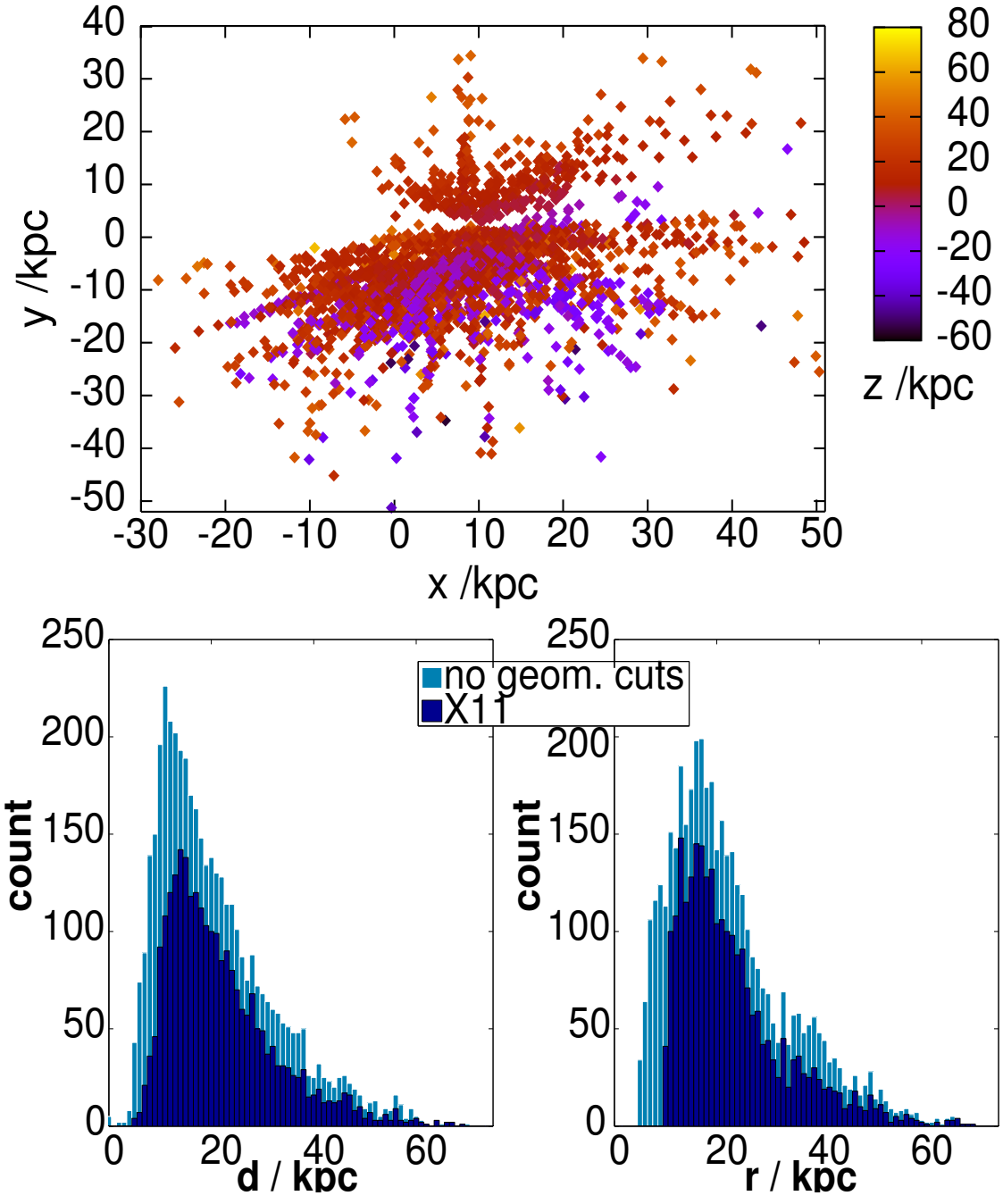


Figure 4.4: Same as Fig 4.2, but for the X11 sample. Here 526 out of 2563 stars (21%) belong to the souther galactic hemisphere.

1993). Furthermore, the popular value of $M_g = 0.7$ was originally calculated by Layden et al. (1996) for halo RR Lyrae star and then associated to BHB stars by Yanny et al. (2000), who compared RR Lyrae and BHB stars in the globular cluster Pal 5. However, the justification of this approximation depends on the blue response of the filter in use (Yanny et al. 2000). Preston, Shectman & Beers (1991) report that BHB stars in the blue end of $\mathcal{R}$ are 0.7 mag fainter in $g$ than the RR Lyrae stars[5]. Therefore the uncertainty reported for the absolute magnitude of BHB stars needs to be interpreted generously given the non-standardization of the blue filters and the consequent difficulty in comparing estimates from different surveys.

Some previous studies attempted to provide a magnitude calibration for BHB stars in the *ugriz* filter system of SDSS, based on a combination of theoretical expectations and calibrations on globular clusters (e.g. Sirko et al. 2004; Xue et al. 2008). Yet those calibrations have not been provided in a user-friendly form, such that many current studies (e.g. Deason, Belokurov & Evans 2011a; De Propris, Harrison & Mares 2010; Niederste-Ostholt et al. 2010) still use the approximation of a constant absolute magnitude $M_g = 0.7$.

More importantly, the classic calibrations (like Sirko et al. 2004) rely on globular clusters, bearing the usual uncertainties in their distance moduli and reddenings. Further we know that (a) the Horizontal Branch (HB) stars of clusters vary between clusters for reasons that are incompletely understood (e.g. clusters of the same metallicity can either be very red or blue: specifically metal-poor *clusters with high central density present bluer HB morphologies,* Suntzeff, Kinman & Kraft 1991; Buonanno 1993) and (b) there are systematic differences between globular-cluster and field halo stars (e.g. the binary fraction among extreme HB stars is an order of magnitude lower in globular clusters than among field stars, Han 2008).

Given these problems, it is crucial to create a formula that fulfils the following criteria:

- it covers the principal dependences of the HB magnitudes (i.e. colour and metallicity)
- it is provided as a simple analytical formula that gives the absolute magnitude from observables;
- it is validated by/calibrated on field stars.

[5]Both Layden et al. (1996) and Preston, Shectman & Beers (1991) report results for the $V$ magnitude, which Yanny et al. (2000) usefully converts into the $g$ magnitude via the colour transformation of Fukugita et al. (1996).

The aim of this section is to construct such a calibration in the *ugriz* filter system of SDSS and assess its performance w.r.t. currently available calibrations by testing it on BHB field stars. At first glance the last point seems out of reach, considering that there are virtually no reliable Hipparcos parallaxes for BHB stars and especially none for stars within the magnitude range of the SDSS. The sample we have available (FS12 and X11)[6] cover a distance range out to $\sim 50\,\mathrm{kpc}$ and the error in proper motion measurements for single stars is approximately as large as the true signal from solar reflex motion and velocity dispersion. However, by large number statistics the proper motions still bear valuable information and we can obtain statistical distance estimates for field BHB stars via the $V$-$W$ correlation method (see §3.3). The statistical uncertainty simply sets our formal error bars, but we do need to be concerned with the systematic terms: to cope with the known systematics of frame-dragging, we will use the proper-motion correction derived by Schönrich (2012) on the Schneider et al. (2010) quasar sample in Galactic angles $(l, b)$. This correction shrinks the statistically predicted mean fractional distance error by 0.02 (see §4.3.2) and is associated with a systematic which is anyway smaller than the noise by roughly a factor of 20 (see 5.2.1.1).

Finally, models of stellar evolution predict the locus of the HB. Hence, our empirical calibration will also provide a test of these models. We find that our calibration agrees quite closely with BASTI isochrones (Pietrinferni et al. 2004, 2006), and makes stars significantly more luminous than Dartmouth isochrones predict (Dotter et al. 2007, 2008).

### 4.3.1 Calibration construction for the BHB

Stellar evolution models predict a steep decline of the intrinsic magnitude of BHB stars towards bluer colours (e.g. BASTI, Dartmouth isochrones) in concordance with photometry of globular clusters. Further, they predict metal-poor BHB stars to be systematically brighter than their metal-rich counterparts. The ideal calibration would encompass the main dependences with the simplest possible form: in particular, the metallicity dependence, modelled best by a polynomial for the increasing trend in

[6]To validate our calibration, we use two samples drawn from SDSS DR9 as laid out in §4.2.1 and 4.2.2. We do not impose the geometric cuts that removes the disk though as we are not interested in isolating the halo kinematics, but rather test distance estimates for a specific type of stars, whichever component they belong to. All the same, the Sgr stream still needs to be removed as that can bias the $V$-$W$ correlation method: the stream is a coherent local feature and can therefore produce kinematic anomalies such as "sky-falling" just by the fact that it is an unrelaxed structure.

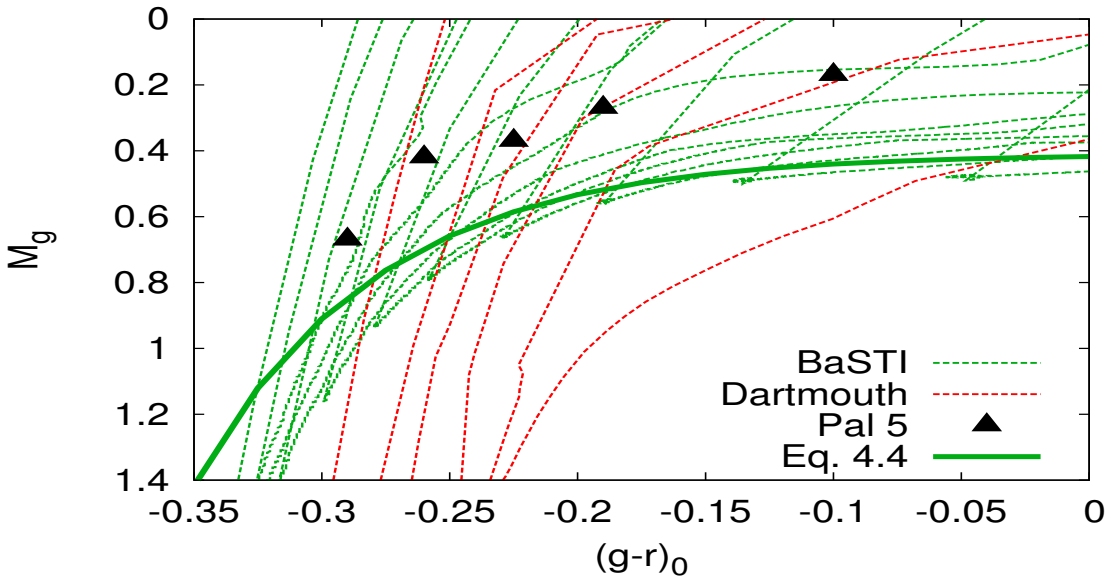


Figure 4.5: Colour-magnitude diagram showing the BaSTI isochrones (dashed green lines), BHB stars in Pal 5 (black triangles) and our calibration 4.4 (solid green line). Dartmouth isochrones are plotted for comparison in dashed red lines. The magnitude difference between isochrones and cluster data, had already been suspected by Sirko et al. (2004) to be due to uncertainty in the distance modulus of Pal 5.

metallicity, and the steep decline on the blue side, which we model by an exponential in $(g-r)_0$ colour:

$$M_g((g-r)_0, [Fe/H]) = a_0 \exp(a_1(g-r)_0) + p([\mathrm{Fe/H}]), \tag{4.3}$$

where $p$ is a polynomial in one variable, $p \in \mathbb{R}_n[[\mathrm{Fe/H}]]$. Parameters of (4.3) are empirically adjusted to mimic both the trends seen in metal-poor clusters and the ones predicted by the isochrones: among the ones available the BASTI isochrones drive our calibration towards a better performance on field stars than for example the Dartmouth ones (see §4.4.1). One must account for photometric and reddening uncertainties though, which make it likely that a star with assigned blue-hook colours will conceivably be still relatively luminous: thus the proposed calibration shall bend a bit less steeply towards the blue hook than the isochrones. In Fig. 4.5 we plot the BaSTI Isochrones at $[\mathrm{Fe/H}] = -1.5$ and the BHB stars belonging to Pal 5 according to the selection of Yanny et al. (2000). Our calibration is adjusted by eye to follow the isochrones, but bend less steeply as the data in Pal 5 suggest. Over-plotted are also the Dartmouth isochrones at the same metallicity: as we quantify in §4.4.1, they would drive the calibration towards the faint side.

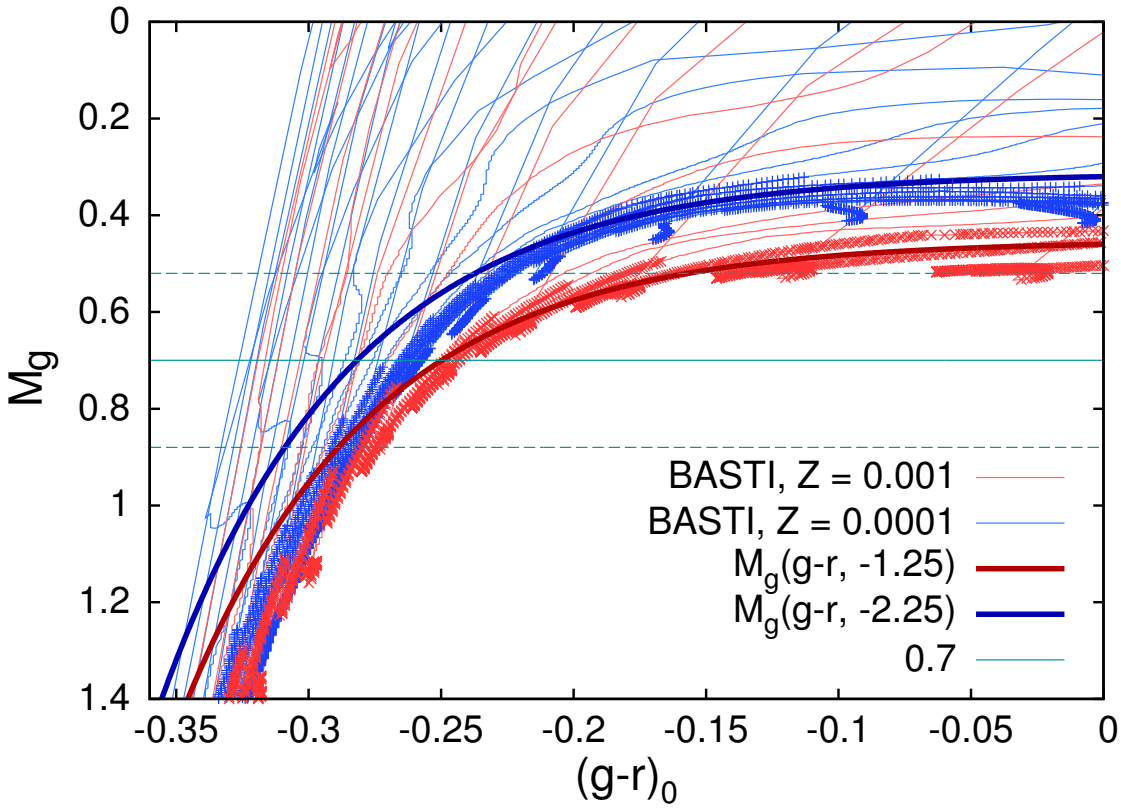


Figure 4.6: Luminosity as a function of colour for the Blue Horizontal Branch at different metallicities. Thin lines represent the predictions from the BASTI isochrones (Pietrinferni et al. 2004, 2006) as do the dots that are associated with a mass spacing of $3 \cdot 10^{-6} M_{\odot}$. Thick lines depict our colour and metallicity dependent approximation for $M_g$. Blue lines refers to metal-poor stars ([Fe/H] $\in ]-3, -2]$) and red lines to metal-rich stars ([Fe/H] $\in ] -2, 0]$). The light blue line depicts the popular approximation of constant $M_g = 0.7$ and light blue dotted lines represent an uncertainty of $\pm 0.18$.

Our calibration will not be a fit to the data, but rather it will be validated by data. The crucial difference is that a priori we do not know if a data sample is perfectly unbiased and how eventual systematics can bias the parameters of (4.3). On the other hand a validation method can rely on the limited information content, which proves to be systematics-free. In the BHB star sample of Xue et al. (2011) for example, we flag a peculiar kinematics associated with metal-poor high $c_\gamma$ stars (see §5.4.1): these stars affect the correlation between the $V$ component of velocity and the angle term $\sin(l)\sin(b)\cos(b)$, but have negligible effect on the relationship between $W$ component of velocity and the same angle term arising from the rest of the sample. Further, it will turn out (see §4.3.2) that the accuracy of our calibration inspired by theoretical models, is within the measurement errors indistinguishable from the optimum fit to the field star data, and therefore there is no reason to "risk" a fit to the data. The end product is:

$$M_g((g-r)_0, [\mathrm{Fe/H}]) = 0.0075\ \exp(-14.0(g-r)_0) + 0.04\,([\mathrm{Fe/H}] + 3.5)^2 + 0.25, \quad (4.4)$$

which we show in Fig. 4.6 for two metallicities: $[\mathrm{Fe/H}] = -2.25$ (thick blue line) and $[\mathrm{Fe/H}] = -1.25$ (thick red line). Also shown in Fig. 4.6 are the BASTI isochrones (Pietrinferni et al. 2004, 2006). The discrepancy between the constant-luminosity approximation (horizontal light blue line in Fig. 4.6) and both the isochrones and our calibration exceeds the frequently adopted uncertainty band of $\varepsilon_{M_g} \sim \pm 0.18$ (light blue dotted lines in Fig. 4.6)[7]. Note that equation (4.4) is *a* possible calibration: the parameters are fitted by eye and multiple combinations of these as well as other functional dependences for both colour and metallicity are likely to provide a consistent description. Given the performance on the selected data though (see following section), there is no gain in exploring alternative solutions in this work.

Equation (4.4) does not depend on the $u-g$ colour as in our tests we could not detect any significant systematics for it once the $g-r$ and [Fe/H] dependences had been taken into account. The ideal calibration would account for more detailed chemical dependences like alpha enhancement though and particularly the abundance of Helium, which changes the absolute magnitude of a star. However, this information is generally not available for BHB stars because their Helium abundances can be very low due to depletion mechanisms (e.g. Faulkner 1967).

[7] An uncertainty of $\pm 0.18$ is commonly associated with the calibration of the horizontal branch, due to neglect of the temperature dependence (Sirko et al. 2004), or ($\pm 0.15$, 0.2) associated with the constant absolute magnitude approximation (e.g. Deason, Belokurov & Evans 2011a; De Propris, Harrison & Mares 2010)

### 4.3.2 Validation on field stars

We validate our calibration on our two samples of halo field BHB stars: on both we obtain a satisfactory accuracy:

$$f_{\rm FS12} = 0.01 \pm 0.04 \quad \text{and} \quad f_{\rm X11} = 0.02 \pm 0.03.$$

Given that the fractional average distance error from (4.4) is consistent with zero within the statistical uncertainties, our calibration can be regarded as the first direct measurement of the magnitude of field BHB stars in SDSS-SEGUE.

Use of the proper motion[8] correction of Schönrich (2012) reduced the average systematic distance error from $f = -0.03 \pm 0.04$ to $f = 0.01 \pm 0.04$ for the FS12 sample and from $f = -0.04 \pm 0.03$ to $f = 0.02 \pm 0.03$ for X11 using our calibration (4.4) for the distance estimate. Hence, the correction affects the average distance calibration by about 5% and hardly affects the relative distances between single bins.

In general the statistics are robust against outliers in the sense that when we remove stars with extreme proper-motion (i.e. tangential velocities above 900 km$s^{-1}$: less than 3% in FS12 and than 6% in X11), $f$ does not change for FS12 and changes by less than 0.01 for X11. The inclusion of the Sgr stream makes very little difference ($\Delta f \sim 0.02$) to our estimates. Indeed, adding the stream region to the samples, we find $f = 0.04 \pm 0.03$ on X11 and a similar difference on FS12: $f = 0.02 \pm 0.04$.

## 4.4 Comparison of other distance/magnitude assignments

Using the same method that allowed us to validate our calibration against field halo BHB stars (see §3.3), we can directly assess other distance calibrations.

### 4.4.1 BASTI versus Dartmouth isochrones

In Fig. 4.7 we plot the discrepancy between the absolute magnitude predicted by the Dartmouth isochrones and our calibration as a function of colour for four synthetic populations of $\sim 2700$ HB stars, each with different metallicity. The synthetic populations were generated with the Dartmouth web tool[9] and input parameters $([\alpha/Fe], \max M, \langle M \rangle, \sigma_M) = (0.4, 4.0, 0.5, 0.1)$, where $M$ is mass in solar masses. In

[8] We checked that the line-of-sight velocity correction given in the same paper has virtually no effect on our distance estimates.

[9] http://stellar.dartmouth.edu/models/shb.html

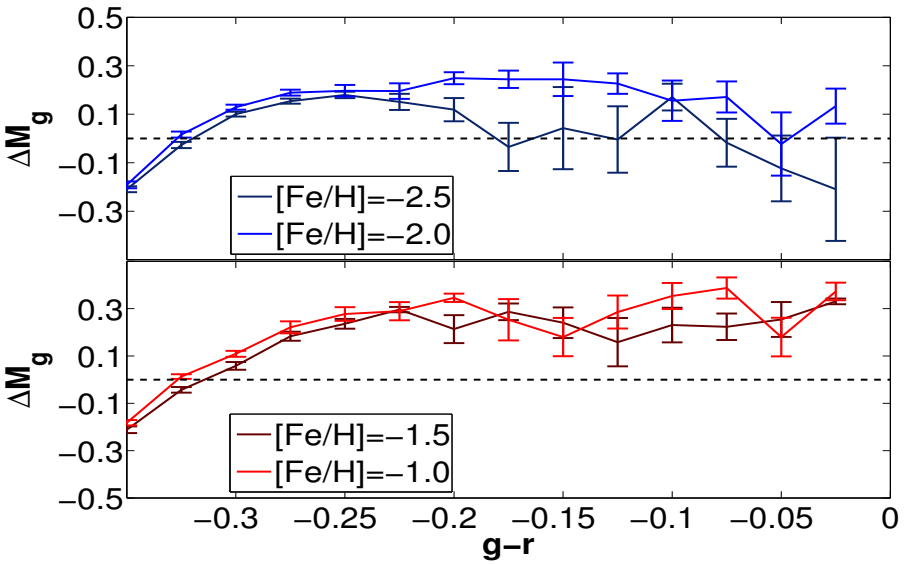


Figure 4.7: Discrepancy between the absolute magnitude predicted by the Dartmouth Isochrones and the one estimated via equation 4.4 as a function of colour for four synthetic populations of $\sim$ 2700 HB stars each with different metallicity. Upper panel: metal poor stars; bottom panel: metal rich stars. Error bars show the standard error on $\Delta M_g$ in each colour bin.

the colour range $g-r \in [-0.25, 0]$ (where our sample stars are concentrated) the offset is systematic and increases with metallicity: $\Delta M_g \sim 0.1$ for metal-poor objects and $\Delta M_g \sim 0.3 - 0.4$ for metal-rich objects.[10] Given that we achieve a fractional average distance error consistent with zero well within $1\sigma$ on our sample of field stars, estimating the stars' absolute magnitudes with the Dartmouth isochrones corresponds to underestimating distances by up to 4 % for metal-poor stars and by up to 17 % for metal-rich ones (see eq. 4.6 below).

In conclusion, the close agreement between our formula and the BASTI isochrones (see Fig. 4.6) implies that at least in the *ugriz*-system, they are superior to the Dartmouth isochrones.

### 4.4.2 Sirko et al. 2004

Sirko et al. (2004) estimated the absolute magnitudes of BHB stars as functions of colour for two metallicities ([M/H] $= -1$ and [M/H] $= -2$) by integrating the Kurucz model atmospheres[11] and then finding the point on the theoretical track that was closest to the observed star in colour space. They validated their calibration with

[10] Varying $[\alpha/Fe]$ by $\pm 0.2$ does not alter our conclusions.
[11] see http://www.kurucz.harvard.edu.

two globular clusters: NGC 2419 and Pal 5. With literature distance moduli for the clusters they found that their theoretical predictions were too bright by 0.13 mag for NGC 2419 and and too faint by 0.2 mag for Pal 5. They considered that these offsets could be accounted for by errors in the adopted distance moduli.

When we compare our calibration with theirs, it turns out that their relation is significantly fainter than ours: $f_{\rm Sirko} = -0.07 \pm 0.04$ on X11. Hence Sirko et al. (2004) predict shorter distances than our formula, which on our data produces a fractional average distance error consistent with zero well within $1\sigma$ (see §3.1). Moreover, application of the calibration of Sirko et al. (2004) (from their Table 2) requires several approximations: one needs to bin in effective temperature, surface gravity, colour and metallicity. Consequently the calibration is undefined for more than half of the stars in the sample.[12]

Our calibration is an analytic formula and not a complicated fitting process to the theoretical models. Since it produces significantly smaller distance errors than the calibration of Sirko et al. (2004), it is not only more convenient, but also more accurate.

### 4.4.3 Xue et al. 2011

The catalogue of BHB stars provided by Xue et al. (2011) includes distances estimated with a method very similar to that of Sirko et al. (2004): i.e. the authors determine the point on the theoretical tracks that is closest to the observed star in $u-g$, $g-r$ space and from the former estimate $M_g$. On X11, i.e. the catalogue of Xue et al. (2011) with the Sgr stream removed, their distances appear to be weakly biased towards distance underestimates: $f = -0.05 \pm 0.03$.

Marginalizing over metallicity, the fit to the colour dependence is comparable between their calibration and ours, within the formal errors. However, our formula (4.4) proves to be significantly more accurate at low metallicity when we marginalize over colour: Fig. 4.8 shows the different trend in metallicity of $f$ for the two distance

[12] To avoid binning the data, we compared the two formulas at fixed metallicity. We use [Fe/H] while Sirko et al. (2004) use [M/H], but this has no relevance in our discussion. When equating [Fe/H] and [M/H] and defining $\Delta M_g = M_g(\text{Sirko}) - M_g(\text{FS12})$, we find $\Delta M_g \in [0.07, 0.35]$ for metal-rich stars and $\Delta M_g \in [0.15, 0.41]$ for metal-poor stars. When we assume [M/H] = [Fe/H] + 0.3, which roughly corresponds to the Salaris, Chieffi & Straniero (1993) approximation [M/H] = [Fe/H] + $\log(0.638\exp([\alpha/\text{Fe}]) + 0.362)$ with $[\alpha/\text{Fe}] \sim 0.4$, the estimates above are unchanged. In conclusion, given their fainter magnitude, Sirko et al. (2004) calibration will produce shorter distances w.r.t. (4.4) by a $\Delta f$ of at least $-0.03$ and up to $-0.17$. Thus, it is reasonable to expect that on our samples their calibration would produce a distance error in the range $f \in [-0.15, -0.02]$, regardless of how data are binned.

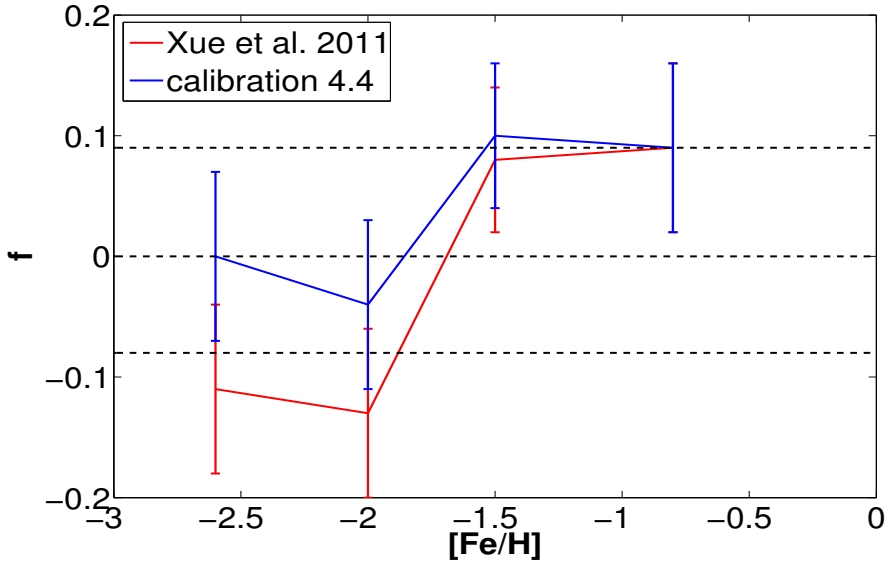


Figure 4.8: Average fractional distance error as function of metallicity on X11 sample, using both the calibration of Xue et al. (2011, red line) and equation 4.4 (blue line).

calibrations on X11 sample. In the range [Fe/H] $\in [-3, -1.8]$ the calibration of Xue et al. (2011) is associated with an average fractional distance error of $f \sim -0.12 \pm 0.04$, while equation (4.4) yields $f \sim -0.02 \pm 0.04$.

### 4.4.4 Deason et al. 2011b

Deason, Belokurov & Evans (2011b) propose a calibration of BHB stars derived from 10 star clusters published in An et al. (2008) covering a range in metallicity of [Fe/H] $\in [-2.3, -1.3]$. They detect a colour dependence, but no obvious metallicity dependence: however, given the considerable uncertainties in the adopted distance moduli and reddenings ($0.20 \pm 0.04$ mag or $\sim 10 - 15\%$, e.g. Gratton et al. 1997) and the systematic differences between clusters and field halo stars, this result is not enough to rule out a detectable metallicity dependence on field BHB stars. The calibration of Deason, Belokurov & Evans (2011b) is systematically fainter for metal-poor stars and brighter for metal-rich stars than ours and leads to a worse performance on the FS12 sample.

On the X11 sample we estimate a global fractional average distance error of $f = -0.02 \pm 0.03$ using their calibration ($f = 0.02 \pm 0.04$ for metal-rich stars and $f = -0.08 \pm 0.06$ for metal-poor stars), while on FS12 Deason, Belokurov & Evans (2011b) gives $f = -0.03 \pm 0.04$ ($f = 0.01 \pm 0.05$ for metal-rich stars and $f = -0.05 \pm 0.06$ for metal-poor stars). When investigating the detailed morphology of the halo, this

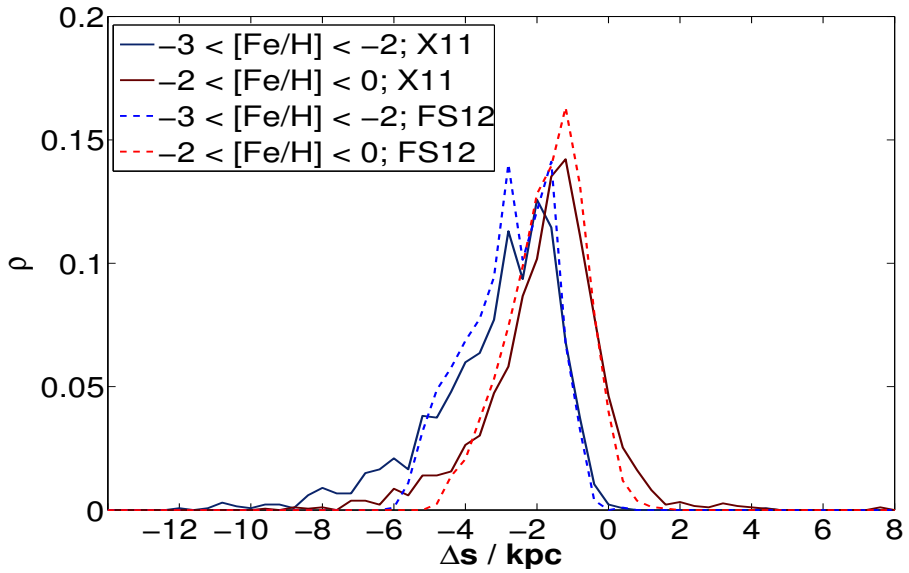


Figure 4.9: Density distribution ($\rho$) of the difference in distance calibrations for BHB stars from SDSS DR9: $\Delta s = s(M_g = 0.7) - s(M_g((g-r)_0, [\mathrm{Fe/H}]))$. Solid lines refer to X11 sample, dotted ones to FS12. The assumption of a constant absolute magnitude underestimates stellar distances w.r.t. the colour and metallicity dependent estimate for $M_g$ given by equation 4.4. The effect is more pronounced for the most metal poor objects (blue lines).

metallicity-related bias is both significant in magnitude and relevant for the research goal. E.g. substructures can be blurred and distorted by several kpc. In particular, the distances for metal-poor stars are underestimated, while that for the metal-rich ones are slightly overestimated: this gives rise to a kinematic bias acting in opposite directions.

### 4.4.5 Systematics of the approximation $M_g = 0.7$

Fig. 4.9 shows the discrepancy in heliocentric distance between the approximation $M_g = 0.7$ and our calibration on the two samples FS12 and X11: the constant absolute magnitude approximation tends to underestimate stars' distances by several kpc, the discrepancy is most severe for the most metal-poor objects.[13] A direct assessment of the constant absolute magnitude approximation via the method in §3.3 measures a global distance underestimate of $f = -0.13 \pm 0.03$ on X11 sample and of $f = -0.11 \pm 0.04$ on FS12 sample, the error being worse for metal-poor objects. In a

[13]W.r.t. Fig. 4.6, it might be helpful to know that for both samples the blue-end coincides with the tail of their $(g-r)_0$ colour distribution so that most of the stars are in the colour range where calibration (4.4) is more luminous than 0.7 mag.

kinematic analysis for example, this systematics can distort the velocity distribution of the sample and the fact that the effect is stronger on metal-poor objects can introduce a false gradient in the kinematics of sub-samples of stars with different metallicities.

We show that the method presented in §3.3 is sensitive enough to detect the metallicity and colour dependence of absolute magnitudes: for this purpose we bin both the spectroscopic and the photometric samples separately in colour and metallicity to measure the expected dependences on each quantity while marginalising over the other.

#### 4.4.5.1 Colour dependence

We divide X11 and FS12 into metal-poor, [Fe/H] $\in [-3, -2]$, and metal-rich stars, [Fe/H] $\in [-2, 0]$: each sub-sample is then split in colour bins, the sizes of which are adjusted to provide enough statistics for each colour band, and therefore meaningful error bars ($\geq 400$ stars per bin).

In Fig. 4.10 we plot $f$ against $(g-r)_0$: the red lines in the upper and lower panel are associated with the metal-rich stars in the X11 and FS12 samples respectively. The average relative distance error $f$ oscillates with colour: the trend is most pronounced in the X11 sample. This behaviour reflects the dependence of their absolute magnitudes on colour. It is not matched by the metal-poor part of either of the samples: the blue lines in Fig. 4.10 show that both the photometrically selected metal-poor stars (lower panel) and the ones spectroscopically selected (upper panel) exhibit an increase of $f$ with colour. The discrepancy in the colour-dependence at different metallicity is the footprint of the metallicity dependence of $M_g$. We remark that the metal-poor stars in the X11 sample (Fig. 4.10, upper panel, blue line) exhibit exactly the behaviour predicted in the previous section and are indeed associated with the poorest fit by a constant absolute magnitude. Also, it is at these metallicities that our distance calibration differs most from the one associated with $M_g = 0.7$: Fig. 4.9 clearly shows that our approximation precisely corrects for the underestimate in distance caused by the assumption of a too faint constant magnitude.

To assess the significance of these dependences w.r.t. the traditional level of uncertainty associated with the absolute magnitude of BHB stars (e.g. $\varepsilon_{M_g} = 0.18$, Sirko et al. 2004; De Propris, Harrison & Mares 2010; Deason, Belokurov & Evans 2011a; Newberg, Yanny & Willett 2009) we need to relate the fractional distance error to the effective error in absolute magnitude, $\varepsilon_{M_g} = M'_g - M_g$. Given that

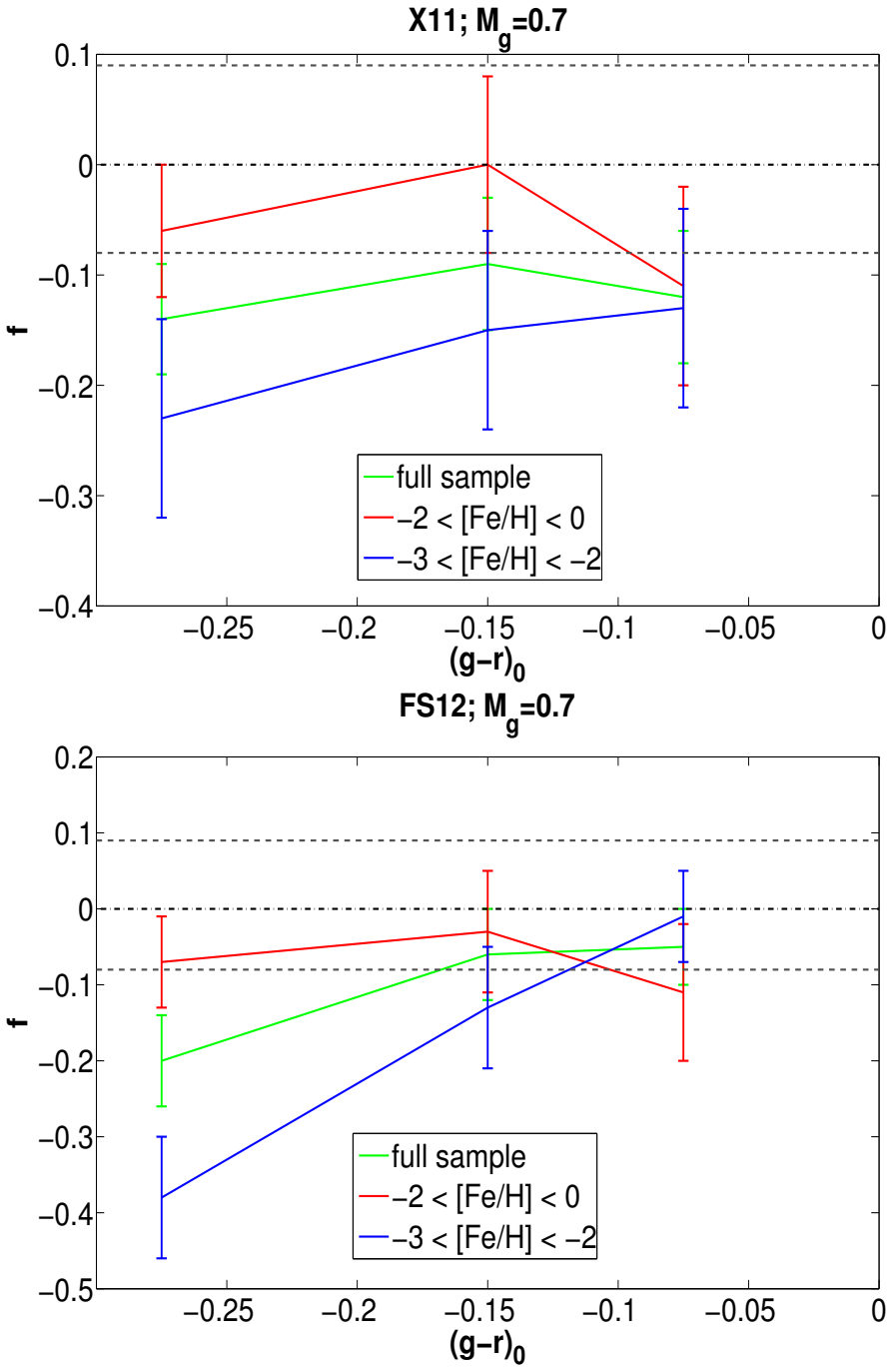


Figure 4.10: Fractional average distance error as a function of colour at different metallicities under the assumption that $M_g = 0.7$ of the spectroscopic sample X11 (upper panel) and the photometric sample FS12 (lower panel). The green lines are the full sample, while the red lines are associated with metal-rich stars and the blue ones with the most metal-poor objects. The actual horizontal coordinate has been offset for each line for the error bars to be distinguishable. Every data point is associated with the centre of a colour bin: its size varies in order for every bin to contain a comparable amount of stars and therefore produce meaningful statistics. Grey dotted lines represent an absolute magnitude uncertainty of $\pm 0.18$ (e.g. Sirko et al. 2004).

$$M_g = g_0 - 5\log_{10}\left(\frac{s}{0.01\,\mathrm{kpc}}\right), \tag{4.5}$$

and $s' = (1+f)s$, where $s$ is the real heliocentric distance, then

$$M'_g = g_0 - 5\log_{10}\left(\frac{s'}{0.01\,\mathrm{kpc}}\right) = M_g - 5\log_{10}(1+f) \tag{4.6}$$

so: $\varepsilon_{M_g} = -5\log_{10}(1+f)$ . An absolute-magnitude uncertainty of $\pm 0.18$ will therefore correspond to $f \in [-0.08, 0.09]$. Fig. 4.10 shows that the fractional distance error trends as a function of colour for different metallicities, significantly exceed this band (grey dotted lines). This implies that the effects of approximating the absolute magnitude of BHB stars with $M_g = 0.7$ go beyond the generally acknowledged uncertainty of this approximation.

Further, the effects of distance misestimates superpose as they ought to: in Fig. 4.10, the green lines that represent the full samples are consistent with being averages of the blue and red lines associated with the metal-poor and metal-rich sub-samples respectively. Also, both samples confirm that the blue end experiences the worst fitting by a constant absolute-magnitude distance scale. The dramatic drop in the metal-rich side of the photometric sample (Fig. 4.10, bottom panel, red line) is however beyond predictable behavior and we therefore flag this end as possibly affected by biases outside our control. Unfortunately, the low statistics associated with this sample do not allow a more detailed investigation of the problem.

Fig. 4.11 shows the dependence of $f$ on $(g-r)_0$ colour on both samples for our calibration and the constant absolute magnitude approximation. The performance of (4.4) on the spectroscopic sample X11 (blue solid line) is almost perfectly satisfactory, while FS12 (blue dotted line) appears slightly problematic at the blue end as expected. We cautiously avoid extensive comment on the photometric sample as regards colour dependence due to the low statistics available for the blue end and the large uncertainties in the stellar parameters at that end.

#### 4.4.5.2 Metallicity dependence

Fig. 4.12 shows the dependence of $f$ on metallicity for both distance calibrations. $|f|$ increases towards underestimates in distance (negative $f$) with decreasing metallicity in both samples. The discrepancy between the two distance calibrations is also affected by metallicity: still in Fig. 4.12, we see that the offset between the blue lines (associated with the colour and metallicity dependent calibration) and the red lines

(constant absolute magnitude) is largest at the metal-poor end for both the FS12 and X11 samples.

The fact that the most metal-poor objects suffer the largest inaccuracy in distance (already visible in Fig. 4.10, upper panel), and experience the largest discrepancy between the two distance calibrations explains why in Fig. 4.11 the offset between the two distance calibrations is significantly larger for the X11 sample than for the FS12 one: the former sample is more metal-poor than the latter and the different mixture of metallicities translates into a larger difference between $M_g = 0.7$ and our formula (4.4).

Finally, we note that when we adopt the colour and metallicity-dependent distance calibration, $f$ lies within the traditional uncertainty band associated with $\varepsilon_{M_g} = \pm 0.18$ (corresponding to $f = -0.08$ and $f = 0.09$ respectively) for both the spectroscopic and photometric sample as regards both colour and metallicity. In particular we tested the hypothesis that $f$ is uncorrelated with metallicity when (4.4) is adopted: a chi-squared test confirms that the discrepancy between the data and a constant trend is not enough to reject the null-hypothesis, at 10-20 % significance level (s.l.) for X11 and 30-50 % s.l. for FS12.

On the other hand the same hypothesis is clearly rejected for the trends associated with $M_g = 0.7$ implying a statistically significant correlation between $f$ and [Fe/H]: at the 1% s.l. for X11 and at the $5 - 10\%$ s.l. for FS12.

### 4.4.6 Improvement w.r.t. previous calibrations

In Fig. 4.13 we show a comparison between the calibrations considered above: the plot is an extension of Fig. 4.6 in the sense that it now shows the trend of $M_g$ with $(g - r)_0$ colour at two different metallicities also for the calibrations of Sirko et al. (2004); Xue et al. (2011) and Deason, Belokurov & Evans (2011b). We do not re-plot the isochrones to avoid crowding in the figure.

We find that our calibration is the only one among those listed above not to suffer the "falling sky" pathology of SBA12 (see §3.3), with the other calibrations showing the most severe effects on the metal-poor sub-sample, as expected from the previous assessment. In Fig. 4.14 we show the average correlation between $W$ and $\sin l \sin b \cos b$ for the metal-poor objects in the X11 sample for our calibration, the one of Sirko et al. (2004); Xue et al. (2011) and for the constant magnitude approximation. The performance of the calibration of Deason, Belokurov & Evans (2011b) is almost identical to the ones of Sirko et al. (2004) and Xue et al. (2011) and we avoid plotting it in order not to crowd the figure. We note that the offset from zero in the mean $W$

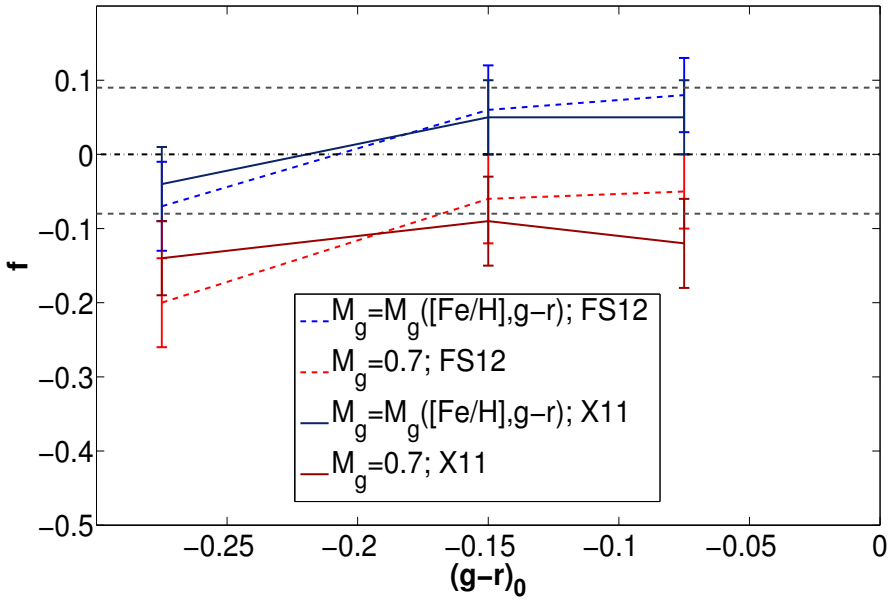


Figure 4.11: Fractional average distance error as a function of colour for BHB stars from SDSS DR9. Blue lines refer to the colour and metallicity dependent distance calibration (4.4), red lines depict the constant $M_g = 0.7$. Solid lines and lighter colours are associated with the spectroscopic sample X11, dotted lines and darker colours with the photometric sample FS12. The discrepancy between the two distance calibrations is larger for X11 due to the sample being more metal-poor than FS12 (see text). Grey dotted lines represent the quoted absolute magnitude uncertainty of $\pm 0.18$.

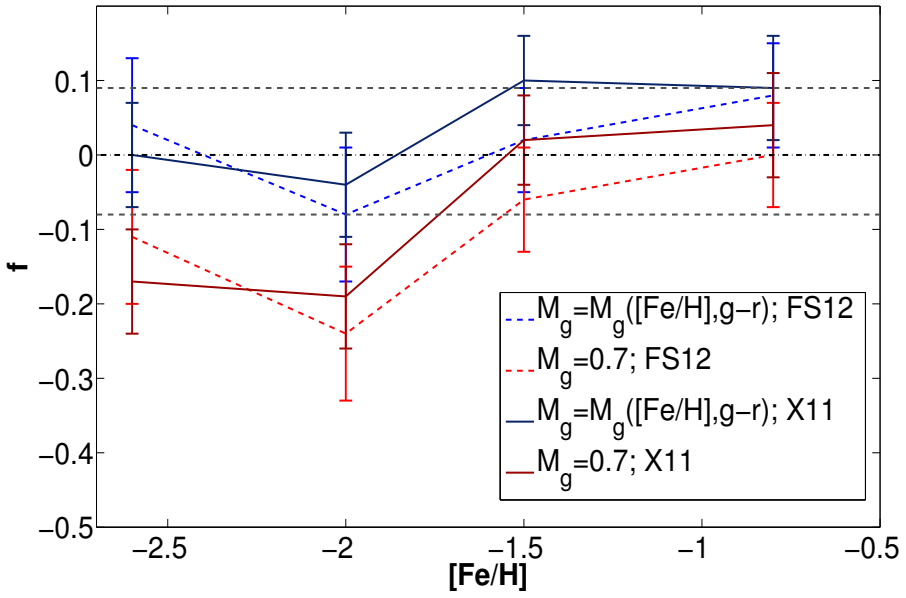


Figure 4.12: Same as Fig. 4.11, but here we plot the fractional average distance error as a function of metallicity for the two samples.

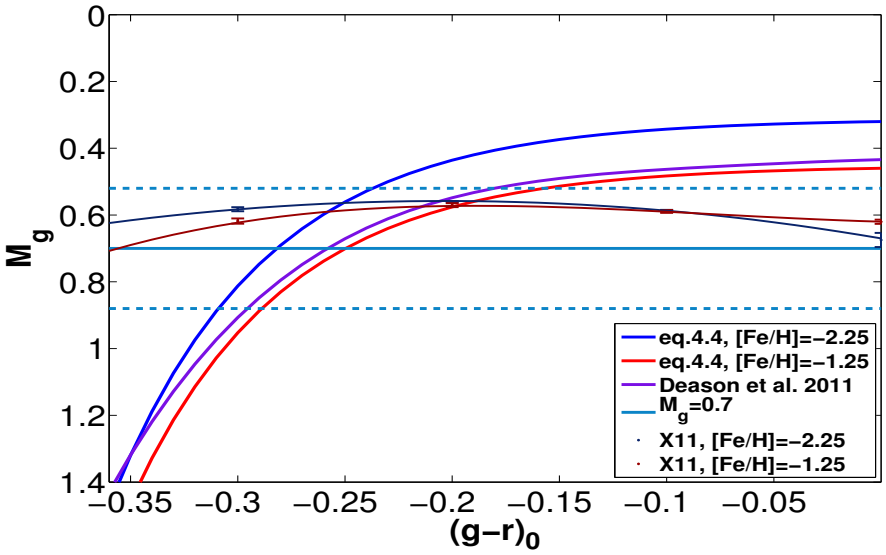


Figure 4.13: Same as for Fig. 4.6 but including the calibrations of Sirko et al. (2004), Xue et al. (2011) and Deason et al. (2011b). The calibration of Deason et al. (2011b) is independent of metallicity and hence shown only once.

disappears accounting for a $v_{\rm los}$ aberration of $\sim 10 - 15\,{\rm km}s^{-1}$. While this is not a proof that SEGUE $v_{\rm los}$ suffer from a systematic offset, the coincidence requires further investigation: we reserve this for future work.

## 4.5 Conclusions

We complied two samples of BHB stars from SDSS DR9: one is based on a combination of photometric and spectroscopic indexes, which are conservatively chosen to minimise the risk of contamination of other spectral types ($\sim 10\%$). A second sample is constructed from the one of Xue et al. (2011): with stars drawn from DR8, they use Balmer-lines shape measurements for a cleaner discrimination of the stars' spectral types. We position cross-match their sample with DR9 and adopt the most recent astrometric and spectroscopic values available, masking out the Sagittarus stream.

We have proposed and validated a very simple analytic formula that properly describes the colour and metallicity dependence of the absolute magnitudes of BHB stars. Our calibration both provides a user-friendly tool to compute the absolute magnitudes of BHB stars from observables in the *ugriz*-system and achieves the best available performance on field BHB halo stars. Given the agreement between our formula and the BASTI isochrones, we conclude that these isochrones successfully

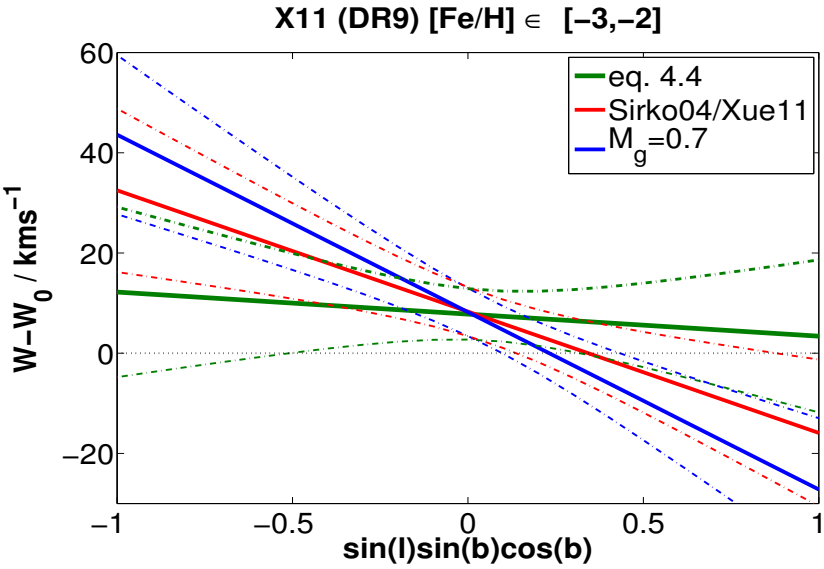


Figure 4.14: Mean correlation between the $W$ component of velocity and the angular term $\sin l \sin b \cos b$ for metal-poor stars ([Fe/H] $< -2$) in the X11 sample. Solid lines are the mean trends associated with the different calibrations: green for our eq. 4.4, red for the calibration of Sirko et al. (2004), Xue et al. (2011) and blue for the constant absolute magnitude approximation. With the same colour code, dashed lines represent the $1\sigma$ contours.

predict *ugriz* colours for the blue part of the HB. The Dartmouth isochrones are fainter than the BASTI models by $\sim 0.1-0.4$ mag and therefore lead to underestimate the distances of field BHB stars.

The quality of proper motions in the SDSS is sufficient to assess the average distance scale. Even though the large distances covered by the two samples imply large random errors in kinematics from proper motions, our considerations are based on averages across the sky. The linear estimator of SBA12 adopted (or see §3.3 in this thesis) is therefore not affected by the inaccurate knowledge of the size of proper-motion errors, while it is mildly vulnerable to proper motion systematics. We corrected for them following Schönrich (2012) and showed that this correction leads to a stretch of the distance scale of about 5%.

Another issue is the unknown level of contamination by blue stragglers and main-sequence stars. This contamination will bias our distance estimates towards signalling an overestimate. Hence our formula providing formally unbiased distance estimates may still be rather on the faint side despite being brighter than the alternative calibrations. Our magnitude approximation may be too flat at the blue end of the horizontal branch, where it declines towards the blue hook. We have actively chosen this mild

deviation in the light of reddening and photometric uncertainties.

The fact that the absolute magnitude of BHB stars depends on both colour and metallicity implies that any calibration that does not account for this dependence is a priori biased. We have shown that the impact of a constant magnitude approximation on distances exceeds the 10% level ($f \in [-0.38, 0.04]$ or equivalently $\Delta M_g \in [-0.09, 1.04]$, see Fig. 4.10, 4.12) and that a calibration which is dependent on colour, but not on metallicity (e.g. Deason, Belokurov & Evans 2011b) introduces a systematics in distance (of order of 0.23 mag or respectively 10%) which correlates with metallicity. This effect is of particular concern in kinematics studies for it can produce a false kinematic gradient across populations with a metallicity-spread.

The available calibrations of BHB magnitudes depend on both colour and metallicity (Sirko et al. 2004; Xue et al. 2008, 2011) require interpolation of theoretical isochrones for each magnitude estimate, and were only validated on a handful clusters, which themselves have a distance uncertainty. We showed that on field BHB stars they significantly under-perform our suggested calibration (on metal-poor stars $f$ can be as high as 0.17) and actually fail the "falling sky" test of SBA12.

The simple calibration formula we presented in equation 4.4 proves to be more accurate than any of the above on field BHB stars, achieves an accuracy that matches the threshold of the noise determined by the sample size, and is the only one that passes the "falling sky" test of SBA12. Larger information content of the data (e.g. better metallicity estimates or a new dimension like Helium content) and better statistics would be required to improve the accuracy of our formula and assess its performance in the blue-hook region.

We remark that the method we use to assess distance calibrations can be applied to any filter set and sample of field stars. The only requirements are that the sample comprises a sufficiently large number of objects with acceptable proper motion quality, is sufficiently extended in Galactic coordinates and especially is not associated with a strong kinematic selection or is dominated by a low number of streams.

# Chapter 5

# Rotational Signature

## 5.1 Introduction

The stellar halo represents an observable tracer of the putative dark matter halo which is thought to govern our cosmological evolution (White & Rees 1978). The present observational evidence favours a hierarchical formation scenario, where the stellar halo was formed through accretion of protogalactic fragments (Searle & Zinn 1978; Bullock & Johnston 2005), especially after the detection of abundant structures with SDSS (Bell et al. 2008; Belokurov et al. 2006). Simulations of stellar halos support the picture of a halo mostly made of accreted structures (e.g. Cooper et al. 2010), but Freeman (2012) remarks that a formation process via dissipation cannot be completely ruled out (Eggen, Lynden-Bell & Sandage 1962; Samland & Gerhard 2003).

A transition from a flattened inner to a more spherical outer halo has sometimes been traced back to a halo formation in different stages and via different physical processes (Hartwick 1987; Preston, Shectman & Beers 1991). Since it has been argued that dynamical friction is more efficient for prograde infall (Quinn & Goodman 1986; Byrd, Saarinen & Valtonen 1986) there is a popular expectation for the outer halo to have a different rotational signature from the inner halo (Murante et al. 2010), and in particular to be counter-rotating (Norris & Ryan 1989).

The kinematics of the stellar halo could serve as a probe of formation history, but it is still highly controversial: on the one hand there are claims for a globally counter-rotating halo (e.g. Majewski 1992). More recently Carollo et al. (2007, 2010) using solar neighbourhood main sequence stars claimed a inner pro-rotating halo opposed to an outer counter-rotating halo, while with a sample of blue horizontal branch field halo stars Deason, Belokurov & Evans (2011a) found a metal-rich pro-rotating halo opposed to a metal-poor counter-rotating halo with such a dichotomy being coherent

at all radii. On the other hand (Chiba & Beers 2000) and Sirko et al. (2004) argue for a non-rotating outer halo, but do not distinguish for different metallicity.

The uncertainty in this debate arises from several causes: the challenge of a clean selection of halo stars, the question if such samples give a fair representation of the halo, as well as uncertainties in determining uniquely their kinematics. Stars in the solar neighbourhood have well-determined kinematics (e.g. GCS, RAVE: Nordström et al. 2004; Steinmetz 2003), however, they contain hardly any halo stars prohibiting an uncontaminated and unbiased halo sample (Kinman 1995). To explore the halo in situ[1], intrinsically brighter objects are needed (Sandage 1970), but for those mainly line-of-sight velocities are available and many distance schemes associated with these stars present metallicity dependent systematics which can produce a false gradient in kinematics across metallicity bins (see §4.4).

In the second part of chapter 3 we have complied a series of kinematic estimators that can extract rotational information from stars with only radial velocity and tentative proper motions. The tight correlation between distance biases and kinematics explored in the first part of the chapter, motivated the construction of a new distance calibration for the two samples of BHB stars selected to trace the stellar halo (see §4.2.1, 4.2.2). Our calibration is inspired by theoretical models of stellar evolution, but is directly validated on the field stars available via the $V$-$W$ correlation method (see §4.3.2): the accuracy it achieves hits comfortably the formal error set by the sample size. In this chapter we will therefore benefit from the efforts above to deliver a clean picture of the rotational profile of the stellar halo and, where possible, set order to the conflicting results present in the literature on this topic.

## 5.2 Four estimates of rotation

We present below four independent estimates of the rotation of the Milky Way stellar halo according to the four methods outlined in 3.4.1 and 3.4.2, traced with the two different samples of BHB stars selected in §4.2.1 and 4.2.2.

The likelihood analysis yields a very weak rotation of both samples, consistent with non-rotation. On FS12 we do not observe any significant trend in the rotational parameter $\eta$ (see §3.4.1) across metallicity: $\eta_{[\mathrm{Fe/H}]<-2} - \eta_{[\mathrm{Fe/H}]>-2} = 0.05 \pm 0.21$, while there is a gradient in the X11 sample: $\eta_{[\mathrm{Fe/H}]<-2} - \eta_{[\mathrm{Fe/H}]>-2} = 0.32 \pm 0.18$. This mismatch disappears on the sub-sample where FS12 and X11 overlap ($\sim 1400$ stars).

[1]By in situ we refer to stars that are actually observed there, rather than stars which are extrapolated to be there via orbit-integration.

| | all | [Fe/H] $\in [-3,-2]$ | [Fe/H] $\in [-2,0]$ |
|---|---|---|---|
| **FS12** | | | |
| $\eta$ | $0.83 \pm 0.08$ | $0.84 \pm 0.17$ | $0.79 \pm 0.12$ |
| $\langle \pi^{-1}(v_{\rm los}) \rangle$ | $17 \pm 9\,{\rm km}s^{-1}$ | $11 \pm 15\,{\rm km}s^{-1}$ | $22 \pm 12\,{\rm km}s^{-1}$ |
| $\langle v_\phi \rangle_{U-{\rm est}}$ | $0 \pm 11\,{\rm km}s^{-1}$ | $2 \pm 17\,{\rm km}s^{-1}$ | $-1 \pm 14\,{\rm km}s^{-1}$ |
| $\langle v_\phi \rangle$ | $6 \pm 6\,{\rm km}s^{-1}$ | $-3 \pm 10\,{\rm km}s^{-1}$ | $13 \pm 8\,{\rm km}s^{-1}$ |
| **X11** | | | |
| $\eta$ | $1.04 \pm 0.08$ | $1.22 \pm 0.15$ | $0.90 \pm 0.10$ |
| $\langle \pi^{-1}(v_{\rm los}) \rangle$ | $-3 \pm 7\,{\rm km}s^{-1}$ | $-18 \pm 11\,{\rm km}s^{-1}$ | $9 \pm 9\,{\rm km}s^{-1}$ |
| $\langle v_\phi \rangle_{U-{\rm est}}$ | $-13 \pm 10\,{\rm km}s^{-1}$ | $-7 \pm 16\,{\rm km}s^{-1}$ | $-18 \pm 13\,{\rm km}s^{-1}$ |
| $\langle v_\phi \rangle$ | $-7 \pm 5\,{\rm km}s^{-1}$ | $-4 \pm 8\,{\rm km}s^{-1}$ | $-9 \pm 7\,{\rm km}s^{-1}$ |

Table 5.1: Estimates of rotation using both $v_{\rm los}$ information only and full 3D motion on two samples drawn from SDSS DR9. $\eta$ is the rotation parameter from model (3.10) and implies prograde rotation for $\eta < 1$ and retrograde motion for $\eta > 1$; $\eta \in [0.79, 1.22]$ corresponds to $\langle v_\phi \rangle \in [-0.08, 0.07] v_{\rm circ}$ at 10 kpc and $\langle v_\phi \rangle \in [-0.05, 0.05] v_{\rm circ}$ at 50 kpc (with $\beta = 0$). $\langle \pi^{-1}(v_{\rm los}) \rangle$ is the rotation estimate from the $v_{\rm los}$-estimator and $\langle v_\phi \rangle_{U-est}$ the one from the $U$-estimator (see §3.4.2.2).

The $v_{\rm los}$ estimator confirms all the above conclusions. Fig. 5.1 presents these results together with the anisotropy estimates from the model in the different cases: FS12 favours an anisotropy consistent with zero, while X11 suggests a very weak radial bias: $\beta = 0.08 \pm 0.05$. It is worth remarking that as pointed out in §3.4.1.1, the model is intrinsically inadequate to estimate the anisotropy due to the $L^{-2\beta}$ term - therefore, the obvious discrepancy to the radial bias in local samples with full 3D information is just a sign for its failure. We note that adopting either a constant distance scale ($M_g = 0.7$, to be consistent with D11) or the colour-metallicity dependent calibration (4.4) does not alter the above conclusions.

When we apply the $U$ and $v_\phi$-3D estimators, they converge to the same conclusion of a rotational signal consistent with zero. On one side, the fact that two estimators agree implies that whatever the nature of the errors affecting the proper-motion, they are either random or not significant enough to bias the average the estimators rely on. Indeed, were there still significant systematics, the estimators would diverge. Further, the $U$-estimator provides also the $U$-component of the Solar velocity. On FS12 we measure $U_0 = 15.7 \pm 7.4\,{\rm km}s^{-1}$ and $U_0 = 7.8 \pm 6.8\,{\rm km}s^{-1}$ on X11: within the error bars both agree with the accepted value of $U_0 = 11.1\,{\rm km}s^{-1}$ (Schönrich, Asplund & Casagrande 2011), again supporting the reliability of the 3D estimators.

On FS12 all estimators agree in measuring no or very weak rotation in both metallicity subsamples (see Table 5.1). However, on the X11 sample the 1D-estimators

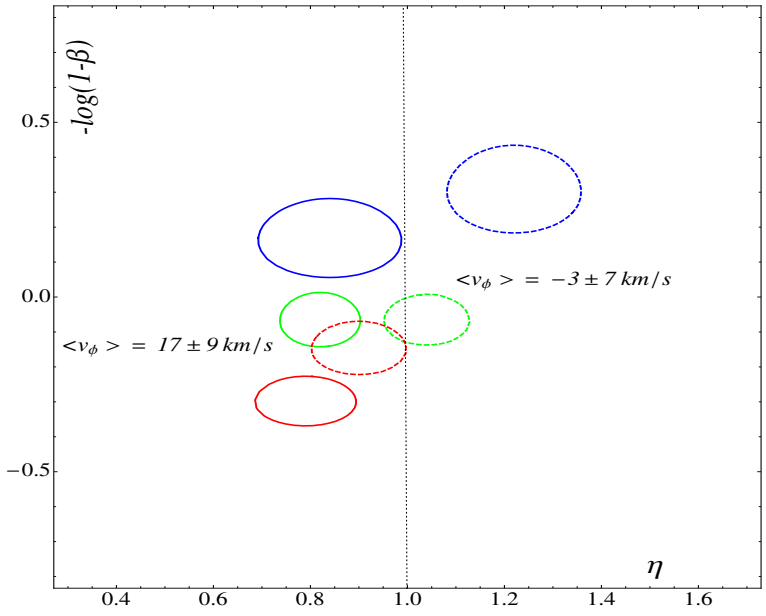


Figure 5.1: Rotation and anisotropy of BHB stars from SDSS DR9 estimated via likelihood of the model (3.10) given the data. Green line is the one sigma contour for the full sample, red line is associated with the metal rich stars and the blue one with the most metal-poor objects. Full lines are associated to the sample selected via (4.1), dashed lines refer to the spectroscopically selected sample from X11. The model-independent $v_{\rm los}$ estimator (see §3.4.2.1) agrees with the findings of the model. Both methods use $v_{\rm los}$ information only.

measure a slight, though significant, discrepancy between the kinematics of metal-poor and metal-rich stars[2].

For the entire halo population, the estimators on the two samples seem to point towards zero-rotational. On FS12, both the 1D and the 3D estimators agree that there is no significant discrepancy in the rotation of metal-poor objects w.r.t. metal-rich ones. On X11 the situation is less clear as the 1D estimators support the existence of a gradient in rotation across metallicity bins, while the 3D estimators disprove it.[3] This issue arises on the sample where we had already flagged 50% of the stars as potentially problematic: in §5.4.1 we critique this discrepancy in detail.

[2]We call metal-poor stars the ones with [Fe/H] $\in [-3,-2]$ and metal-rich the ones with [Fe/H] $\in [-2,0]$. The ratio is 9:20 for both samples.

[3]Note though that an aberration in the l.o.s. velocity would shift the estimate of the 1D estimator leaving nearly unaffected the ones of the 3D estimators (see §5.2.1.1). An aberration of $\pm 5\,(10)\,{\rm km}s^{-1}$ (following what measured by Schönrich (2012) for blue metal poor stars, though for a different portion of the sky) will shift the $\pi^{-1}(v_{\rm los})$-estimate of $-18 \pm 11\,{\rm km}s^{-1}$ by $\pm 9.7\,(19.4)\,{\rm km}s^{-1}$ while moving the 3D-estimator by only 1 (2) ${\rm km}s^{-1}$.

### 5.2.1 Notes on uncertainties

#### 5.2.1.1 Uncertainties from datasets

Proper motions of stars have significant systematic errors, which may affect the $3D$-estimators. The main systematic that affect the proper motions are "frame-dragging" and chromatic aberration (Kaczmarczik et al. 2009). The astrometric frame in Sloan is fixed using galaxies, and can be "dragged" by the inclusion of stars in the "galaxy" sample (see discussion in Schönrich 2012). The correction developed by Schönrich (2012) is designed for usual stellar colours and the chromatic aberration for the bluer BHB stars may be different. However, the Schneider et al. (2010) quasar sample harbours only very few objects with observed $g - r < 0$ (about 10%) providing us with a meagre basis for tests. While for $g - r > 0$ the sample displays almost perfect stability, the blue end deviates at the $1 - 2\sigma$ level towards smaller corrections. Fortunately, the correction on the DR9 sample shifts the average rotational velocity $\langle v_\phi \rangle$ by a moderate $\sim 0.6 - 1.6\,\mathrm{km}s^{-1}$ for the $\langle v_\phi \rangle$-estimator and by $\sim 1.1 - 1.6\,\mathrm{km}s^{-1}$ for the $U$-estimator towards retrograde motion. Further, the proper motion systematics are approximately 0.2 mas/yr. This leads to a maximum 20 $\mathrm{km}s^{-1}$ systematic uncertainty for single stars at $d = 20\,\mathrm{kpc}$, however the vector of the uncertainty is not aligned with the rotational signal in proper motions (see Fig. 5.2) and does not favour any preferred direction of motion, hence it introduces no bias[4]. Anyway, in most cases we are only interested in the rotation difference between two sub-samples, which is nearly unaffected. Therefore, even with the proper-motion uncertainty the mean $v_\phi$ provides valuable information to check the robustness of the model-dependent rotation estimates.

The heuristic exercise of perturbing our $v_{\mathrm{los}}$ by a systematic $\pm 5\,\mathrm{km}s^{-1}$ shows that the 3D-estimators are rather robust against systematics in the l.o.s. velocity:

$$\langle v_\phi \rangle_{v_{\mathrm{los}} \pm 5\,\mathrm{km}s^{-1}} - \langle v_\phi \rangle_{v_{\mathrm{los}}} \sim \pm 1\,\mathrm{km}s^{-1}. \tag{5.1}$$

The 1D $v_{\mathrm{los}}$-estimator is on the contrary affected by a shift larger than the actual systematic introduced:

$$\langle \pi^{-1}(v_{\mathrm{los}}) \rangle_{v_{\mathrm{los}} \pm 5\,\mathrm{km}s^{-1}} - \langle \pi^{-1}(v_{\mathrm{los}}) \rangle_{v_{\mathrm{los}}} \sim \pm 9\,\mathrm{km}s^{-1}. \tag{5.2}$$

[4]Further evidence in this sense comes from comparing the rotational signal between inner and outer halo (see Table 5.3): the signal in the $r < 15\,\mathrm{kpc}$ region remains unperturbed further out ($r > 15\,\mathrm{kpc}$) confirming that the proper- motion errors of stars in this region do not bias the measured rotational signal.

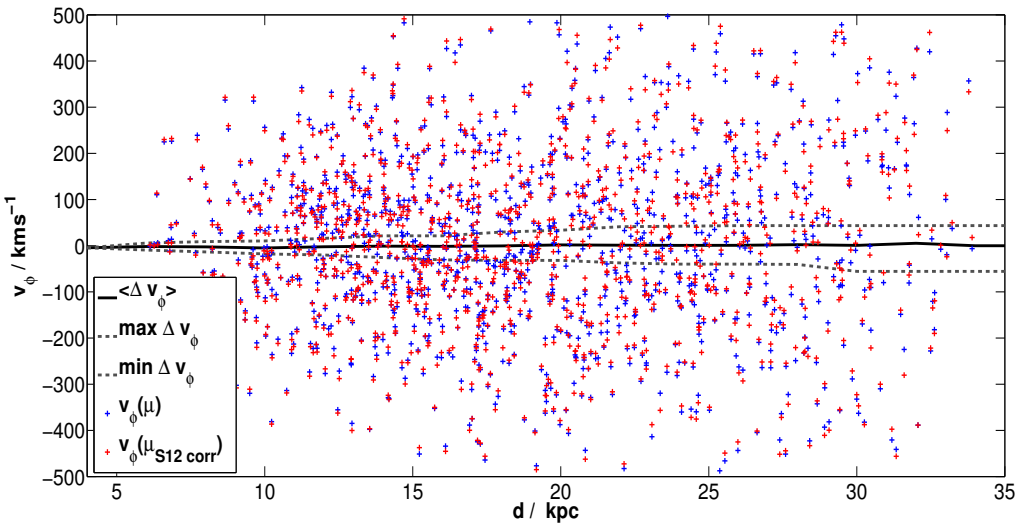


Figure 5.2: Rotational velocity of each star in the sample as function of distance, without (blue) and with (red) the proper-motion correction of Schönrich (2012). The solid black line shows the trend of the average discrepancy between the two values of $v_\phi$ as a function of distance. Dotted line trace the maximum and minimum discrepancy.

This strong deviation might be surprising at the first glance, since the left and the right side of the sky should balance each other. However, the sample is strongly asymmetric, and worse, the geometric factors in the 1D-estimator deprojecting the observed mean motion into a rotational signal amplify any systematic on the line-of-sight velocities. This confirms the necessity of validating rotational estimates via 3D-estimators and supports the conclusions of the latter whenever they disagree with the 1D-estimators.[5]

As a further check we repeated all experiments on earlier data releases DR7 and DR8, before SDSS underwent significant changes especially to the astrometry (Adelman-McCarthy et al. 2008; Abazajian et al. 2009; Aihara et al. 2011a,b; Munn et al. 2004) and all the above statements are confirmed well within one sigma. For those previous data releases the selection via photometry was obtained with a more generous filter in surface gravity: $\log g \in [2.0, 3.5]$, accounting for the higher surface gravity estimates from the SSPP by on average $\sim 0.2$ dex in DR8 w.r.t. DR9 (Ahn et al. 2012). The careful handling of the high surface gravity stars is justified by the risk of contamination from other A-type stars when $\log g > 3.2$ (e.g. blue stragglers,

[5]For example the kinematic trend across different metallicity bins found in the X11 sample and supported by the 1D-estimators only will in fact result in being most probably an artefact (see §5.4.1).

| | $u-g>0.8$ | | $u-g>1.15$ | |
|---|---|---|---|---|
| | $\eta$ | $\beta$ | $\eta$ | $\beta$ |
| all | $0.83 \pm 0.08$ | $-0.04 \pm 0.07$ | $0.83 \pm 0.08$ | $-0.07 \pm 0.08$ |
| rich | $0.75 \pm 0.12$ | $-0.24 \pm 0.11$ | $0.79 \pm 0.12$ | $-0.35 \pm 0.11$ |
| poor | $0.97 \pm 0.16$ | $0.11 \pm 0.11$ | $0.84 \pm 0.17$ | $0.15 \pm 0.11$ |

Table 5.2: Estimates of the rotational signature and anisotropy of the Milky Way stellar halo with BHB stars from SDSS (DR9) selected according to two different cuts in the $(u-g)_0$ colour (the other filters stay as in 4.1): the tighter cut in $(u-g)_0$ aims at excluding a suspected blue hook. $\eta \in [0.75, 0.97]$ corresponds to $\langle v_\phi \rangle \in [0.01, 0.09] v_{\rm circ}$ at 10 kpc and $\langle v_\phi \rangle \in [0.007, 0.056] v_{\rm circ}$ at 50 kpc (with $\beta = 0$, even though adopting the values for $\beta$ reported in the table does not change these limits by more than 7%.).

Sirko et al. 2004). All the same, applying the more generous filter $\log g \in [2.0, 3.5]$ also to the FS12 DR9 sample does not alter our results: the rotation measured by the model shifts by $\sim 5\%$ and the estimators by $1 - 5\,{\rm km} s^{-1}$, but not in a systematic way. With this cut the sample is enlarged by $\sim 200 - 400$ stars for $(u-g)_0 > 1.15$ and $(u-g)_0 > 0.8$ respectively.

#### 5.2.1.2 Systematics in the analysis

We repeat the same analysis applying the more generous cut in $u-g$ colour discussed in §3.1 to include what we identified with the onset of the blue hook in the colour-colour plane. For our sample we find that the effect is negligible: the discrepancy in the rotational signature is less than 7% (respectively $\Delta\eta < 0.13$) and about $4 - 11\%$ (respectively $\Delta\beta \sim 0.04 - 0.11$) in the anisotropy. Results are reported in Table 5.2.

With regard to a possible error in the BHB absolute magnitude (see §3), we re-analyse our sample accounting for shifts in $M_g$ of order of $\pm 0.20$ and find that both the rotational signature and the anisotropy are practically unaffected: we measure no systematic variations in either $\eta$ or $\beta$, with the maximum discrepancy being less than 4% for the rotation parameter and 5% for the anisotropy.

We then contaminate our sample with 10% main-sequence A stars: we measure no appreciable shift in either the rotational signature or the anisotropy previously estimated.

Given the on-going debate on the value of $R_0$ and of the Sun's peculiar velocity (see discussion in Schönrich 2012), we test our results by considering the following variations in the radial position of the Sun: $R_0 = 8.0$, 8.2, 8.5 kpc and find no difference in either the rotation or anisotropy retrieved. The same holds for alterations in the U-component of the Sun's velocity that span values from 10 km/s up to 16 km/s.

What does affect the rotational signature estimated through our analysis is the Sun's azimuthal velocity: raising it from the adopted $v_{c\odot} = 220$ km/s to $v_{c\odot} = 250$ km/s, makes the rotation signature $\eta$ rise by 0.5 towards pro-rotation. However, the significance of this shift is more intuitive looking at the direct estimators: the $+30\,\mathrm{km}s^{-1}$ change in $v_{c\odot}$ translates into a shift in the rotation estimate of $< 10\,\mathrm{km}s^{-1}$ towards pro-rotation.

## 5.3 Comparison with Deason et al. 2011

Deason, Belokurov & Evans (2011a), hereafter D11, presented a study on the kinematics of the Milky Way stellar halo using SDSS BHB stars from DR7 and retrieved a metal-poor counter-rotating halo and a metal-rich pro-rotating halo. They identified BHB halo tracers by cutting in colour, surface gravity, effective temperature and position to remove the disk and the Sagittarius stream as we do in (4.1), except they also include stars with[6]:

$$\begin{cases} g \geq 18 \\ 3.5 \geq \log g < 4.0. \end{cases}$$

### 5.3.1 Discussion of D11's selection criteria

The inclusion by D11 of stars with $\log g \in [3.5, 4.0]$ and $g \geq 18$ exposes them to the risk of contamination by non-BHB stars (Sirko et al. 2004; Lee et al. 2008a). Therefore, the rotational signature retrieved with their sample is more prone to be biased: according to the literature value they buy their larger sample size with a contamination by miss-selected stars of $\sim 25\%$ (compared to 10% for our sample, see §4.2.1) and the inclusion of faint stars makes their sample span distances from the GC up to 90 kpc, while we cover a sphere of radius $30 - 40$ kpc. Therefore the geometrical biases outlined in §3.4.3.1 will be more significant for the sample of D11 than for ours.

We confirm that if we drop our tighter cuts in surface gravity and apparent magnitude, we retrieve the results of D11[7]. Also, we try the heuristic exercise of perturbing

[6]Note that given the higher average surface gravity value in SDSS data releases prior to DR9 by $\sim 0.2$ dex (Ahn et al. 2012), for DR7 the surface gravity cut in (4.1), reads: $2 < \log g < 3.5$ (see §5.2.1.1 for a discussion).

[7]With the sample drawn from SDSS DR7, filtered as prescribed by D11 and assuming $M_g = 0.7$, we find the same two populations (a metal-poor one with $-3 <$ [Fe/H] $< -2$ and a metal-rich population identified by $-2 <$ [Fe/H] $< 0$) and estimate $\beta$ and $\eta$ for them separately, according to the likelihood (3.12). We confirm the trend presented by D11 of a pro-rotating metal-rich population versus a counter-rotating metal-poor one and match their numerical estimates of $\eta$ and $\beta$ with the

the distances in the sample of D11 by 20% and verify that the inclusion of faint stars makes the rotational signature highly susceptible to distance errors.

We assume a misclassification of $\Delta M_g = \pm 0.40$, corresponding to a distance underestimate of $\pm 20\%$[8]. Fig. 5.3 shows that for distance overestimates (dotted lines) the rotational signatures are noisier and the discrepancy between the metal-poor (blue lines) and the metal-rich (red lines) sample is exacerbated, while for distance underestimates (dashed lines) they shift towards non-rotation and the signal reduces. On the other hand, when we challenge our sample with a perturbation in the adopted absolute magnitude of similar intensity, we measure no scatter greater than one sigma in the rotational signatures previously estimated. It seems helpful to remark that distance errors do not change the sign of rotation, but rather enhance (for distance overestimates) or shrink (for distance underestimates) the measured rotational signature: this is a pure geometrical consequence for measuring rotation via the component of velocity that at the limit of infinite distance is perpendicular to the direction of rotation. Therefore, if there is any rotational signal a distance bias will inflate it or reduce it, but if the system is non-rotating, errors in distance will have no effect on the estimated streaming motion. Distant stars are particularly sensitive to distance systematics for the effect of a bias scales with distance (see Fig. 3.5). This explains why our sample is so little affected by perturbations in the distance scale, while the signature found by D11 is not robust in these respects.

In conclusion, the rotation measured by D11 is driven by stars with $\log g \in [3.5, 4.0]$ and $g \geq 18$. When we remove these stars, with the same analysis we measure no rotation. We note that this exclusion does not bias the properties of our sample: for example the first and fourth panels of Fig. 5.4 show that the distributions of $v_{\rm los}$ and of $T_{\rm eff}$ in FS12 are unaffected by the tighter filters than D11. The $g < 18$ cut very reasonably shifts the peak of our distance distribution towards a shorter mean, but has no effects on the other properties. Further, this does not affect our conclusions since the signal claimed by D11 is constant with radius. Further, the right-top panel of Fig. 5.4 tells us that the removal of $\log g \geq 3.5$-stars diminishes mainly objects from the metal-rich sub-sample and far less from the metal-poor sub-sample. Since the controversial signature resides on the metal-poor side of the sample, the signal

exception for the anisotropy parameter for metal-rich stars. D11 find $(\eta, \beta)_{\rm rich} = (0.8 \pm 0.1, 0.1 \pm 0.1)$ and $(\eta, \beta)_{\rm poor} = (1.4 \pm 0.1, 0.2 \pm 0.1)$ while we find $(\eta, \beta)_{\rm rich} = (0.7 \pm 0.1, 0.5 \pm 0.2)$ and $(\eta, \beta)_{\rm poor} = (1.5 \pm 0.1, 0.3 \pm 0.1)$. We suspect this difference traces back to their undescribed handling of the singularity at $L = 0$ by D11 and to the 15% difference in sample size.

[8]We remark that regardless of the theoretical nature of this exercise, these values for $\Delta M_g$ are not at all unrealistic (see Chapter 4).

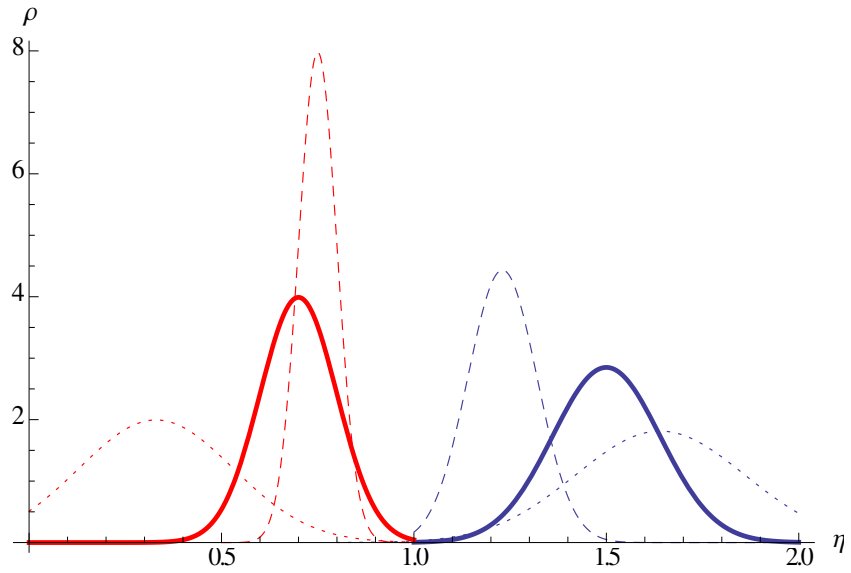


Figure 5.3: Probability density of $\eta$ marginalized over $\beta$ for BHB stars from DR7 (see 4.1): blue lines for metal-poor stars and red lines for metal-rich ones. Full lines represent D11's results, who assume that $M_g = 0.7$ for each star in the sample. Dashed lines show how the rotational signatures change if the distances have been underestimated by 20%, dotted lines represent the signatures associated with a distance overestimates of the same order.

should be enhanced by the altered proportions, but this does not happen. It is worth remarking though, that these high surface gravity stars do bias the overall metallicity distribution in D11: when we compare the metallicity distribution of the FS12 and D11 sample with the control sample of Xue et al. (2011) (to which no such cut in $\log g$ was applied), we observe that the mean in D11 sample is shifted by $\sim 0.2$ dex w.r.t. X11, while the mean metallicity in FS12 and X11 samples coincide (see Fig. 5.5).

Besides the concerns associated with the classification of these stars, if the signature measured with D11's more generous filter were a global property of the sample considered, removing *random* elements from it would widen the spread around the mean, but keep the latter the same[9]. On the contrary the comparison shows a shift in the means and a decrease in the formal errors, confirming the suspicion that the objects with $\log g \in [3.5, 4.0]$ and $g \geq 18$ are non-BHB stars.

[9]Random is here understood as statistically non significant and the latter consideration is a consequence of the assumption that there are no spectral contamination in the population.

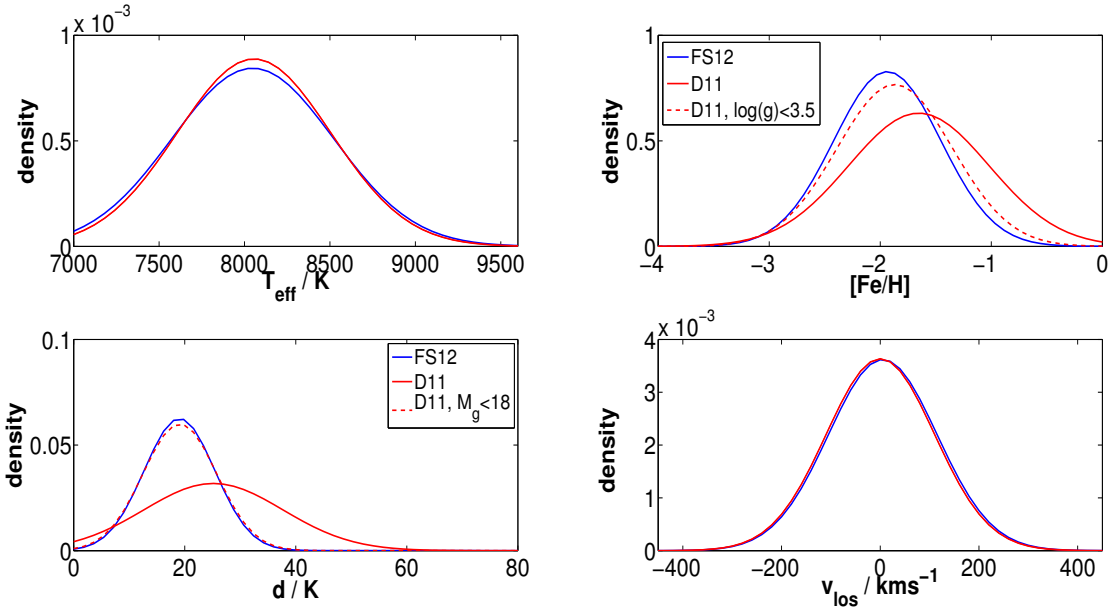


Figure 5.4: From top left to bottom right, density distributions of effective temperature, metallicity, heliocentric distance and l.o.s. velocity for the FS12 sample (blue lines) and the Deason, Belokurov & Evans (2011a) sample (red lines). The dotted red lines in the second and third plot refer to the Deason, Belokurov & Evans (2011a) sample with only one additional condition: $\log(g) < 3.5$ and $M_g < 18$ respectively.

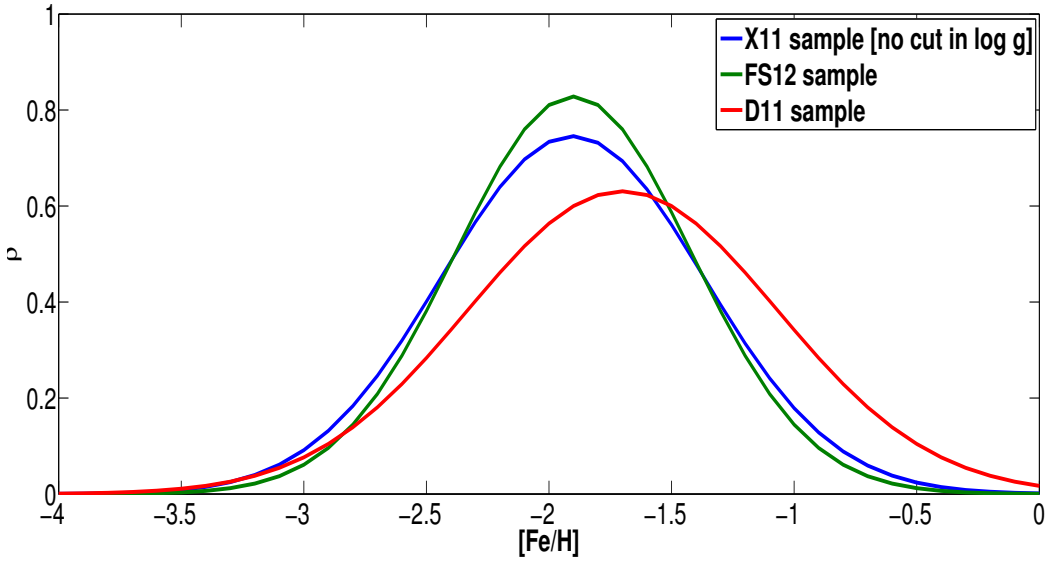


Figure 5.5: Metallicity distribution for the Deason, Belokurov & Evans (2011a), Xue et al. (2011) and FS12 samples. The more geneours filters of Deason, Belokurov & Evans (2011a) bias the mean metallicity of the tracer sample towards higher metallicities.

### 5.3.2 Consistency between model and data

To assess whether there is any systematic in the analysis of D11 due to the more generous filters in apparent magnitude and surface gravity, we run the consistency check introduced on their sample (see §3.4.3.1). In the subsequent subsection §5.3.2.2 we favour a disc contamination as a possible origin of the detected inconsistency.

#### 5.3.2.1 Spatial dependence of the rotational signal

We bin $(l, b)$-space to have cells with useful statistics according to the $(l, b)$ distribution of BHB stars within SDSS. A fine grid causes the noise on the estimate of $\eta$ to increase and progressively makes whatever discrepancy across bins statistically less significant. Oppositely, if we increase the bin size (and hence decrease the grid refinement) we tend to wash out any local discrepancy. As a general procedure one can vary the grid size and look for discontinuites at the different scales[10]; the test is only qualitative and aims only to inform whether the data-model pair is consistent, therefore once any discrepancy is found the exercise is over. In our study the first iteration of this search was a regular $4 \times 4$ grid, which made the formal error in each bin of the same order and already provided a definitive output for the test, thus it was also our final choice.

There is of course a geometrical effect one needs to account for: the sphericity of the reference frame causes no rotation at the poles ($b = \pm 90$) and maximum rotation on the galactic plane. However given the significantly big bins chosen, this produces a smooth fall of the rotational signature measured from the plane up to the poles, rather than jumps. From one bin to the next the mean rotation velocity is therefore expected to be consistent within the error bars. The only place where we can have a real physical discontinuity in the rotational signature is the region which includes the GC.

Fig. 5.6 shows the bins in $(l, b)$-space with an arrow in each bin pointing in the direction of the recovered rotation for the metal-poor component. The arrow lengths are associated to the strength of the rotational signature, and the numerical values of $\eta$ are reported above the arrow. We see that the best values for the rotational parameter are consistent, within the errors, with a non rotating stellar halo ($\eta = 1$). Fig. 5.7 shows the same analysis for the metal-rich population: here the situation

[10]Given the low statistics available, it would be unseemly to talk about an optimal grid size though, also because the optimality with respect to what is debatable.

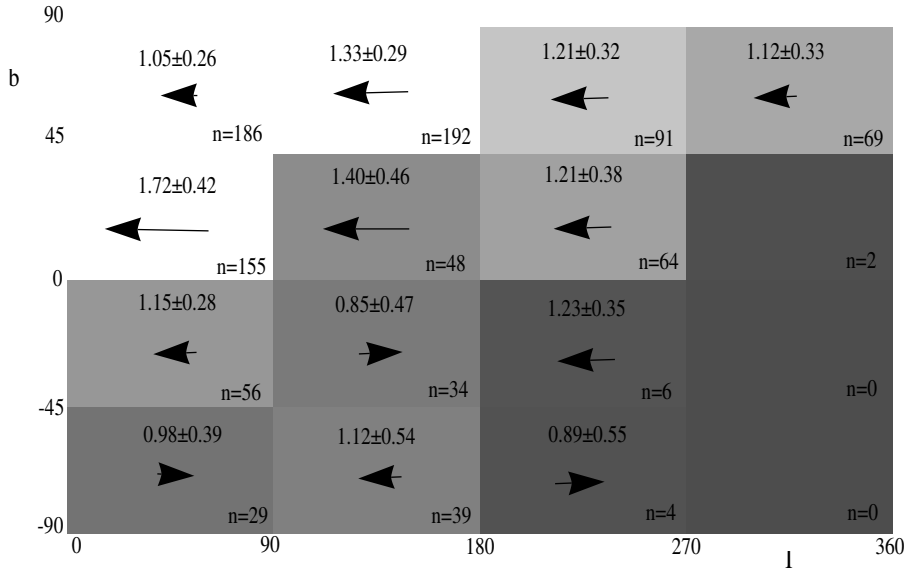


Figure 5.6: Analysis of the rotation parameter $\eta$ on metal-poor BHB stars from SDSS DR7 as selected in D11 and performed locally in $(l, b)$-space. Arrows pointing right indicate pro-rotation and vice versa for arrows pointing left. Their magnitude is a visual indication of the rotation strength while the actual value is reported above them together with the associated formal error. Bins are coloured according to the density of stars (reported in the bottom-right corner of every bin). The colour scale is normalized per each population.

becomes pathological as there are jumps from pro-rotating to counter-rotating motions. Compare for example the bin centred on $(135, -68)$ with the one on $(135, -23)$ or the bin centred on $(225, 23)$ with the one on $(225, 68)$.

From this, we conclude that the model (3.10) is unable to fit the sample of D11 and the parameters estimated using this model cannot be used to achieve any further understanding of the system. Indeed, if the model is a bad representation of the data, the parameters estimated from it, will be driven by the noise that arises from the discrepancy between the model and the true distribution, rather than by the real signal in the data.

A second consistency check is presented in Appendix C, where we show that for a distribution function like (3.10) and for fixed $\beta$, the best-fitting value for the entire population should be a linear combination of the best-fitting values of any sub-samples one breaks it into, where the coefficients are the fractional sizes of each sub-sample with respect to the total size. Our results satisfy this constraint, but the ones reported by D11 do not. From their Fig 4. $\eta_{\rm rich} = 0.8$, $\eta_{\rm poor} = 1.4$, $\eta_{\rm total} = 1.1$. For fixed $\beta = 0.1$, having 32% metal poor stars and 60% metal rich stars, one should

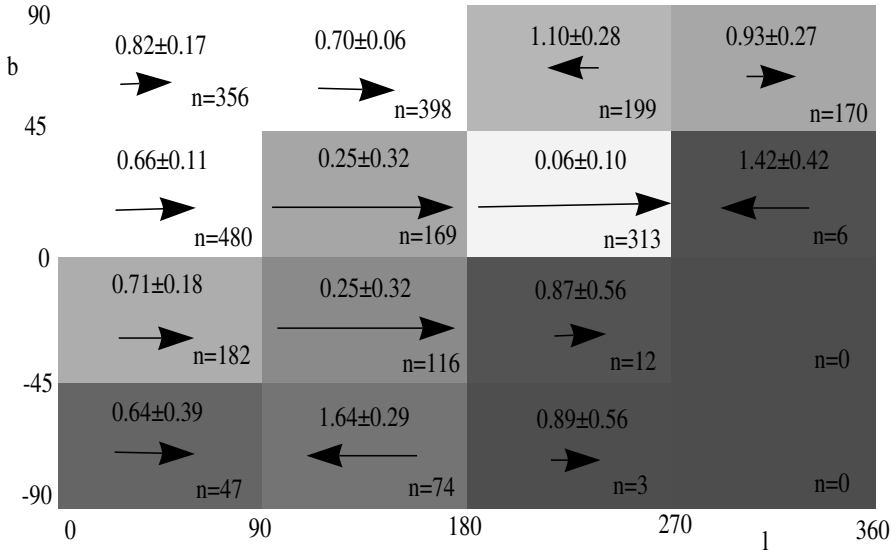


Figure 5.7: Same as in Fig. (5.6) for metal-rich stars. Note the jump in signature between the bin centred in $(135, -68)$ and the one in $(135, -23)$. Or between the bin centred in $(225, 23)$ and the one in $(225, 68)$ bin.

have $\eta_{\text{total}} = 0.9$ (if one considers the 3549 stars) or 1.0 (if one previously excludes from the total sample stars with "no assigned metallicity", i.e. $\sim 300$ stars).

The U-estimator (see 3.4.1) supports the conclusions of the two sanity checks just introduced: on the sample of D11 we measure an overall very weak counter-rotation actually consistent with no-rotation at all: $\langle v_\phi \rangle = -4 \pm 10$ km/s. However, with this sample, the radial velocity of the Sun is offset by more than 1 sigma: $U_\odot = 1.4 \pm 6.7$ km/s.[11]

#### 5.3.2.2 Likely disc contamination in D11's sample

The prograde rotation measured by D11 on their metal-rich sample is linked to the spatial inconsistency of their data-model: Fig. 5.7 shows that the systematic prograde rotation of metal rich objects occurs at low latitudes and fades away at higher latitudes. A priori this would be consistent with either a disc contamination, the lower halo being set into stronger rotation by friction with the disc or the debris of an accretion event.

[11]Note that unlike the double power-law model, the direct estimators rely on averages over the whole sky and we will therefore not use them to estimate rotation on a spatial subdivision of the data. Should they show any discontinuity of rotation across the different bins, one could not distinguish whether the discrepancy pattern is a consequence of the failure of the assumptions the methods are based on or it is a genuine discontinuity in the rotation profile of the smooth component.

In an attempt to break the degeneracy, we vary the lower altitude cut. When we vary the z-cut from 4 to 6 kpc (cutting about $\sim 10\%$ of the sample) in the D11 sample, the $v_{\rm los}$-estimator shows a shift in the mean rotational signal of metal-rich stars from $34.1 \pm 7.4\,{\rm km}s^{-1}$ to $23.5 \pm 8.2\,{\rm km}s^{-1}$. At a cut of 14 kpc the $v_{\rm los}$-estimator has lost its significance but confirms the systematics of the shift in $\langle v_\phi \rangle$ by dropping to $12.5 \pm 16.6\,{\rm km}s^{-1}$. The prograde signal is therefore consistent with coming from stars close to the Galactic disc. How reliable is it? To answer this, we look again at the 3D rotation estimators and compare to the other, cleaner datasets.

In the cleaner FS12 sample we find a slower rotation of $22.2 \pm 12.0\,{\rm km}s^{-1}$ for $|z| > 4\,{\rm kpc}$ and a statistically insignificant signal of $11.6 \pm 12.6\,{\rm km}s^{-1}$ for $|z| > 6\,{\rm kpc}$, while the X11 sample, which should have the lowest contamination with blue stragglers/main-sequence stars, has no prograde rotation at all (see Table 5.1). Hence, with increasing data quality the signal diminishes.

On the other side the 3D estimators do not display any rotation at all. In the most strongly rotating case of D11 we find:

$$\begin{aligned} \langle \pi^{-1}(v_{\rm los}) \rangle &= (34 \pm 7)\,{\rm km}s^{-1}, \\ \text{but } \langle v_\phi \rangle_{U-est} &= (-2 \pm 10)\,{\rm km}s^{-1}, \quad \langle v_\phi \rangle = (3 \pm 6)\,{\rm km}s^{-1} \end{aligned}$$

where the values are taken for $|z| > 4\,{\rm kpc}$. This inconsistency could be the consequence of part of a stream in the sample, but even then its rotation wouldn't be necessarily prograde. Further, the 3D estimator $\langle v_\phi \rangle$ shows a hint of rotation in the near-disc region and interestingly gets into better agreement with $\pi^{-1}(v_{\rm los})$ when we shorten all distances, while the $U$-estimator stays consistently at 0. All this points to the simpler and therefore favourable hypothesis of a contamination with fainter disc objects.

It is worth remarking that for the sample of D11, disc contamination is likely to exceed the geometric limits of the disc itself: the a priori comparison with the sample of Xue et al. (2011) flagged half of the stars as probable non-BHB. If a blue straggler/main sequence star is mistaken for a BHB member, its distance is significantly overestimated, so disc contamination due to these objects can affect the BHB sample well above the actual disc limits.

In summary even on the D11 sample the 3D estimators refute any reliable evidence for a clean prograde rotation, while the difference between the estimators points to a contamination by disc objects. This does not mean that no prograde tendency in the near-disc region of the Galactic halo exists, but it is not detectable at the accuracy of the Deason, Belokurov & Evans (2011a) sample or other samples.

### 5.3.3 Rotation gradient

#### 5.3.3.1 In metallicity

The gradient in rotation that D11 measure between samples of different metallicity could at first sight be interpreted as support of the claim of a double-component stellar halo, even though D11 w.r.t. their result advocate the need of redefining the Sun's azimuthal velocity to $250\,\mathrm{km}s^{-1}$ and impute the rotation gradient to contamination from an accretion event.[12] Besides the fact that with stringent quality cuts we do not measure any rotation discrepancy between metal-poor and metal-rich stars (see §5.2), we investigate if there is any further ground in the sample of D11 that could support the thesis of a double-component stellar halo. We have searched systematically for evidence of bimodality in other distributions and can report that no bimodality is evident in any of the following planes: [Fe/H] versus any of $u-g$, $g-r$, $\log(g)$, $g$-mag or $v_\phi$; $v_\phi$ versus any of $g$-mag, $u-g$ or $r$; or $r$ versus $u-g$.

#### 5.3.3.2 In radius

On the other hand, Carollo et al. (2007, 2010) and Kinman et al. (2007, 2012) suggest the existence of a double halo in terms of real space position: an inner halo dominant at 5-10 kpc and an outer (counter-rotating) halo dominant beyond 20 kpc. We investigated this possibility both with our samples and with the selection of D11, but find no evidence for such bimodality in either the radial density profile or in the radial distribution of the rotational signature (e.g. outer halo counter-rotating and inner halo pro-rotating). Fig. 5.8 shows the mean-streaming motion as a function of galactocentric radius: we split the sample into 5 kpc-wide bins, values being shown every 2.5 kpc so that every second value is independent. The top panel refers to the FS12 sample and the bottom one to the low $c_\gamma$ stars in X11 (see discussion in the following section): to determine the significance of the deviations from a zero constant rotation we run a *p-test* on both trends. Assuming a certain null-hypothesis (model) is true, the p-value is the probability to measure a deviation from it as large as the one observed. The levels of significance to reject a null-hypothesis are most commonly taken to be $1-5\%$. The p-value for the trends shown in Fig. 5.8 confirms

[12] We remark that drawing from the SDSS DR7 catalogue and applying the same cuts as in D11, lead to obtain a slightly different sample size with respect to D11. With the colour, gravity, temperature and height filter we have 5523 stars (D11, 5525) and when masking out Sgr stream and any close star ($r > 10$ kpc) the sample shrinks to 3500 stars (D11, 3549). Of these, 2525 fall in the metal-rich population (D11, 2125) and 975 in the metal-poor population (D11, 1135) even though we are using the same metallicity indicator (Wilhelm, Beers & Gray 1999).

that the deviation from a constant zero rotation is statistically insignificant for all estimators on all samples:[13]

| $p-$value | $v_{los}-$est | $U-$est | $v_\phi-$est |
|---|---|---|---|
| FS12 | 0.53 | 0.75 | 0.50 |
| X11 | 0.80 | 0.82 | 0.26 |

In Table 5.3 we report the actual estimates of the mean streaming motion for two spatial components, one of stars closer than 15 kpc in galactocentric radius and one of stars further than that. All estimators show[14] a halo weakly rotating if at all and that there is no appreciable signal between the components either spatially or chemically selected. Note that the prograde rotation found by Kafle et al. (2012) at small radii in their metal-rich sample of BHB stars drawn from the catalogue of Xue et al. (2011), is based on $v_{los}$ information only: we find no grounds to support it looking at the more robust 3D estimators (see §5.2.1.1). Even though this could be the kinematic trace of a debris from an accreted galactic fragment, it is helpful to note that a plausible 10-15 $kms^{-1}$ aberration in $v_{los}$ (Schönrich 2012) would cancel any hint of rotation also in the $v_{los}$ estimators, leaving basically unaltered the 3D estimators. Further, the discrepancy in the rotational signal that Kafle et al. (2012) see between metal-rich and metal-poor stars, is enhanced both by an inappropriate choice of distance scale, and by the kinematic bias afflicting high $c_\gamma$ stars in the sample of Xue et al. (2011) (see below).

## 5.4 Signature in the Xue et al. sample: critique

As in the case of Deason, Belokurov & Evans (2011a), the data-model pair for the X11 sample fails the consistency check introduced in §3.4.3.1, so we need to investigate the sample in detail.

In Fig. 5.9 we see that X11's sample strongly enhances the high temperature-high $b_\gamma$ region previously unexplored by the photometrically selected sample. Also, the relation between $T_{\rm eff}$ and $b_\gamma$ clearly changes in this region w.r.t. lower temperatures. Separating the two trends by eye (green line in Fig. 5.9) the rotation discrepancy between metal-rich and metal-poor stars below the boundary disappears (metal-rich : $\eta = 0.76 \pm 0.15$; metal-poor: $\eta = 0.94 \pm 0.15$). On the other hand, the stars above the

[13] In our analysis we use 1 minus the cumulative distribution of the $\chi^2$-distribution to estimate the p-value. Note that because our bins overlap by half their size, we effectively have half the degrees of freedom we would have if the bins were all independent.

[14] The discrepancy between the numerical estimates reflect the different sensitivities of the estimators to the sample geometry.

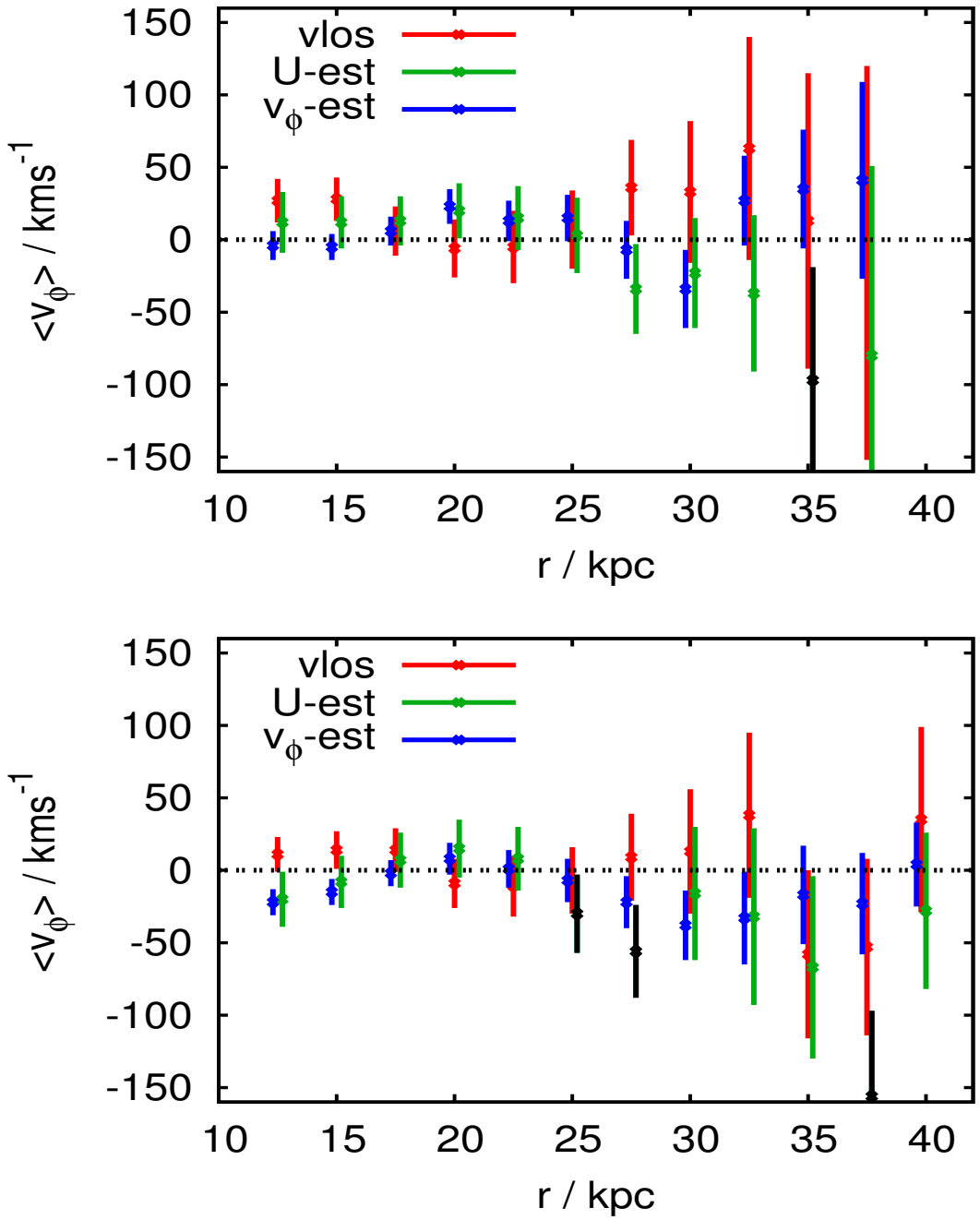


Figure 5.8: Mean streaming motion as function of galactocentric radius for the FS12 sample (top panel) and stars in the X11 sample with $c_\gamma < 0.95$ (bottom panel). In one bin out of eleven for FS12 and three out of twelve for X11, the $U$-est returned a solar value in disagreement with the literature: the level of systematics in those bins is therefore relevant and the rotation estimate cannot be trusted there (dark green points).

| $v_\phi(\,\mathrm{km} s^{-1})$ | $v_\phi$-estim | U-estim | $v_{\rm los}$-estim |
|---|---|---|---|
| $r < 15\,\mathrm{kpc}$ | | | |
| | $-3 \pm 8$ | $7 \pm 19$ | $22 \pm 13$ |
| $-2 < [\mathrm{Fe/H}] < 0$ | $10 \pm 11$ | $9 \pm 24$ | $31 \pm 16$ |
| $-3 < [\mathrm{Fe/H}] \leq -2$ | $-21 \pm 13$ | $6 \pm 32$ | $7 \pm 21$ |
| $r \geq 15\,\mathrm{kpc}$ | | | |
| | $7 \pm 7$ | $-4 \pm 11$ | $4 \pm 11$ |
| $-2 < [\mathrm{Fe/H}] < 0$ | $12 \pm 10$ | $-6 \pm 15$ | $14 \pm 15$ |
| $-3 < [\mathrm{Fe/H}] \leq -2$ | $0 \pm 11$ | $-1 \pm 18$ | $-8 \pm 17$ |

Table 5.3: We test the hypothesis of a double component stellar halo. With the BHB sample from SDSS DR9, the model-independent estimators of sections §3.4.2.1, 3.4.2.2 measure no significant signal both between the kinematics of the closer stars as opposed to the furthest ones and when we discriminate for their metallicity.

line show a strong counter-rotation as well as a signal of $\Delta\eta \sim 0.66 \pm 0.19$ (metal-rich: $\eta = 1.11 \pm 0.06$; metal-poor: $\eta = 1.77 \pm 0.18$).

The few stars above the line in the photometrically selected sample show the same trend. Their strong under-representation (20% versus 50% in X11) explains the discrepancy towards X11 (left panel in Fig. 5.9).

### 5.4.1 Kinematics of high $c_\gamma$ stars

This dichotomy can be linked to the steepness parameter $c_\gamma$, which is a line indicator characteristic of luminosity type. While it does not directly select on kinematics we see a strong systematic difference within the metal-poor stars between high and low $c_\gamma$. In Fig. 5.10 we plot the de-projected l.o.s. velocities as a function of the line indicator $c_\gamma$ (red squares) and show the least square fit to $\pi^{-1}(v_{\rm los})(c_\gamma)$ in the two regions (green lines). We do not apply any kinematic cut in fitting the $\pi^{-1}(v_{\rm los})$-trends, but for the plot we limit $\pi^{-1}(v_{\rm los}) \in [-300, 300]\,\mathrm{km} s^{-1}$ to make the fit difference in $c_\gamma$ better visible. Metal-poor stars with $c_\gamma \geq 0.95$ display a significant retrograde signal that is absent in the low $c_\gamma$ stars. We note that these stars are not more distant than stars with $c_\gamma < 0.95$; indeed the density distributions of the two sub-samples show no appreciable differences. The metal-rich stars do not seem to suffer from this systematic.

We bisect our sample in both in $c_\gamma$ and [Fe/H] and for each of the four sub-samples determine the cumulative distributions in $v_{\rm los}$ and $\pi^{-1}(v_{\rm los})$. We then determine the Kolmogorov-Smirnov statistics to ask whether the former distributions can be drawn

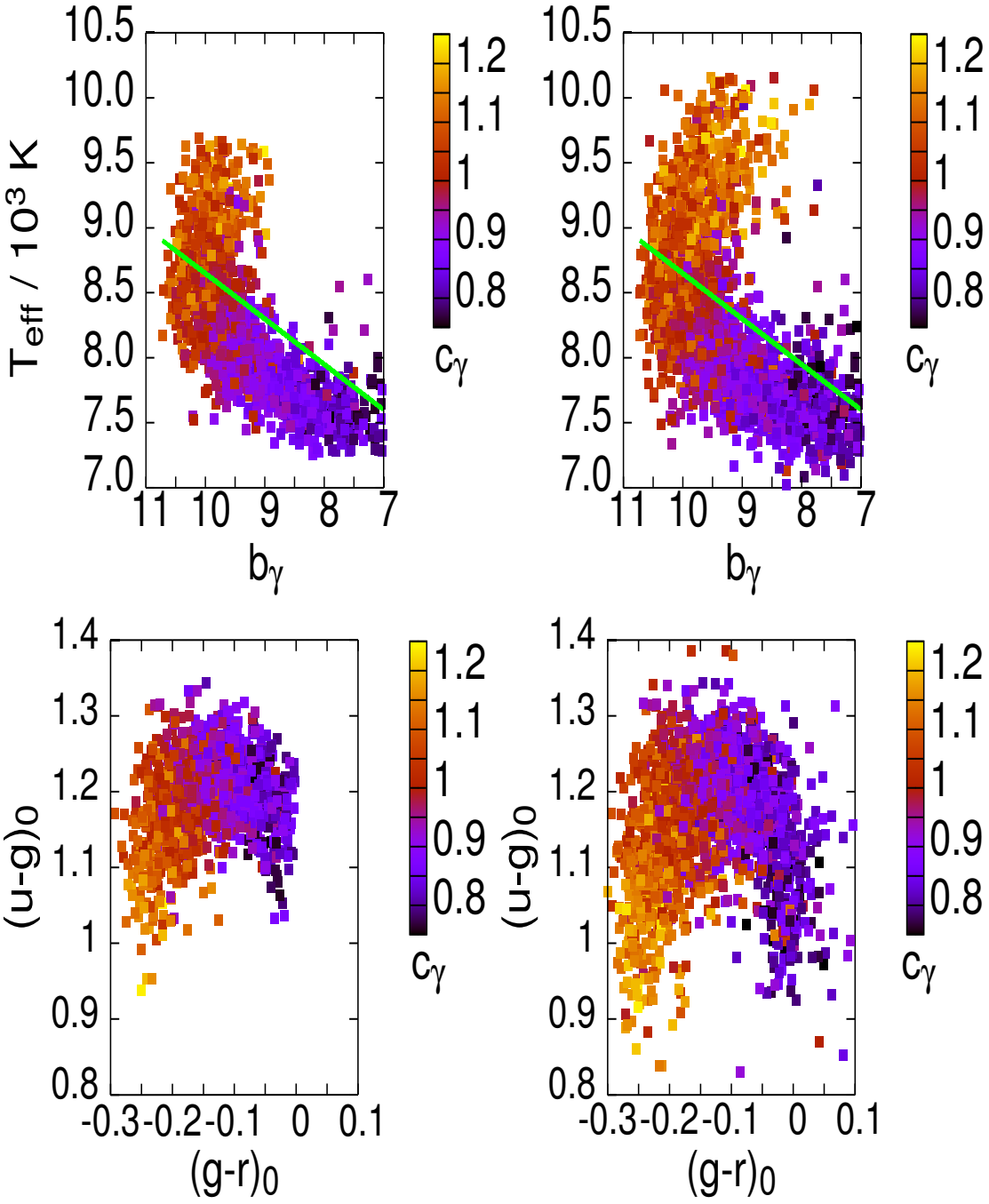


Figure 5.9: BHB stars that appear in both X11's sample and in the sample photometrically selected via (4.1) (left panels) and the full X11's sample (right panels). In the top panels we plot $b_\gamma$ versus $T_{\rm eff}$: the green line indicates what by eye appears to be a change in trend. X11's sample clearly enhances the region above the line: in the colour-colour plane (bottom panels) that is associable with the blue hook. The colour code reflects the value of the $c_\gamma$ Sersic parameter.

| $v_\phi(\,\mathrm{km s}^{-1})$ | $\eta$ | $v_\phi$-est | U-est | $v_{\rm los}$-est |
|---|---|---|---|---|
| $c_\gamma < 0.95$ | | | | |
| | $0.90 \pm 0.17$ | $-13 \pm 8$ | $-20 \pm 16$ | $10 \pm 11$ |
| $-2 <$ [Fe/H] $< 0$ | $0.93 \pm 0.16$ | $-13 \pm 11$ | $-27 \pm 21$ | $11 \pm 14$ |
| $-3 <$ [Fe/H] $\leq -2$ | $0.87 \pm 0.19$ | $-12 \pm 13$ | $-11 \pm 24$ | $8 \pm 17$ |
| $c_\gamma \geq 0.95$ | | | | |
| | $1.11 \pm 0.13$ | $-2 \pm 7$ | $-7 \pm 13$ | $-10 \pm 9$ |
| $-2 <$ [Fe/H] $< 0$ | $0.91 \pm 0.10$ | $-5 \pm 8$ | $-10 \pm 16$ | $8 \pm 11$ |
| $-3 <$ [Fe/H] $\leq -2$ | $1.4 \pm 0.19$ | $2 \pm 10$ | $-2 \pm 21$ | $-36 \pm 14$ |

Table 5.4: Estimates of the rotational signature of BHB stars spectroscopically selected by X11. We find that stars with high $c_\gamma$ present a pathological correlation between line cuts and kinematics and reproduce the popular result of a metal-rich pro-rotating halo and a metal-poor counter-rotating halo. Unbiased stars instead confirm our findings of a weakly pro-rotating or non-rotating at all halo and the absence of signal between populations with different metallicity.

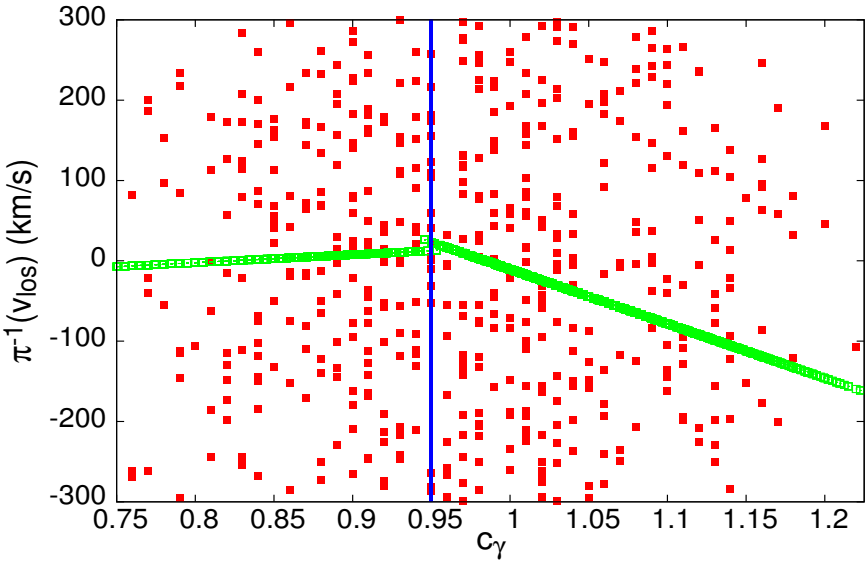


Figure 5.10: Deprojected l.o.s. velocity for metal-poor BHB stars spectroscopically selected by X11 as function of the line indicator $c_\gamma$. Only stars with $\pi^{-1}(v_{\rm los}) \in [-300, 300]\,\mathrm{km s}^{-1}$ are shown (red squares), but the least square fit to the $\pi^{-1}(v_{\rm los})$-trends plotted here with green lines have been obtained on the full sample.

from a common underlying population. Fig. 5.11 shows Kolmogorov-Smirnov tests on the discrepancy between the distribution of $v_{\rm los}$ (upper panel) and $\pi^{-1}(v_{\rm los})$ (lower panel) for's the four sub-samples. We plot the difference in the cumulative probability distributions of metal-poor and metal-rich stars in the two $c_\gamma$ bins (low $c_\gamma$: magenta line; high $c_\gamma$: purple line) as well as the difference between sub-samples of same metallicity, but different $c_\gamma$ (metal-rich stars: red line; metal-poor stars: blue line). For each test we show the 5% significance threshold with horizontal lines of the same colour. The discrepancy between two distributions is statistically significant if this threshold is crossed.

Indeed, we find that the most significant deviation occurs between metal-poor and metal-rich stars with $c_\gamma \geq 1.0$ (or, almost equivalently, for $b_\gamma > 9.5$) and between metal-poor stars with different $c_\gamma$. By contrast, stars with $c_\gamma < 1.0$ present no kinematic signal between sub-samples of different metallicity. In addition the significance of the signal between metal-poor stars in different $c_\gamma$ bins has exactly the same structure as the one between metal-rich and metal-poor stars with high $c_\gamma$ (and same significance). Therefore, the Kolmogorov-Smirnov test confirms that the signature in the X11 sample is indeed caused by high $c_\gamma$ metal-poor stars. This inconsistency within the metal-poor sub-sample, raises severe doubts about the reality of the observed differences.[15]

Even though with the current data we cannot rule out the possibility that these stars are associated with a real sub-structure in the halo, the $v_{\rm los}$-estimator (see §3.4.2.1) measures a non-zero gradient on these stars, pointing again to a pipeline problem.

These stars ($\sim 500$) are hot and very metal-poor so their spectra harbour nearly no metal lines that can support a radial velocity measurement, so this anomaly occurs in exactly the group of stars where we have the least trust in the velocity measurement. As can be seen from Table 5.4, neither of the two 3D estimators show an appreciable retrograde signal. Proper motions seem not to contribute to the rotational signal and the weak one measured appears to come from $v_{\rm los}$ alone. In addition the $V$-$W$ estimator measures an underestimate in distance of $f = -0.10 \pm 0.09$ for these stars. Since there cannot be a significant contamination with more luminous stars (a contamination with blue stragglers or blue hook stars would give the opposite sign), this points to unusual kinematics in the vertical direction rather than a simple retrograde motion.

[15]When we opt for the pipeline adopted metallicity instead of the WBG metallicity, the discrepancy remains and in fact is exacerbated.

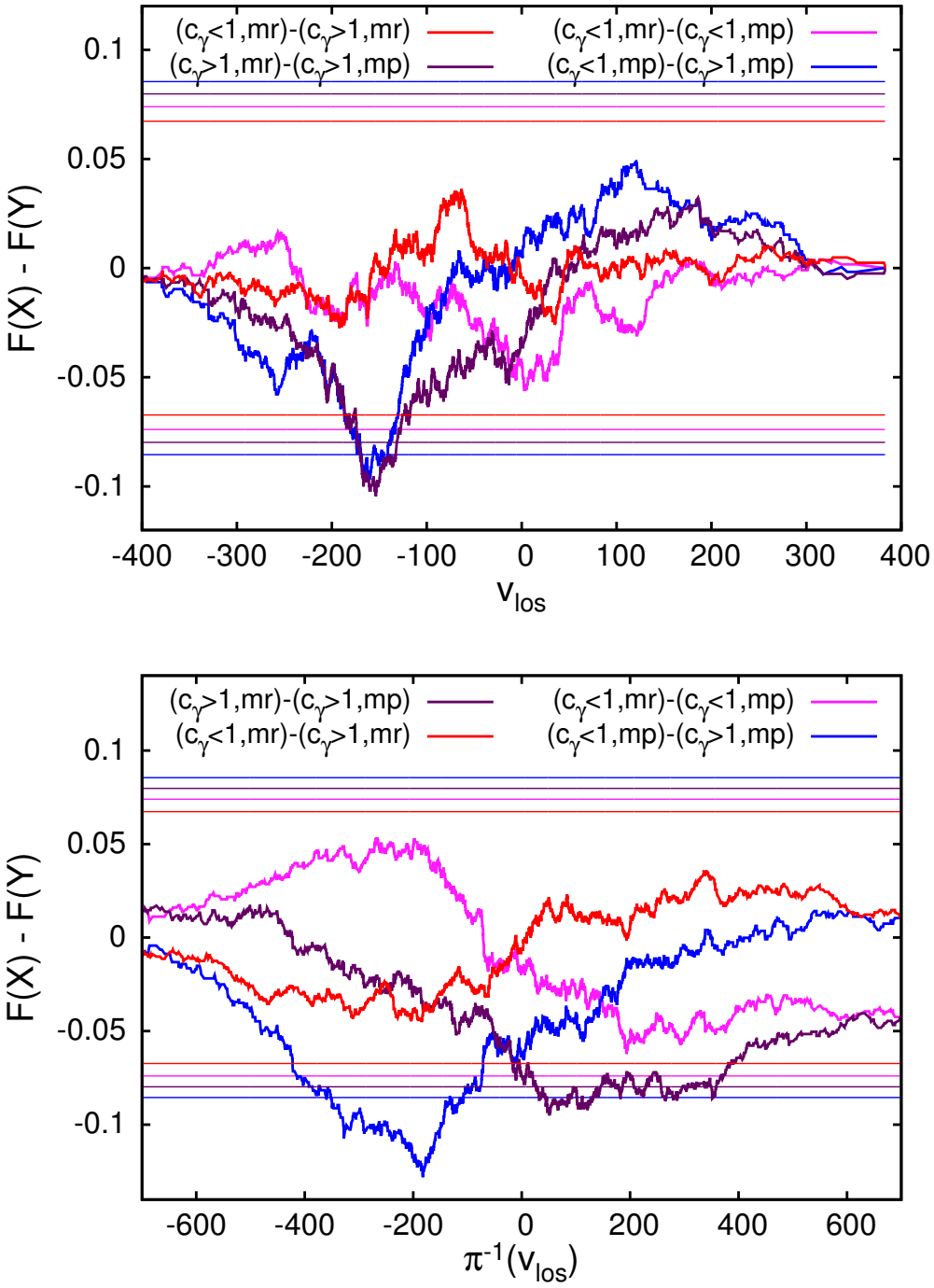


Figure 5.11: Differences in the cumulative distributions of l.o.s. velocity (upper plot) and de-projected $v_{\rm los}$ (bottom plot) between sub-samples of different metallicity in two $c_\gamma$ bins. Red lines represent the discrepancy between metal-rich stars with different $c_\gamma$, blue lines the same but for metal-poor stars. The magenta lines depict the difference between the velocity distributions of metal-rich and metal-poor stars with high $c_\gamma$, the purple lines the same but for stars with low $c_\gamma$. The horizontal lines are associated with the corresponding 5% significance levels of the Kolmogorov-Smirnov test.

In conclusion, we find that metal-poor stars with high $c_\gamma$ do show a retrograde signal and we identify them as the source of the line-of-sight velocity distribution anomaly observed by Beers et al. (2012); Hattori et al. (2013); Kafle et al. (2012). By their inconsistency with their low $c_\gamma$ counterparts we flag these either as a pipeline problem or a special structure rather than as a second smooth component.[16] If it were real, this structure would predominantly contain only hot BHB stars and have large systematic $W$ velocities across the sky while not showing a consistent retrograde motion. On the other hand, trustable stars (low $c_\gamma$) confirm our previous conclusions of a weakly pro-rotating or non-rotating halo irrespective of metallicity. Equivalent behaviours emerge when splitting the sample according to a line cut associated with $b_\gamma$.

## 5.5 Conclusions

We have analysed the kinematics of BHB stars drawn from SDSS/SEGUE (DR7, DR8, DR9), for which the only reliable measured component of velocity is $v_{\rm los}$, using both a selection via photometry and the sample of Xue et al. (2011). Using a very simple model and three model-independent estimators we find no evidence for rotation. Also, we measure no discrepancy in rotation between metal-poor and metal-rich stars. Following Carollo et al. (2007, 2010) we look for a discrepancy in the kinematics of the inner and an outer halo, but find none in any of the possible selections. In both samples we trace back previously claimed rotational differences between different metallicity bins to stars that don't pass stricter quality criteria.

Our analysis is not restricted to the simple double power-law model used by Deason, Belokurov & Evans (2011a), but relies on three direct model-independent estimators: one uses the de-projected $v_{\rm los}$ and two make use of the full 3D motion. The latter require knowledge of proper-motion and while individual BHB stars are too remote to have sufficiently accurate proper-motion estimates, the errors being

[16] We also note that the discrepancy in metallicity distribution that Beers et al. (2012) advocate as evidence for a double-halo arises when comparing stars closer and further than 10 kpc. The shift of 0.3 dex is purely due to stars between 5 and 10 kpc, which still have a non negligible probability of being disk contamination (hence why both in Deason, Belokurov & Evans (2011a) and here they are tossed out). The tangentially biased metal-poor halo found by Hattori et al. (2013) on the other hand, is a consequence of the pseudo-rotation of high $c_\gamma$ stars, but the robustness of their measurement is also undermined by neglected treatment of possible contaminations. A minuscule fluctuation arising from a distance misestimate for example, can be enhanced to arbitrary level profoundly biasing the inferred anisotropy, hence the duty to quantify their vulnerability to distance errors and to deviations from the assumption of sphericity.

comparable to the signal, on the entire sample proper motions prove useful by large-number statistics. We checked that the 3D-estimators are robust both against change in the used Sloan data release and against systematics of the magnitude of the correction derived in Schönrich (2012). Systematics of higher magnitude would bias the 3D-estimators differently: the fact that they agree and that we recover the correct radial reflex motion $U_0$ confirms that the systematic floor is within the formal errors and the proper-motions useful. Moreover, the use of the 3D-estimators is crucial to test the results from the 1D-estimates that are highly vulnerable to pipeline systematics. A systematic shift in line-of-sight velocities causes a larger shift in the measured rotation, since the errors are amplified by the geometric factors and only partly balanced by having both sides of the sky. Further, a non-zero "rotation" in the one-dimensional estimators does not prove a true retrograde or prograde net motion, but can be caused by failure of the assumption of zero mean $U$ and $W$ component.

Model-wise, overestimates in distance cause the portion of velocity space that we expect to be occupied to shrink, so certain stars are pushed out beyond the escape velocity with the consequence that the correct model assigns them zero probability. In this case the stars are eliminated from the group that contributes to estimates of the population's kinematics. This effect is not mirrored for distance underestimates, making the bias more severe. Unless observational uncertainties are appropriately folded with the model, again the 3D estimators are a necessary confirmation of any claim of net rotation.

Re-analysing the photometrically selected sample of Deason, Belokurov & Evans (2011a) we can recover their difference between prograde metal-rich and retrograde metal-poor populations. However, the signal is inconsistent across the sky and among the different estimators. In particular, the weak prograde signal in the metal-rich subsample shrinks considerably when we increase the geometric disk cut from $|z| > 4$ to $|z| > 6$ kpc. From the point of theory, it is plausible that due to accretion or frictional effects with the disc a confined region in the Galactic halo could display a mildly prograde rotation. However, the observations point to a disc contamination, given that the signal strongly weakens for the cleaner samples and that the 3D estimators do not confirm the rotation. We note that the additional constraints only reduce the level of contamination, while they do not a priori bias the inferred kinematics; when we repeat the same analysis on X11 we measure in fact homologous trends (see below). We further argue that the anisotropy measured via the adopted model is unreliable due to the unphysical $L^{-2\beta}$ term in the distribution function, which for $\beta < 0$ produces a bimodal distribution in the azimuthal velocity component.

We repeat our analysis on the spectroscopically selected sample of Xue et al. (2011). On the intersection of the two samples we do not find any hint for rotation. However, on the entire X11 sample we detect a weak offset in $v_\phi$ between metal-poor and metal-rich stars. While there is no detectable difference between the velocity distributions of metal-poor and metal-rich stars at low $c_\gamma$, we identify the metal-poor stars with high $c_\gamma$ as the source of this discrepancy. A Kolmogorov-Smirnov test reveals that the $v_{\rm los}$ distribution of these stars is not only offset significantly from that of metal-rich stars, but also from that of their metal-poor counterparts with low $c_\gamma$. Furthermore the retrograde behaviour in $v_{\rm los}$ is not confirmed by the 3D estimators, and anomalous $W$ velocities are detected. In fact these systematics are confined to metal-poor hot BHB stars that have very weak to non-detectable metal lines, which are crucial for reliable $v_{\rm los}$ determination.

There are two possibilities: either this is a pipeline problem or these stars belong to a very metal-poor population containing only hot BHB stars. A systematic shift in line-of-sight velocities is not unlikely, given that the determination of line-of-sight velocities is very difficult in the smooth and nearly metal-line free spectra of hot, metal-poor BHB stars and that the average motion of stars towards the Galactic North and South pole indicates a significant pipeline systematic that increases towards blue, metal-poor stars in general. If the component were real, it would not be in smooth retrograde motion, but rather likely form a dispersed stream-like structure with large systematic $W$ velocities. High resolution spectral analysis of the suspected culprits will be required to decide if the SEGUE $v_{\rm los}$ are trustable for those stars, and if so, to identify the peculiar abundance pattern that can be expected for such a unique metal-poor structure.

We tested our results against a variety of sources of uncertainty such as selection criteria, absolute magnitude estimate, contamination by main-sequence A stars, the Sun's parameters, the effect of reddening from the disk[17] and distance errors up to 20%. The results hold and prove to be susceptible only to variations in the Sun's circular velocity, for which we provide none the less a trend: changing $v_{c\odot} = 220\,{\rm km}s^{-1}$ to $v_{c\odot} = 250\,{\rm km}s^{-1}$ shifts the rotational signature towards pro-rotation by less than $10\ {\rm km}s^{-1}$.

The failure of the DF based model used here teaches us that significant efforts have to be devoted to the creation and selection of appropriate models and strategies

[17]We studied the effect that the choice to cut for $|z| < 4\,{\rm kpc}$ has on our results. We consider different cuts at $|z| = 1, 2, 3, 5, 10\,{\rm kpc}$ and observe no change in the signatures retrieved with our sample.

for fitting them to data. So far in fact, there is no a priori way to guess which characteristics a model should possess to fit satisfactorily the data and in fact, there is no agreed general definition of goodness of fit. There are ways to compare a model's ability to represent a certain sample (e.g. by computing the likelihood of the data given the model), but even with the best model among the ones available, there is no way to assess its intrinsic uncertainty.

We believe this area is still largely unexplored and should be considered the key step to cover for any meaningful analysis to be pursued.

# Part III

# Dynamics

# Chapter 6

# $f(\mathbf{J})$-models for spheroids

Surveys with unprecedented coverage of the Sky and measurement accuracy have opened to unravelling the mechanisms of formation and evolution of our Galaxy (e.g. SDSS Ahn et al. 2012) and of external galaxies (e.g. SAURON, Davies et al. 2001). Such a rich variety of information requires a coherent framework for its interpretation so that falsifiable predictions can be made. Dynamical models become therefore a necessity for extracting science from observations.

The class of analytical models available is effectively very narrow (see Binney & Tremaine 2008, for a summary) and is confined to toy profiles that either do not possess realistic features such as anisotropy (e.g. $f(\mathcal{E})$ models) or flattening (e.g. $f(\mathcal{E}, L)$ models), or present intrinsic divergences that constrain the model to be only limitedly meaningful. For example models with a distribution function (DF) of the type $L^{-2\beta} f(\mathcal{E})$, are still widely used (e.g. Evans, Hafner & de Zeeuw 1997; Deason, Belokurov & Evans 2011a; Watkins, Evans & van de Ven 2013) even though they produce an unrealistic bimodality in velocity space ($\beta < 0$) or an infinite phase-space density as $L \to 0$ ($\beta > 0$); still no clearly superior alternatives are available.

Numerical models that fit realistic systems can be grouped in essentially two classes: models based on the Jeans equations and Schwarzshild models. Jeans' equations (Jeans 1915) require an extra condition to be closed and the only intellectually defendable one, $\sigma_r = \sigma_z$, is ruled out by observations of the solar neighbourhood. The incomplete and widely debated understanding of the anisotropy of galaxy halos makes any alternative attempt to close Jeans' equations, hard to safely accept. All the same, models based on the Jeans equations, for example the anisotropic models in axisymmetric potentials by Cappellari (2008), have been used to exploit the SAURON data within the ATLAS$^{3D}$ project (Cappellari et al. 2011).

Schwarzschild (1979) argued that we should think of a galaxy as a superposition of orbits: one first creates a library of orbits in a given potential and then assigns each a

weight so that their sum fits the observations. If there is no acceptable fit consistent with the non-negativity condition on the weights, the potential is ruled out. These models appear to have been successful in constraining, for example, the dark-matter halos of galaxies (e.g. Rix et al. 1997), but their rigorous use requires a marginalization over all possible models consistent with a given potential (Magorrian 2006) otherwise they produce the paradox that increasing the quality of the data leads to a decrease in the ability to constrain the potential (Valluri, Merritt & Emsellem 2004). Further the uncertainty in the choice of the initial conditions for orbits' integration as well as on the level of smoothing required for a credible solution make the physical interpretation of these models opaque and the computational costs make them inconvenient (see Chanamé, Kleyna & van der Marel 2008, for a review).

Fundamentally, what most dynamical models lack is a clean and immediate interpretation of the physics implied by the model: the multi-Gaussian expansion (Emsellem, Monnet & Bacon 1994) used in the Jeans anisotropic models or the series of weights in Schwarzshild models fragment the system in a puzzle of physically meaningless components.

When conceiving a dynamical model we require it to depend on three integrals of motion to achieve a realistic description of the system, but we also want an immediate intuition about the physics implied by the model. Clarity in this sense is best achieved by defining our distribution function in terms of actions. They constitute a complete and well defined set of integrals (Jeans's strong theorem): the radial action $J_r$ measures the radial oscillation, the vertical one $J_z$ quantifies the oscillation perpendicular to the equatorial plane and $J_\phi$ is the angular momentum. On the contrary there is still debate on the classical integrals such as $L$ or $E_z$ as third integral of motion (Contopoulos 1963; Ollongren 1965; Gebhardt et al. 2000) so that the only viable scheme to obtain $I_3$ is indeed action computation. This and the fact that one of the first two integrals must be the angular momentum anyway, makes it a conceptual muddle to choose anything other than the triplet of actions as the full set of integrals. Moreover, one of the practical advantage in using actions is that they are adiabatically conserved: invariance of actions under slow variations of the potential, implies for example that sub-structures are more easily identifiable in action space (Helmi et al. 1999; McMillan & Binney 2008).

In practice we need to convert observables into angle-actions to fit an $f(\mathbf{J})$ model to the data: Binney (2012a) has recently shown how to compute actions from phase-space coordinates for any axisymmetric potential $\Phi(R, z)$ in a computationally cheap manner. The usefulness of $f(\mathbf{J})$ models has been proven for the thin and thick disk

by Binney (2010, 2012b), who provides a satisfactory account of solar neighbourhood observations through these models. Pontzen & Governato (2013) presented encouraging results on the ability of $f(E, \mathbf{J})$ models consistently to describe dark matter halos from cosmological simulations.

Here we present a procedure for creating $f(\mathbf{J})$ models for a spheroidal component of specified shape and velocity anisotropy in a general axisymmetric potential. We extend isotropic spherical DFs $f(E)$ by writing the Hamiltonian as function of the actions and then allowing a rescaling of the actions in $E(\mathbf{J})$. The dependence on the three actions both guarantees the flexibility of the DF to describe realistic stellar systems and offers a clear interpretation of the physical effects produced by a particular choice of the DF parameters. We also explore the alternative strategy of separating the energy dependence from the density of stars on action-space surfaces of constant energy.

In Section 2 we construct an approximation to the Hamiltonian as a function of the actions in a general axisymmetric potential and test the accuracy of our scheme against a known analytical profile (Isochrone potential) and the scheme of Binney (2012a). Section 3 defines the generalized DFs for spheroidal components, checks their consistency and presents some instructive examples. In section 4 we discuss the alternative strategy of Binney & Tremaine (2008) to separate the energy dependence from the density of actions in writing an $f(\mathbf{J})$ model. Section 5 sums up and looks to the future.

The techniques and the considerations presented in this chapter are the result of a collaboration with James J. Binney.

## 6.1 Obtaining $E(\mathbf{J})$

Techniques for handling angle-action space are summarized in Fig. 6.1: from the left, if we start from phase-space coordinates $(\mathbf{x}, \mathbf{v})$ we can easily determine the energy $E$ and now also the angle-actions $(\mathbf{J}, \theta)$ (Binney 2012a). If we start from angle-action space (centre bottom of the figure) we can use the Torus machine (Kaasalainen & Binney 1994; Binney & McMillan 2011) to compute $(\mathbf{x}, \mathbf{v})$ and can then compute $E$ (dashed red line), but we cannot determine $E$ directly from $\mathbf{J}$ (solid red line).

Estimating $H(\mathbf{J})$ via the Torus machine would require constructing full torii for each triplet of actions for which we want to evaluate the energy, which is computationally very inefficient on top of adding an unnecessary degree of complexity to the problem.

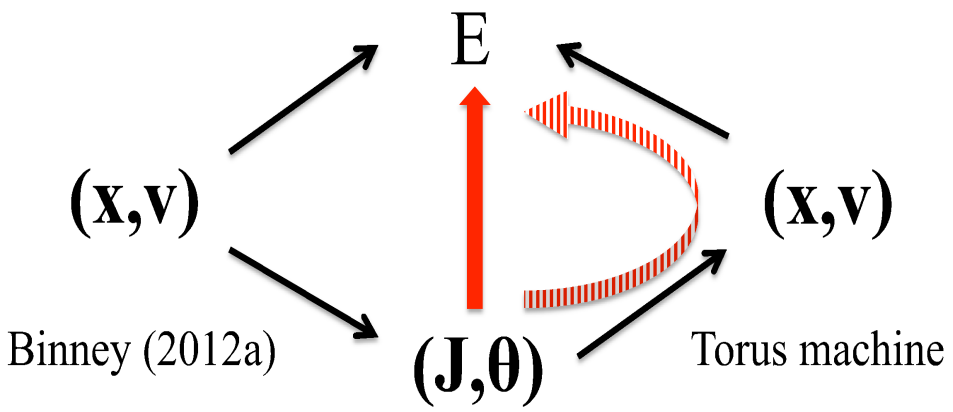


Figure 6.1: A visualization of the techniques available to move from phase-space to action space and vice-versa. The solid red line represents the missing link to compute energy from angle-action coordinates.

Here we approximate $H(\mathbf{J})$ by tabulation on an energy grid $\mathcal{G}_E$ in the form of a library of analytic fits to the surfaces of constant energy $\mathcal{S}_E$.

For the harmonic oscillator or the Kepler potential, surfaces of constant energy are planes in action space. For the Isochrone profile, they are essentially quadratics (which actually are not far from planes). Thus, the natural starting point is to fit a quadratic surface to a sampling of each $\mathcal{S}_E$. Obtaining $E$ at a general point of action space requires the Torus machine so the construction of a library of torii for each choice of the potential, which is precisely what we wish to avoid. We show that very few (of order of 15) special orbits accessible via direct integration, are sufficient to constrain a surface of constant energy.

### 6.1.1 Constrain $\mathcal{S}_E$ via special orbits

Let us consider a surface $\mathcal{S}_E$ of constant energy $E$ in the general axisymmetric potential $\Phi(R, z)$. The edges of $\mathcal{S}_E$ represent special orbits for which at least one of the three actions is zero. Reading the top panel of Fig. 6.2 clockwise, on the edges of a surface of constant energy we can identify shell orbits (no radial action), planar rosettes (no vertical action) and box orbits (no angular momentum). Actions for the first two types of orbits turn out to be particularly easy to compute; box orbits are instead generic orbits, but for reasonably flattened systems their action computation allows a tractable approximation (see below).[1]

[1]A combination of orbit integration and Floquet analysis returns the frequencies of the orbits named above. We tried to include frequency information in the least square fit: comparison between fits that use no frequency information and the ones where the calibrating data set included

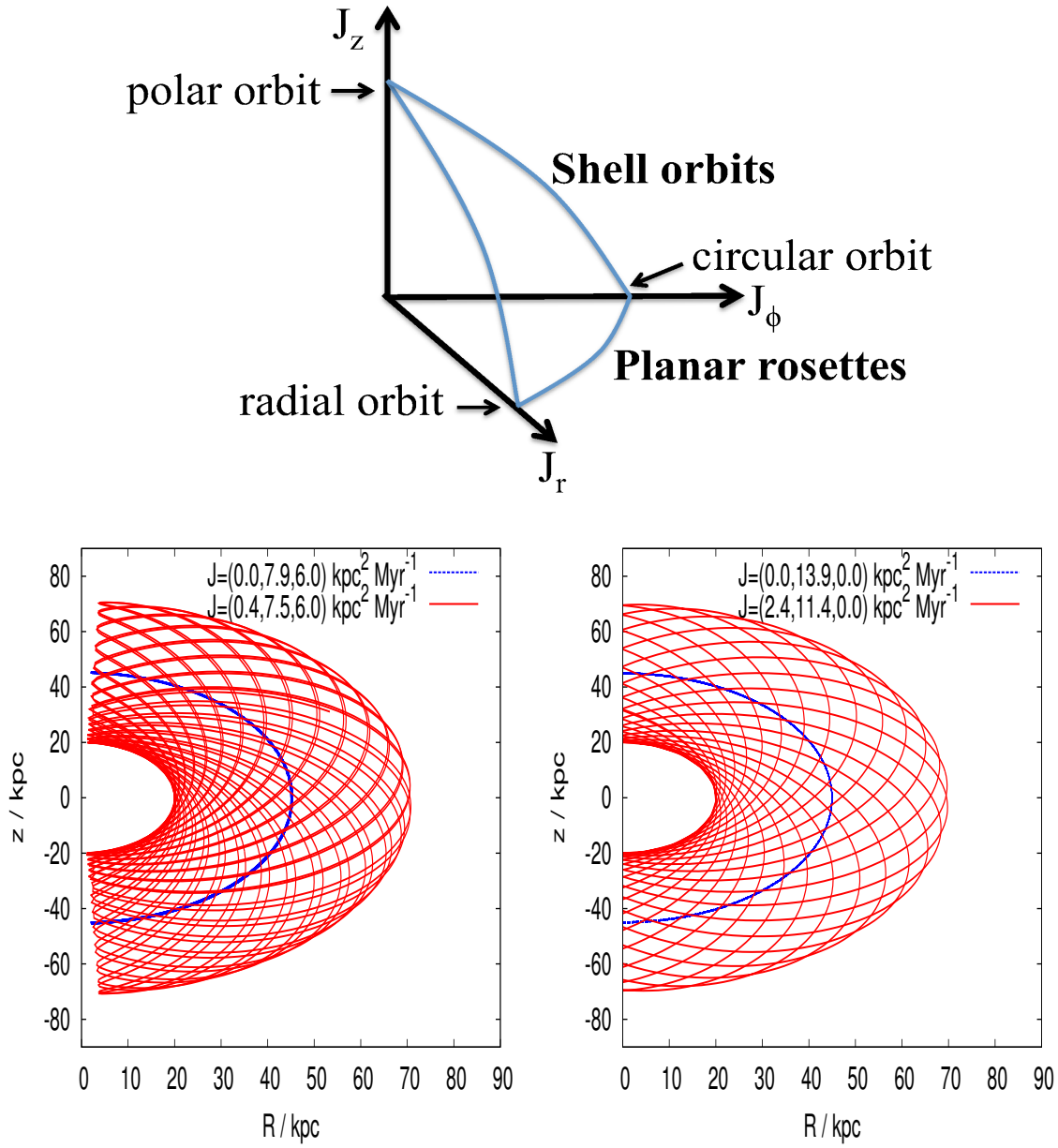


Figure 6.2: Surface of constant energy in action space, with identification of the orbits where at least one action is zero (top panel). The other two panels show the evolution in the $(R, z)$-plane of four orbits in an Isochrone potential ($b = 1.0\,\mathrm{kpc}$, $M = 10^{12} M_\odot$). The bottom left panel compares orbits with non-zero angular momentum, while in the bottom right panel $L_z = 0$: in both plots the red lines refer to orbits with non-zero radial action and dashed blue lines to zero radial action orbits. In the bottom left panel, the dashed blue line is a shell orbit, while in the bottom right panel the blue line represents a polar orbit and the red one a box orbit. All orbits have the same energy of $E = -0.05\,\mathrm{kpc}^2\mathrm{Myr}^{-2} M_\odot$ and were integrated for $\sim 30$ periods.

When the vertical action $J_z$ is zero, orbits are confined in the equatorial plane $z = 0$ and form **rosettes** (e.g. §3.1, Binney & Tremaine 2008). The size of the non-accessible circle at the centre of the rosette refers to the size of the angular momentum and in the limit of zero $L_z$ it shrinks to a point and the star goes in and out with infinite precession time producing a pure **radial orbit**. On the other hand in the limit $L_z \to \max L_z$, $J_r = 0$ and we recover a circular orbit. In general, once a value for the angular momentum has been specified $L_z = J_\phi$, the radial action can be computed via the simple formula (Binney & Tremaine 2008):

$$J_r = \frac{1}{\pi} \int_{R_{\min}}^{R_{\max}} \mathrm{d}R \sqrt{2E - 2\Phi(R, 0) - L^2/R^2} \tag{6.1}$$

where $L = |J_\phi|$.

With zero radial action we have a **shell orbit**, this star oscillates in the meridian plane precessing in the $\phi$-direction with speed $L_z/R^2$. An orbit of this type is shown for an Isochrone potential in the bottom left panel of Fig. 6.2 with a blue line. For comparison we plot in red an orbit with the same energy and angular momentum, but non-zero radial action: the width of the tube that contains the orbit is a monotonic function of the radial action. Neither orbit can approach the centre: the volume of non-accessible region around the centre is instead a monotonic function of the total angular momentum.

In the limit of zero angular momentum the precession time is infinite and the star effectively oscillates in a fixed $(R, z)$-plane leading to a **polar orbit**. The bottom right panel of Fig. 6.2 shows a polar orbit in an Isocrhone potential (blue line): even though quite similar to a shell orbit, the zero-angular momentum allows the orbit to explore lower radii. Again for comparison, we plot in red an orbit with the same energy, zero angular momentum, but non-zero radial action (a box orbit, see below): like in the previous case, the width of the tube that contains the orbit is an increasing function of the radial action.

For both shell and polar orbits, given that there is zero radial action, the orbit crosses the equatorial plane at a fixed $R^*$ (see blue lines in Fig. 6.2, bottom panels). Given $L_z$ we can determine the crossing radius as follows: we take the mid point $R^{(0)}$ of the range $[R^{(0)}_{\min}, R^{(0)}_{\max}]$, where $R^{(0)}_{\min} = 0$ and $R^{(0)}_{\max}$ is such that $\Phi(R^{(0)}_{\max}) = E$. We integrate the equations of motion from $(R, z, p_R, p_z) = (R^{(0)}, 0, 0, \sqrt{2E - 2\Phi_{\text{eff}}(R^{(0)}, 0)})$

frequencies shows that the exercise is not worthwhile.

| **Orbit (Limit)** | **Actions** | **Computation** |
|---|---|---|
| Shell (Circular, Polar) | $(0, J_z, J_\phi)$ | Iterative bisection of $[0, R_{\max}]$ to search for crossing radius, direct integration of $p_z$ on closed orbit |
| Rosette (Circular, Radial) | $(J_r, 0, J_\phi)$ | $J_r = \frac{1}{\pi}\int_{R_{\min}}^{R_{\max}} \mathrm{d}R \sqrt{2E - 2\Phi(R,0) - J_\phi^2/R^2}$ |
| Box (Polar, Radial) | $(J_r, J_z, 0)$ | $J_r = \frac{1}{\pi}\int_{R_{\min}}^{R_{\max}} \mathrm{d}R \sqrt{2E - 2\Phi(R,z) - J_z^2/R^2}$ |

Table 6.1: Methods and formulae to determine a complete triplet of actions given a potential $\Phi(R, z)$ and an energy $E$ for orbits associated with the edges of a surface of constant energy in action space.

until we cross the plane again at $R_c$. If $|R_c - R^{(0)}|$ is below a certain accuracy threshold, we take $R^{(0)}$ to be the crossing radius, otherwise we update the search range:

$$\begin{aligned} R^{(1)}_{\min} &= R^{(0)}_{\min},\ R^{(1)}_{\max} = R^{(0)} \quad \text{if } R_c < R^{(0)} \\ R^{(1)}_{\min} &= R^{(0)},\ R^{(1)}_{\max} = R^{(0)}_{\max} \quad \text{otherwise,} \end{aligned}$$

and iterate the above until convergence to a close orbit. $J_z$ is then computed by direct integration of $p_z$ along the final orbit:

$$J_z = \frac{2}{\pi}\int_0^{z_{\max}} \mathrm{d}z\, p_z \tag{6.2}$$

The iterative search proves to be computationally cheap in our test cases (see §6.1.1.1) with an execution time of less than $10^{-2}$ seconds.

Setting $J_\phi$ to zero produces **box orbits**, which are generic non-closed orbits that move in the $(R, z)$-plane. Actions for these orbits are more difficult to track down: given that they are non-periodic, the two-dimensional manifold they produce in real space is a complete expression of the Hamiltonian and cannot be reduced to a 1D integral like in the previous cases (e.g. shell orbits and rosettes). For spherical potentials though, $L = J_z$ when $J_\phi = 0$ and box orbits are exactly the same as planar rosettes. For reasonably flattened systems it is therefore legitimate to expect that the radial action can still be estimated via (6.1) with $L \simeq J_z$.

**Quadratic fit to $\mathcal{S}_E$** To approximate $\mathcal{S}_E$ with a quadratic of the form:

$$\tilde{E}(\mathbf{J}) := \sum_{l=1}^{3} \left(a_l J_l + a_{l+3} J_l^2\right) + a_7 J_r J_z + a_8 J_z J_\phi + a_9 J_r J_\phi, \tag{6.3}$$

we need $N \geq 9$ orbits at each energy to determine $(a_1, \ldots, a_9)$ and in particular we need at least one point in each $J_i = 0$ plane, $i = r, z, \phi$, in order to determine the coefficients of all the mixed terms $(J_i J_j)_{i \neq j}$ in (6.3). Thus, we will consider the three

corner points and four additional points on each edge. The coefficients of (6.3) for each surface will be determined via linear fit by minimizing

$$\sum_{\zeta}(E - \tilde{E}(\mathbf{J}_{\zeta}))^2$$

where $\zeta$ indexes the sampled points on the surface edges.

#### 6.1.1.1 Examples

We first test the proposed approximation scheme with a spherical potential (Isochrone), quantifying the improvement of a quadratic fit w.r.t. a linear approximation. Then we validate our algorithm on a non-spherical potential, namely a Miyamoto-Nagai profile.

The convenience of the Isochrone potential is that we have an exact expression for $H(\mathbf{J})$ (see §3.5.2 of Binney & Tremaine 2008). Therefore, given an energy $E$ and a point in the $(J_z, J_\phi)$-plane for example, we can compare the analytic value of $J_r$ with the one evaluated from the approximation (6.3). Given $E, J_z, J_\phi$, we therefore compute:

$$\langle \Delta J_r/(J_r + L)\rangle = \langle (J_r - J_r^{\rm fit})/(J_r + L)\rangle \tag{6.4}$$

where $J_r^{\rm fit}$ is the radial action computed by root-finding on $\tilde{E} - E$ and $J_r$ is the analytic radial action[2]. We can also directly test the accuracy of Eq. 6.3 to estimate $E$ given a triplet of actions by comparing $E$ with the value of $H(\mathbf{J})$. Given $\mathbf{J}$, we consider

$$\langle \Delta E/E\rangle = \langle (\tilde{E}(\mathbf{J}) - H(\mathbf{J}))/H(\mathbf{J})\rangle \tag{6.5}$$

where $\tilde{E}(\mathbf{J})$ is Eq. 6.3 computed at $\mathbf{J}$ and all averages are over the domain $J_z, J_\phi \in \{J_z + |J_\phi| < L_{\rm max}\}$, uniformly sampled.

Furthermore, the form in which our approximation has been proposed allows a direct test of its ability to predict the frequency ratios of an orbit:

$$\Omega_i(\mathbf{J}) = \frac{{\rm d}E}{{\rm d}\tilde{E}}\frac{\partial \tilde{E}}{\partial J_i}(\mathbf{J}) \Rightarrow \frac{\Omega_i(\mathbf{J})}{\Omega_j(\mathbf{J})} = \frac{\partial \tilde{E}}{\partial J_i}(\mathbf{J}) \Big/ \frac{\partial \tilde{E}}{\partial J_j}(\mathbf{J}) \tag{6.6}$$

for $i, j = r, z, \phi$. The frequency ratios can then be easily compared with $(\partial H/\partial J_i)/(\partial H/\partial J_j)$ at each point of the surface.

[2]Note that the inversion of (6.3) allows two mathematically possible solutions, even though only one is physically meaningful. If the quadratic fit is an ellipsoidal surface, one of the solution is not in the quadrant of interest (i.e. $J_r > 0, J_z > 0$) and it is clear which solution is the physical one. On the other hand if the surface is an hyperboloid (i.e. at least one of the eigenvalues of the associated quadratic form is negative), then both solutions are in the right quadrant of action space. When that is the case, we must choose the solution that lies on the negative branch of the hyperboloid.

In the case of the Miyamoto-Nagai potential, no analytic formula is available for comparison: thus we check (6.3) against a surface of constant energy generated via the Stäckel approximation of Binney (2012a). To construct the test surface, we uniformly sample position space within $\Phi(R, z) \leq E$ and then choose a velocity which agrees with the prescribed energy[3]. Then we use the machine of Binney (2012a) to map $(\mathbf{x}, \mathbf{v})$ to $(\mathbf{J}, \theta)$. The machine also outputs the associated frequencies for each $\mathbf{J}$ so we can effectively compare the same quantities as for the Isochrone.

For an Isochrone potential, a quadratic fit to the edges of a constant-energy surface proves to be an improvement of two orders of magnitude w.r.t. a linear approximation in predicting energy and radial action:

| | Linear Fit | Quadratic Fit |
|---|---|---|
| $\langle \Delta J_r/(J_r + L) \rangle$ | $(7.5 \pm 0.1) \times 10^{-3}$ | $(4.4 \pm 0.05) \times 10^{-5}$ |
| $\langle \Delta E/E \rangle$ | $(8.0 \pm 0.1) \times 10^{-3}$ | $(3.7 \pm 0.05) \times 10^{-5}$. |

Fig. 6.3 shows the distribution of $\langle \Delta J_r/(J_r + L) \rangle$ and the actual $\Delta J_r$ as a function of $J_r$ (upper and lower panels respectively). The accuracy of (6.3) in encapsulating the physics of the system exceeds the design requirements: the error of (6.6) in estimating the frequency ratios is less than 0.1%: in Fig. 6.4 we plot the discrepancy between the estimates from (6.6) and the analytic formulae.

For a Miyamoto-Nagai potential, the quadratic approximation (6.3) shows a similar performance against the constant energy surface generated via the approximation of Binney (2012a) (see also Fig. 6.5):

$$\begin{aligned} \langle \Delta J_r/(J_r + L) \rangle &= (-3.4 \pm 6.0) \times 10^{-5} \\ \langle \Delta E/E \rangle &= (-3.6 \pm 0.3) \times 10^{-5}. \end{aligned}$$

The average discrepancy in the frequencies ratios computed via (6.6) and the ones obtained via Binney (2012a) is $\sim 3\%$: see Fig. 6.6. We note that this underperformance w.r.t. the Isochrone case is not to be attributed to our approximation, but rather to the test case that now is an approximation itself. Indeed, when we assess (6.3) for the Isochrone potential generating a surface via Binney (2012a) rather than using analytic formulae, we measure a similar 3% accuracy for the frequencies ratios.

[3]The actual distribution in phase-space is not of interest as long as all the points lie on a surface of constant energy in action space.

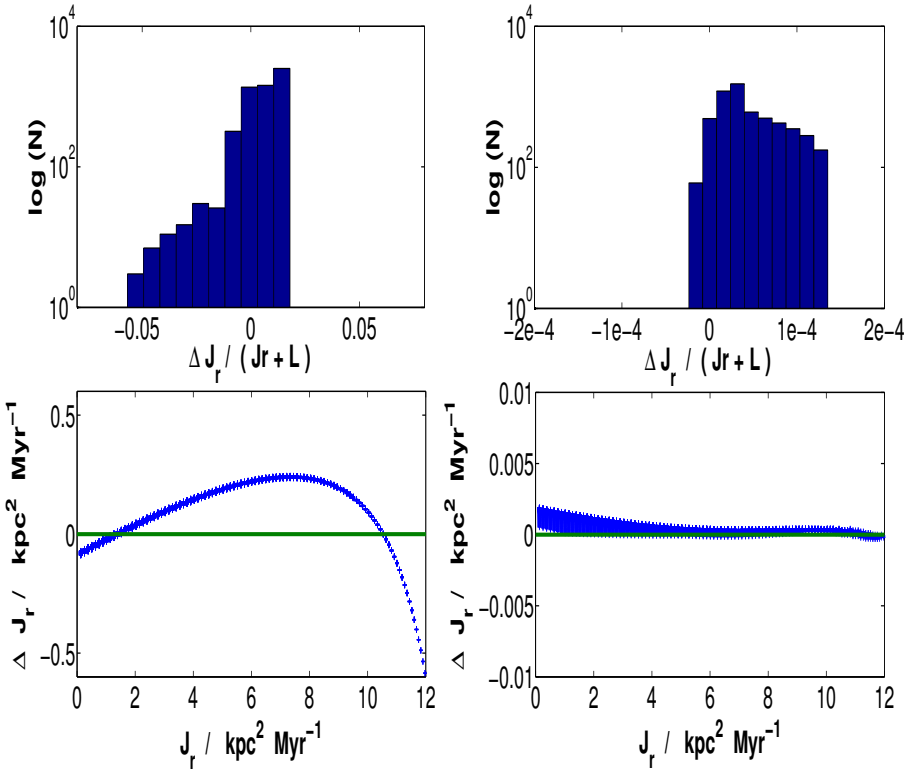


Figure 6.3: Distribution of $\Delta J_r/(J_r + L)$ on a logarithmic scale (top panel) and the actual $\Delta J_r$ against the analytic radial action (bottom panel) for an Isochrone potential ($b = 1.0\,\mathrm{kpc}$, $M = 10^{12}\,M_\odot$, $E = -0.05\ \mathrm{kpc}^2\mathrm{Myr}^{-2}M_\odot$, both distributions are marginalized over $J_z, J_\phi$). Left plots show the performance of the linear approximation, right plots that of the quadratic approximation (6.3).

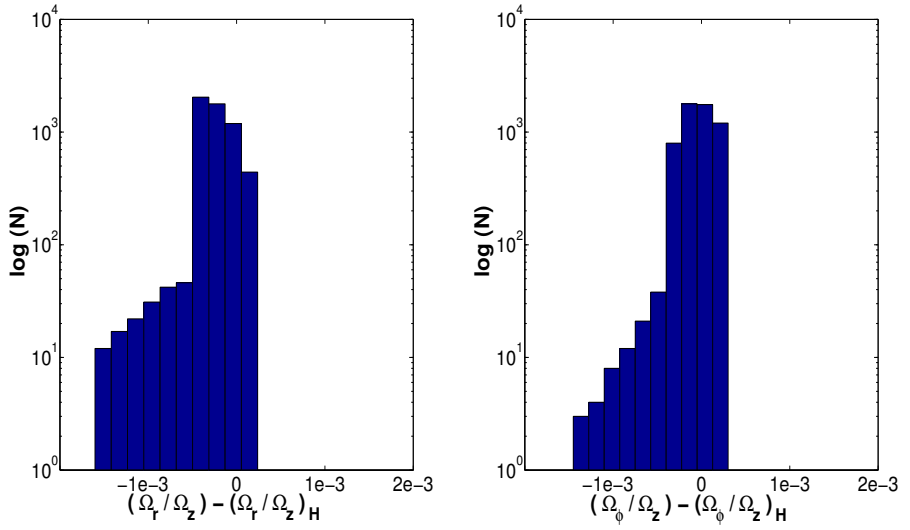


Figure 6.4: Discrepancy between the frequencies ratios computed via (6.6) and ones from the analytic formulae for the Isochrone potential $H(\mathbf{J})$. The distributions are in logarithmic scale and are marginalized over the surface $E = -0.05\ \mathrm{kpc}^2\mathrm{Myr}^{-2}M_\odot$.

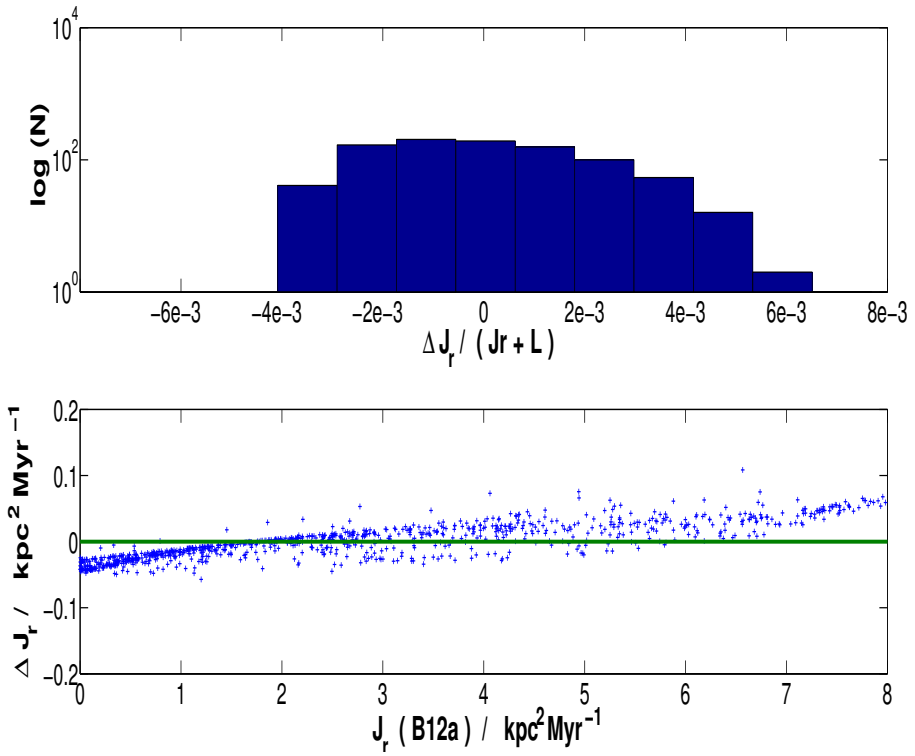


Figure 6.5: Distribution of $\Delta J_r/(J_r+L)$ in logarithmic scale (top panel) and the actual $\Delta J_r$ against the test radial action (bottom panel) for a Miyamoto-Nagai potential ($a = 1.0\,\mathrm{kpc}$, $b = 0.45\,\mathrm{kpc}$, $M = 10^{12}\,M_\odot$). Here, the test surface was generated via the approximation of Binney (2012a) and both plotted distributions are marginalized over the whole surface.

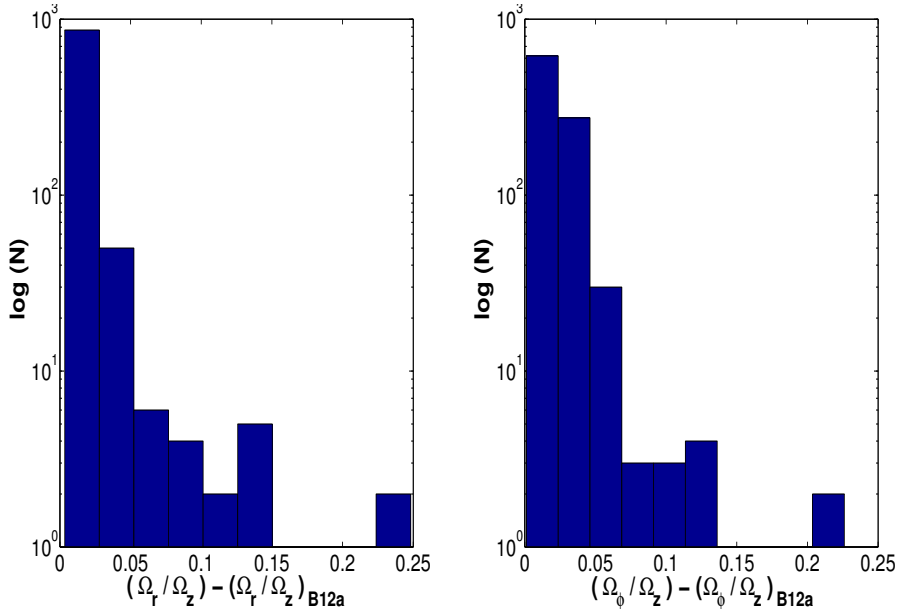


Figure 6.6: Same as Fig. 6.4 but for a Miyamoto-Nagai potential. As in Fig. 6.5 the test surface was generated via the approximation of Binney (2012a).

### 6.1.2 A grid of $\mathcal{S}_E$ to express $H(\mathbf{J})$

We have shown that a constant energy surface can be constrained using a small number of special orbits. We have given an indication of the accuracy of a quadratic fit in approximating $E(\mathbf{J})$ and the frequencies ratios at fixed $E$ via two different examples. We now construct a gird of these surfaces that by interpolation realizes a continuous approximation of $H(\mathbf{J})$.

Let us consider an energy grid $\mathcal{G}_E := \{E_i\}_{i\in I}$ and a corresponding library of constant energy surfaces in action space identified by their coefficients $\mathcal{A} := \{(a_1^i, \ldots, a_9^i)\}_{i\in I}$. We take a grid uniform in energy, in the range $[\Phi(0,0), 0]$ and if at a particular grid point the numerical fitting of the surface does not converge, we shift the value of the energy by 0.05%. To evaluate $H(\mathbf{J}) = E$, $E \notin \mathcal{G}_E$ then, one identifies the surface in the grid that lies closest to $\mathbf{J}$ and determines the coefficients $a_n(E)$ via interpolation on $\mathcal{A}$. Effectively the Hamiltonian becomes an implicit solution of $\tilde{E}(E, \mathbf{J}) = E$, where $\tilde{E}(E, \mathbf{J})$ is our approximation for $H(\mathbf{J})$:

$$\begin{aligned} \tilde{E}(E,\mathbf{J}) \;:=\; & \sum_{l=1}^{3} \left(a_l(E) J_l + a_{l+3}(E) J_l^2\right) + \\ & a_7(E) J_r J_z + a_8(E) J_z J_\phi + a_9(E) J_r J_\phi \end{aligned} \tag{6.7}$$

where $a_n(E) = \mathrm{interp}(a_n^i, E)$ and *interp* is any consistent interpolation scheme. Hereafter $\tilde{E}(\mathbf{J})$ denotes the root of $\tilde{E}(E, \mathbf{J}) = E$.

For example we can use basic linear interpolation and for the Isochrone case, this very simple scheme achieves a remarkable accuracy already with a grid of $\sim$ 300 surfaces:

$$\left\langle \frac{\Delta E}{E_{\max}} \right\rangle = \left\langle \frac{\tilde{E}(E,\mathbf{J}) - H(\mathbf{J})}{H(\vec{0})} \right\rangle = (2.4 \pm 1.4) \times 10^{-4}, \tag{6.8}$$

the average being over action space. In Fig. 6.7 we check that the accuracy of (6.7) scales with the number of surfaces.

## 6.2 Generalized $f(\tilde{E}(\mathbf{J}))$ models

The full dynamics of a stellar system is encapsulated in the knowledge of the underlying potential and of the distribution function of objects we observe. Here, we show how to generalize a spherical $f(E)$ model to an $f(\mathbf{J})$ model for a spheroidal component in a general axisymmetric potential and quantify the contribution of the potential itself in shaping phase-space distribution of the system.

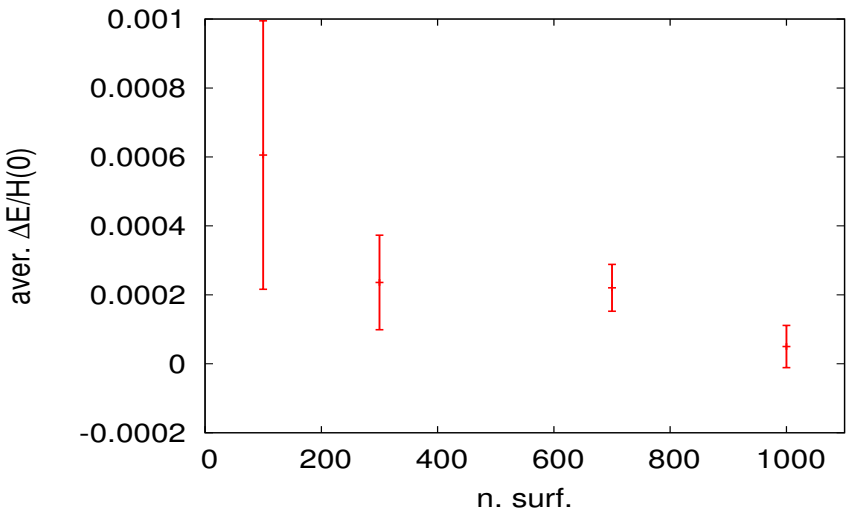


Figure 6.7: Accuracy of the approximation $\tilde{E}(\mathbf{J}) \simeq H(\mathbf{J})$: we plot $\langle \Delta E/E_{\max} \rangle$ as a function of surface library size. The error bars are the statistical error on the mean plotted.

The generalized models are built by making explicit the dependence of $E$ on the three actions with the approximation developed in the previous Section and then, by *rescaling* each action in the DF to control phase-space features such as flattening, anisotropy and rotation in a fixed potential. Complementarily, we choose a DF generated via the above algorithm and vary the underlying potential to measure the variations in the phase-space distribution of the system.

### 6.2.1 Rescaling of the actions

In a spherical potential, each feature in phase-space that departs from the spherical isotropic case is caused by non-uniform distribution of orbits on surfaces of constant energy. An excess of planar rosettes discourages vertical motion favouring flattening, while a higher density of shell orbits has the opposite effect. A strong population of high $L_z$-orbits deforms the system into a prolate or oblate shape according to whether it discourages radial or vertical motion respectively; any oddness in $L_z$, induces rotation.

Even if the functional shape of the DF of a given system is effectively unknown, we shall demonstrate that just the three numbers that specify the proportions in which each type of orbit occurs, are flexible descriptors of a wide range of systems. Equivalently, a system can be well represented by the three action scaling factors.

We define the generalized distribution function:

$$f(\mathbf{J}) := f(\tilde{E}(\tilde{\mathbf{J}})) \tag{6.9}$$

where $f(E)$ is any isotropic DF, $\tilde{E}(\mathbf{J})$ is the approximation constructed in the previous section and

$$\tilde{\mathbf{J}} = (\alpha_r J_r, \alpha_z J_z, \alpha_\phi J_\phi)$$

is the action triplet rescaled according to the vector $\alpha \in \mathbb{R}^3$. Given that $f$ is a monotonically decreasing function of $E$ which is a increasing function of $\mathbf{J}$, scaling factors smaller than one enhance the density of stars with the associated action and vice versa.

While this action rescaling induces features such as anisotropy, flattening and rotation we do not want to change the total number of stars at each energy. That is, we want to keep fixed:

$$\frac{N(E)}{(2\pi)^3} = \int \mathrm{d}^3\mathbf{J}\delta(E - H(\mathbf{J}))f(\mathbf{J}) = f(E)g(E) \tag{6.10}$$

where $f(\mathbf{J})$ is the generalized DF (6.9), $f(E)$ is the initial ergodic DF and $g(E)$ is the volume of phase-space associated with energy $E$:

$$g(E) = \int_{H(\mathbf{J})=E} \frac{\mathrm{d}J_\phi \mathrm{d}J_z}{\Omega_r}. \tag{6.11}$$

In essence, we need to impose that if the probability of a certain type of action is enhanced, that of a second type is decreased so the overall effect is a transfer of probability on surfaces of constant energy. In action space this can be achieved by rocking a surface of constant $f$ around its barycentre: probability can then flow between the axes, but without causing the radial profile to change.

Let us refer to $\mathbf{J}^m$ as a generic barycentre, either geometrical (e.g. mid-point) or physical. We consider the transition from $\tilde{E}(\mathbf{J})$ to $\tilde{E}(\tilde{\mathbf{J}})$ due to a step in action of $\mathrm{d}\mathbf{J} = \tilde{\mathbf{J}} - \mathbf{J}$. We impose the invariance of the constant-$f$ surface at $\mathbf{J}^m$ by requiring:

$$0 = \mathrm{d}H(\mathbf{J}^m) = \sum_{i=r,z,\phi} \frac{\partial H}{\partial J_i}\mathrm{d}J_i. \tag{6.12}$$

where here $\mathrm{d}J_i = (\alpha_i - 1)J_i^m$. Therefore, rearranging for $\alpha_z$ for example, $f(\tilde{E}(\mathbf{J}^m)) = f(\tilde{E}(\tilde{\mathbf{J}}^m))$ *iff*

$$\alpha_z = 1 + \left((1-\alpha_r)\frac{\Omega_r}{\Omega_z}(\mathbf{J}^m)\frac{J_r^m}{J_z^m} + (1-\alpha_\phi)\frac{\Omega_\phi}{\Omega_z}(\mathbf{J}^m)\frac{J_\phi^m}{J_z^m}\right). \tag{6.13}$$

and therefore we have two degrees of freedom. For example, if these are taken to be $(\alpha_r, \alpha_\phi)$, we may choose $\alpha_r \neq 1$, $\alpha_\phi = 1$ and then we are rocking the surface around its mid-point maintaining the $J_\phi$-axis fixed and the result is a transfer of probability between $J_r$ and $J_z$. Hereafter, we shall denote with $\alpha_x$ the scaling factor upon which we impose the consistency condition (6.12): given that this is a function of $\mathbf{J}$, we won't report its value in our tables.[4]

Some instructive special cases come indeed from probability exchange between just two of the $J$-axes: radial **anisotropy** is generated with any transfer from either the $J_z$ or the $J_\phi$-axis to the $J_r$-axis or vice versa for tangential anisotropy. The idea is that high $J_r$-orbits enhance radial motion, which increases $\sigma_r$ and hence the radial pressure. The enhancement of high $J_\phi$-orbits on the other hand, leads to higher $\sigma_\phi$ and the intensification of high $J_z$-orbits produces high $\sigma_\theta$: both imply tangential anisotropy. **Flattening**, can be achieved either with an enhancement of high $J_\phi$ orbits and a impairment of high $J_z$ orbits or by encouraging high $J_r$ orbits w.r.t. high $J_z$-ones. Vice versa **elongation** is induced by a transfer $J_r \to J_z$ or $J_\phi \to J_z$. To produce **rotation** the DF must have a part odd in $L_z$, $f_-$. The odd part can have any form or shape and the only constraint is on its modulus: $|f_-| \leq |f_+|$, where $f_+$ is the even part of the DF.

We briefly consider the set of statistics that will help us to describe a posteriori the phase-space distribution generated by our models. Ivezić, Beers & Jurić (2012) have recently stressed the importance of quantifying the discrepancy between $\sigma_\phi$ and $\sigma_\theta$ since if they are equal, the stellar density distribution would be spherical, if $\Phi$ were spherical. To investigate the validity of this implication we define the position-dependent parameter:

$$\beta_\delta := \frac{\sigma_\theta^2 - \sigma_\phi^2}{\sigma_r^2} \tag{6.14}$$

on top of the classical anisotropy parameter (Binney & Tremaine 2008):

$$\beta := 1 - \frac{\sigma_\theta^2 + \sigma_\phi^2}{2\sigma_r^2}. \tag{6.15}$$

If both $\beta$ and $\beta_\delta$ are zero, then all the second moments are equal and the velocity distribution is isotropic and in a spherical potential, the density distribution will be

[4]Condition (6.12) and the positiveness of two of the actions limit the range of the rescaling parameters to $\sim [0, 3]$; this comes from the fact that when there is no rescaling ($\alpha_i = 1\,\forall i$), $|\alpha|^2 = 3$ and a consistent shift of probability shows experimentally roughly to conserve $|\alpha|$. All the same, within an even smaller domain of rescaling, namely $\alpha \in [0.8, 1.2]^3$ we were able to generate flattening in the range of $[0.7, 1.2]$, anisotropy in the range of $[-0.4, 0.4]$ and rotation up to $450\,\mathrm{km}s^{-1}$ at $10\,\mathrm{kpc}$ in either directions (see details below).

spherical. On the other hand $\sigma_\theta^2 = \sigma_\phi^2$ does not imply that the tangential pressure is the same as the radial one: that is to say $\beta_\delta = 0 \not\Rightarrow \beta = 0$, so $\beta_\delta$ is a non redundant descriptor w.r.t. $\beta$. In conclusion, we adopt the following statistics to describe the stellar system: $q$ to quantify the flattening of the system and $(\beta, \beta_\delta, \langle v_\phi \rangle)$ to track its kinematics.

We now present a series of examples that realize flattening, radial anisotropy and rotation starting from a self-consistent Isochrone profile (with potential parameters $b = 1.0\,\mathrm{kpc}$, $M = 10^{12} M_\odot$). The phase-space features are estimated by direct computation of the first three moments of the distribution function $f(\mathbf{x}, \mathbf{v})$. To compute them one first maps $(\mathbf{x}, \mathbf{v})$ into $(\mathbf{J}, \theta)$ via the scheme of Binney (2012a)[5] and then computes $f$ in action space where it has been defined (see Eq. 6.9); for the $\tilde{E}$ approximation, we use hereafter a library of 300 surfaces of constant energy. The number of stars is approximately conserved by imposing (6.13); choosing the barycentre of a constant energy surface to be the point of equal actions, i.e. $J_i^m = J^m$ for $i = r, z, \phi$, the condition simplifies to[6]:

$$\alpha_z = 1 + \left(\Omega_r(1 - \alpha_r) + \Omega_\phi(1 - \alpha_\phi)\right)/\Omega_z. \tag{6.16}$$

We test the ability of the above simple condition to preserve the stars number by computing:

$$\delta_\rho(r) := 1 - \frac{\rho(r, r)}{\rho_H(r\sqrt{2})} \tag{6.17}$$

where $\rho$ is the zero-th moment of the model and $\rho_H$ is the analytic real-space density of the Isochrone profile.

The first DF we evaluate is the "control case", i.e. the model $\alpha = (1.0, 1.0, 1.0)$, which is the original isotropic DF. In Fig. 6.8 we compare its zero-th and second moments with the ones of the Isochrone profile. An upper bound for the discrepancy in either cases is $\sim 0.5\%$, which proves that the numerics of our implementation is under control. We then consider three generalised models, each associated with a different probability transfer: a $J_r \to J_z$ transfer, its inverse and a $J_z \to (J_\phi > 0)$ transfer, which we obtain by adding an odd part in $L_z$ to the DF such that only the positive $L_z$ survive; the parameters for each model are reported in Table 6.2. On each

[5]For the angle-action computation, the vertical coordinate of the foci of the ellipsoidal reference frame needed by the Stäckel approximation is set at $\Delta = 1.0$ (see §3.5.3, Binney & Tremaine 2008). In the specific case of the Isochrone potential, the actions can be computed analytically: in this case, we use the exact form of $\mathbf{J}(\mathbf{x}, \mathbf{v})$ to speed up the computation and make the approximation of the Hamiltonian the only source of uncertainty.

[6]Choosing $\mathbf{J}^m$ to be the geometrical mid-point of the surface of constant energy, i.e. $\frac{1}{2}(J_r^{\max}, J_z^{\max}, J_\phi^{\max})$, leads to undistinguishable results.

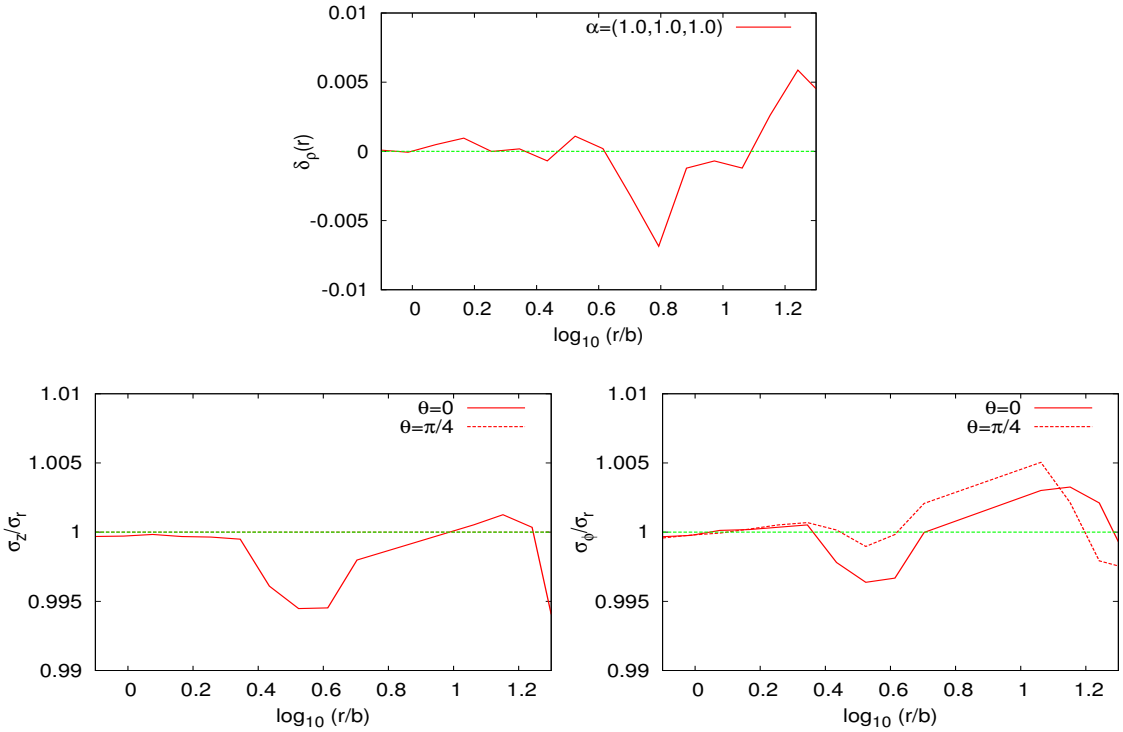


Figure 6.8: Discrepancy between the zero-th moment of the $f(\tilde{E}(\tilde{\mathbf{J}}))$ DF, where $f$ is the ergodic Isochrone DF and the rescaling parameters are $\alpha = (1.0, 1.0, 1.0)$, and the analytic radial profile of the Isochrone model (top panel, $\delta_\rho$ is defined in eq. 6.17). In the lower panels, comparison of the axis ratio of the velocity dispersion tensor. The control case $\alpha = (1.0, 1.0, 1.0)$ serves as test of implementation.

transfer we impose the condition (6.16) approximately to conserve the stars number: we plot $\delta_\rho(r)$ for the three models in Fig. 6.9 and note that the maximum discrepancy for either model is generously below 10%; moreover increasing the size of the surface library, further reduces the dispersion around zero.

Favouring high $J_r$-orbits, while impairing high-$J_z$ orbits, flattens the system (top-right panel of Fig. 6.10), while the inverse elongates it (bottom-left panel). A $J_z \rightarrow J_r$ transfer, increases the radial pressure, and induces a radial bias in the system, while the opposite transfer enhances the tangential pressure and leads to a negative anisotropy. We show $\beta$ and $\beta_\delta$ in Fig. 6.11 as a function of radius in the plane and along the $\theta = 45^o$ line for these two transfers together with the isotropic control case ($\alpha = (1.0, 1.0, 1.0)$, green lines).

An other way to flatten the stellar distribution is also to impair high $J_z$-orbits and favour high $J_\phi$-orbits (bottom-right panel of Fig. 6.10). The anisotropy estimators for the DF associated with the $J_z \rightarrow (J_\phi > 0)$ transfer are plotted in purple

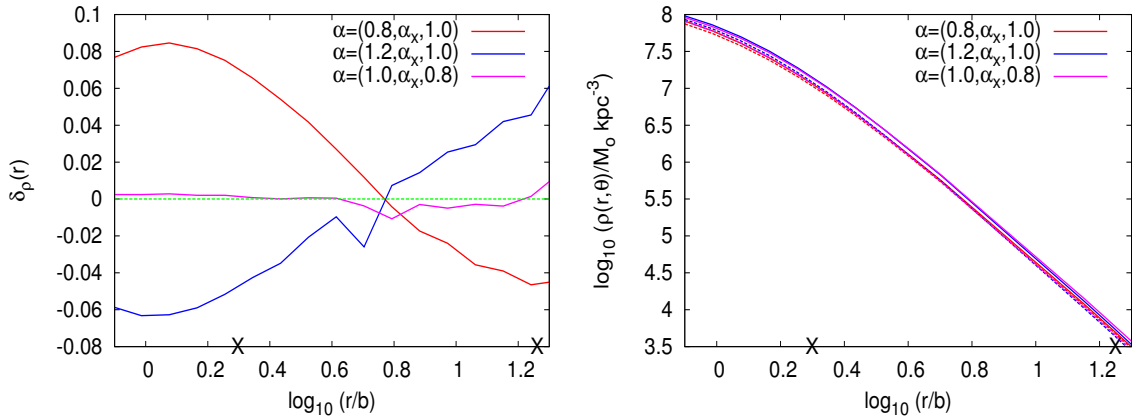


Figure 6.9: Left panel: discrepancy between the zero-th moment of three generalized Isochrone models and the analytic real space density of the Isochrone profile ($\delta_\rho$ is defined in equation 6.17). Right panel: radial density profile of the three models in logarithmic scale; solid lines refer to the profile in the plane and dashed lines to the one along the $\theta = 45^o$ line, though the difference is imperceptible. The two crosses on the x-axis indicate the half-mass and 90%-mass radius.

in Fig. 6.11: breaking the evenness in $L_z$ induces a non-zero streaming motion (see bottom-right panel) and the appearance of an ordered motion causes the azimuthal dispersion to drop, making the system radially biased. It is interesting to note that even though the $J_z \to J_r$ and $J_z \to (J_\phi > 0)$ transfers induce a comparable degree of flattening ($q = 0.9$, q=0.8 respectively), the resultant distribution in velocity space is significantly different (red and purple lines respectively, in Fig. 6.11). In both cases we have impaired the high $J_z$-orbits causing a drop in $\sigma_\theta$: shifting this subtracted probability towards high $J_r$-orbits, increases the radial dispersion leaving the azimuthal one unchanged. The decrease $\sigma_\phi$ experiences for pure geometrical reasons as we move from the plane towards the z-axis, is then enough to make $\beta_\delta$ eventually positive, i.e. $\sigma_\theta > \sigma_\phi$. When the system is set in rotation though, the drop in $\sigma_\phi$ due to the arising of an ordered motion can be even greater than the decrease in $\sigma_\theta$ due to the impairment of high $J_z$-orbits, so that $\beta_\delta$ is systematically positive.

### 6.2.2 Potential bias

Binney & May (1986) showed that due to adiabatic conservation of the actions, when proto-disc material accretes onto a spheroid, if the angular momentum of the material is not aligned with the spheroid's minor axis, the latter will rigidly rotate until the two are parallel. They further showed that the flattening of a stellar system due to the flattening of the underlying potential is more efficient for radially biased systems

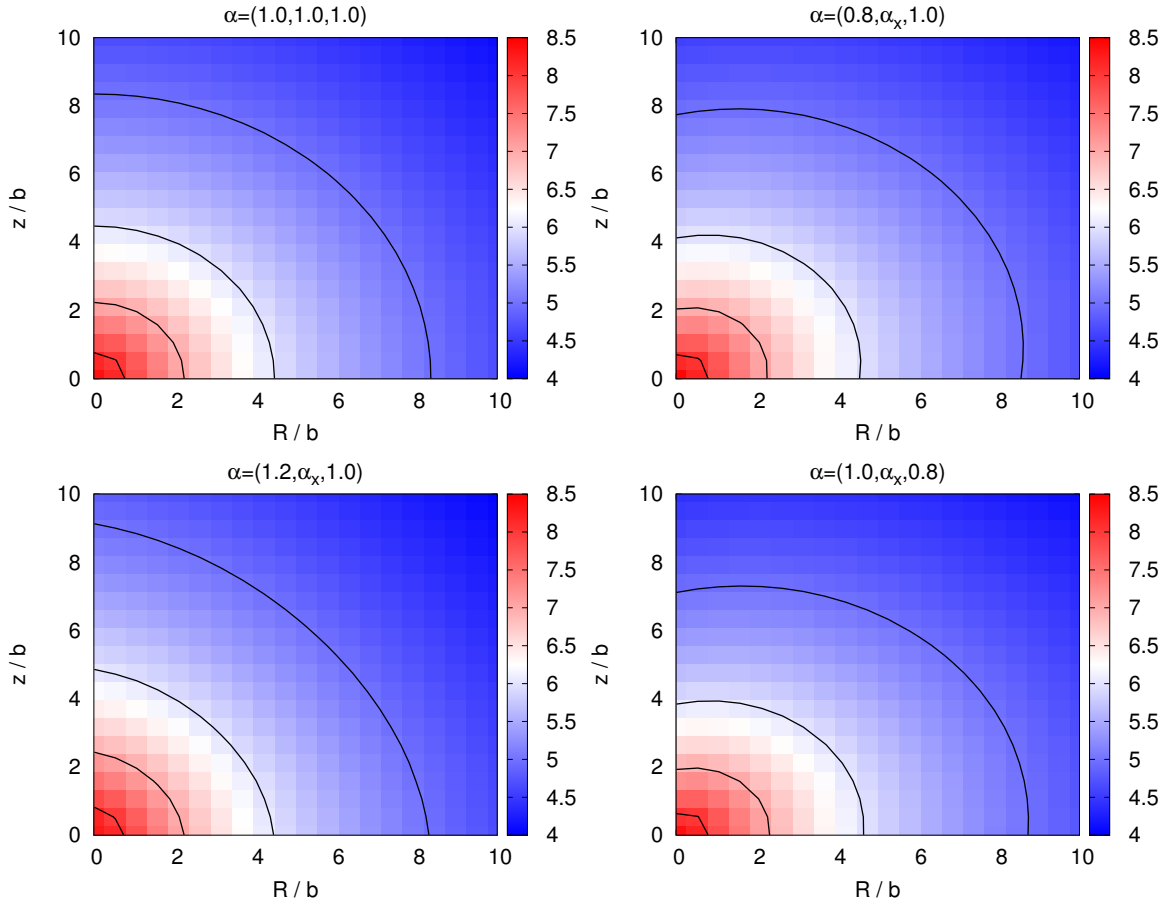


Figure 6.10: Contour-plot of the real-space density for the four generalised Isochrone models of Table 6.2: the control case $\alpha = (1.0, 1.0, 1.0)$ in the top row shows purely the effect of the potential, which is the spherical Isochrone potential of parameters $b = 1.0\,\mathrm{kpc}$, $M = 10^{12} M_\odot$. The density is in units of $M_\odot\,\mathrm{kpc}^{-3}$ and is shown colour-coded in logarithmic scale.

| | $J$ to depress $\rightarrow$ $J$ to enhance | Example $\alpha$ |
|---|---|---|
| **flattening** & | (oblate) $J_z \rightarrow J_r$ | $(0.8, \alpha_x, 1.0)$ |
| **anisotropy** | (prolate) $J_r \rightarrow J_z$ | $(1.2, \alpha_x, 1.0)$ |
| **rotation** | $J_z \rightarrow J_\phi > 0$ (pro-rotation) | $(1.0, \alpha_x, 0.8)$ |
| control case | – | (1.0,1.0,1.0) |

Table 6.2: List of the different probability transfers in action space necessary to produce a particular feature in phase-space. The column on the right shows an example of the weights combination that can induce a given transfer; $\alpha_x$ signals the weight on which condition (6.16) is imposed. The real-space density and anisotropy associated with the generalized DFs defined by an isotropic Isochrone DF and these rescaling parameters, are shown in Fig. 6.10 and 6.11.

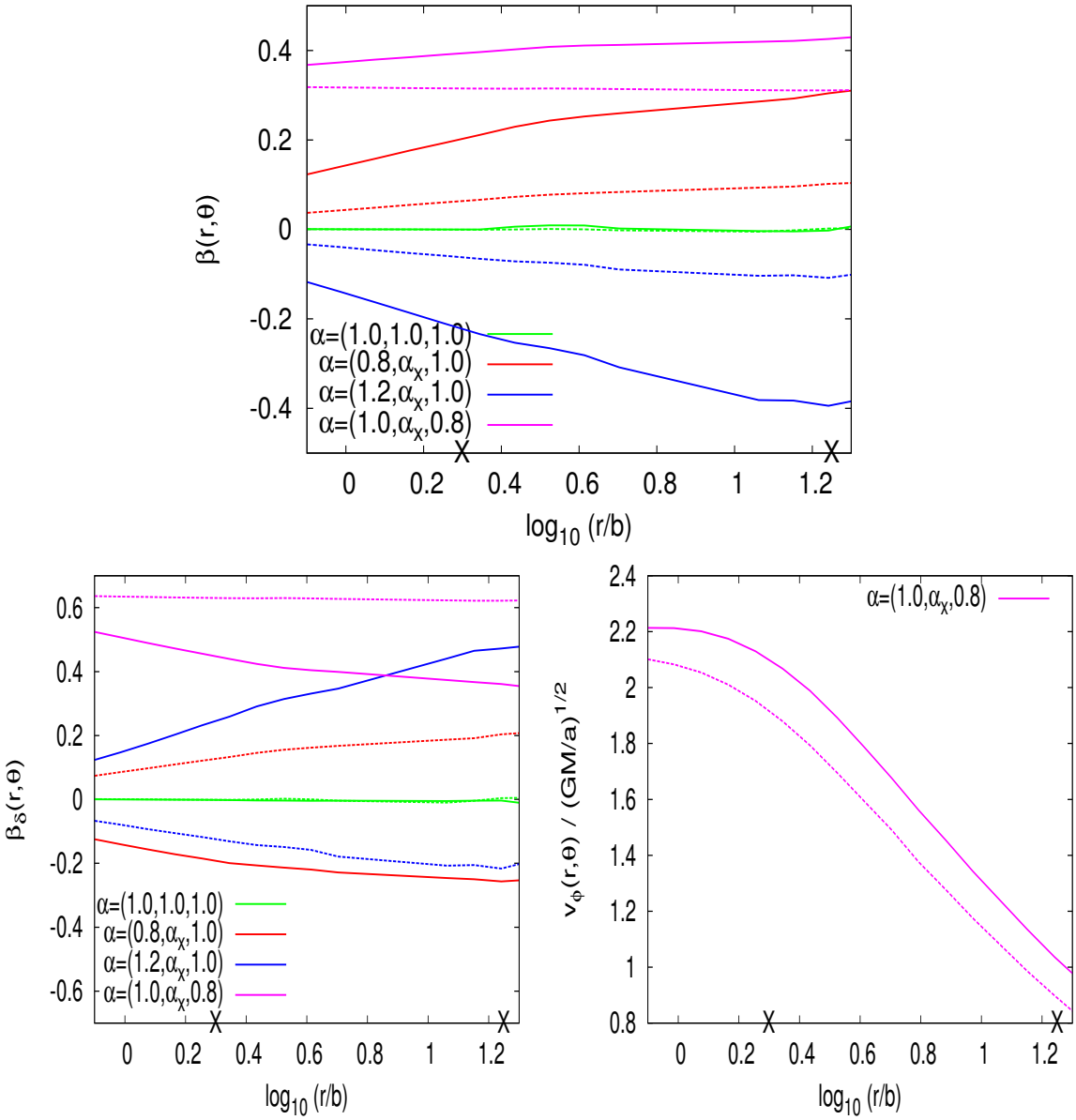


Figure 6.11: Anisotropy and mean-streaming motion as function of radius in the plane ($\theta = 0$, solid lines) and along the $\theta = 45^o$ line (dashed lines) for the four generalised Isochrone DFs of Table 6.2. The green lines (solid and dashed, though almost indistinguishable) refer to the control case $\alpha = (1.0, 1.0, 1.0)$, where we see the effect of the potential only (spherical Isochrone potential, $b = 1.0\,\mathrm{kpc}$, $M = 10^{12} M_\odot$). The two crosses on the x-axis indicate the half-mass and 90%-mass radius.

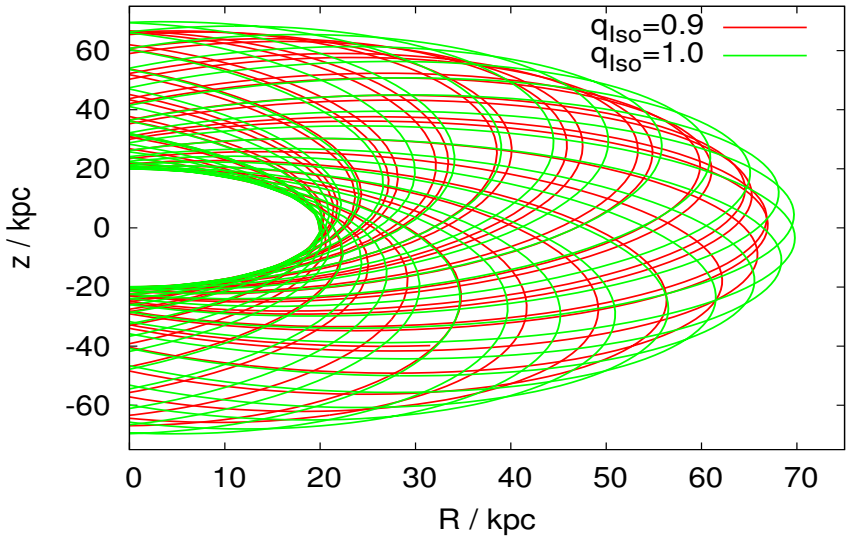


Figure 6.12: Evolution of two $L_z = 0$ orbits in the $(R, z)$ plane of common energy $E = -0.05\,\mathrm{kpc}^2\mathrm{Myr}^{-2} M_\odot$: the green line shows the orbit computed in the spherical Isochrone potential ($b = 1.0\,\mathrm{kpc}$, $M = 10^{12} M_\odot$), while the red line the one in the flattened Isochrone potential ($q = 0.9$).

than systems with an initial isotropic velocity ellipsoid. Here we want to recover this phenomenon from our generalized models and more generally show that our formulation allows a clear understanding of the interactions between the DF and the potential.

When a spherical potential is flattened sufficiently slowly, actions are conserved and consequently orbits are squashed by the potential (see for example Fig. 6.12). A radially anisotropic DF has by construction an enhanced radial pressure so that the two effects constructively superpose. We recompute the four generalised DFs introduced in the previous section, in a Miyamoto-Nagai potential ($a = 1.0\,\mathrm{kpc}$, $b = 0.45\,\mathrm{kpc}$, $M = 10^{12}\,M_\odot$). The top-left panel in Fig 6.13 shows the real-space density contours for the model $\alpha = (1.0, 1.0, 1.0)$, which here serves to trace the effect of the potential alone: the Miyamoto-Nagai potential makes the stellar distribution flattened and more centrally concentrated. When we enhance high $J_r$-orbits via a $J_z \to J_r$ transfer (e.g. model $\alpha = (0.8, \alpha_x, 1.0)$), the associated increase in radial pressure sums up with the flattening of the potential and the stellar distribution results more squashed than for the same DF in the spherical potential or the isotropic DF in the Miyamoto-Nagai potential ($q = 0.7$, $q = 0.8$, $q = 0.9$ respectively and see top-right panel of Fig. 6.13). The opposite J-transfer, namely encouraging high $J_z$-orbits, impairing high $J_r$-ones, induces elongation in the system, but when the DF

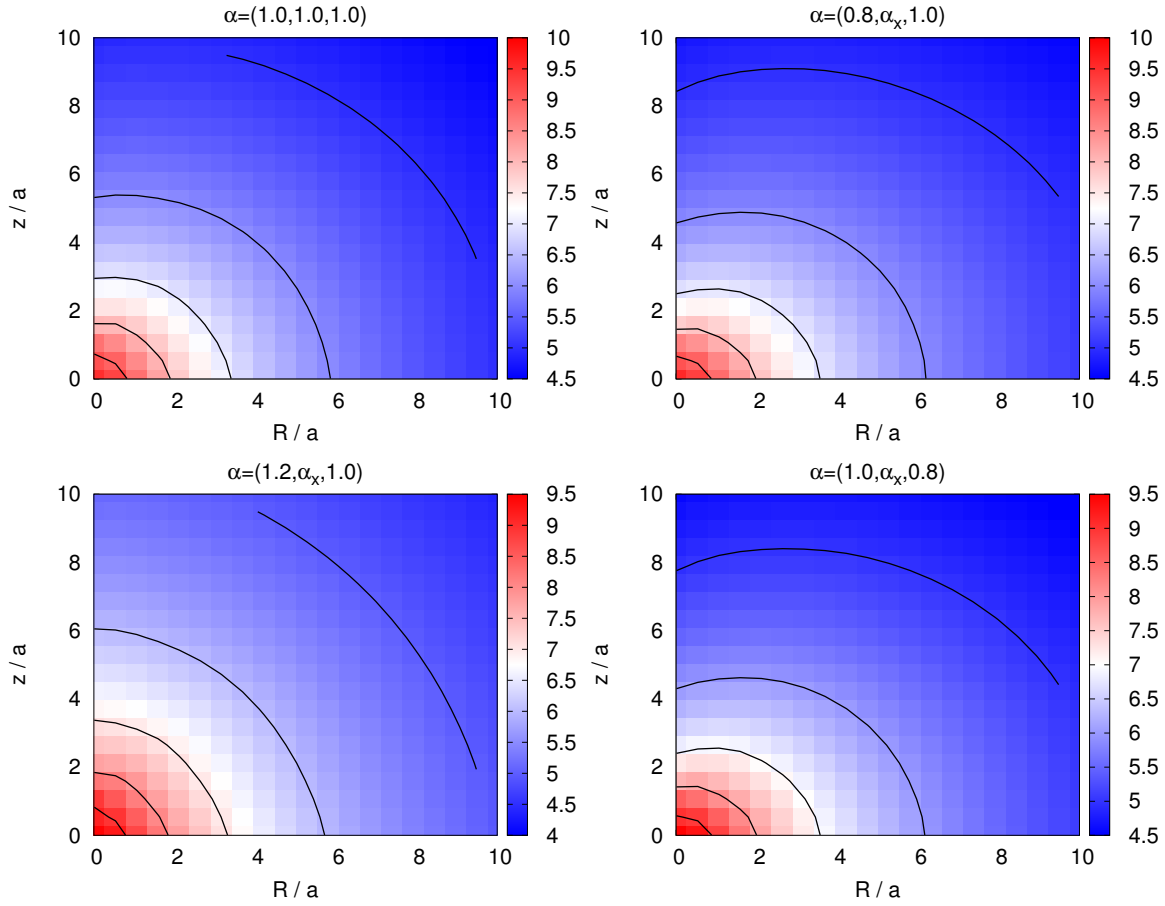


Figure 6.13: Same as Fig. 6.10, but the DFs here were computed in a Miyamoto-Nagai potential of parameters $a = 1.0\,\mathrm{kpc}$, $b = 0.45\,\mathrm{kpc}$, $M = 10^{12} M_{\odot}$.

is evaluated in a flattened potential, the effects cancel and the distribution results almost spherical (bottom-left panel of Fig. 6.13, $q = 1.0$). The bottom-right panel of Fig. 6.13 shows the real space density distribution due to a $J_z \to J_\phi > 0$ probability transfer: the impairment of high $J_z$-orbits favours flattening and the effect is amplified by the fact that the potential is already flattened.

The picture appears rather clean in real-space and it is in velocity space that we can actually perceive in detail the interaction between the probability transfers and the flattening of the potential, as well as the finer structure of the latter. The top panel of Fig. 6.14 shows the anisotropy as a function of radius along two lines of sight (solid lines for $\theta = 0$ and dashed lines for $\theta = \pi/4$). The green lines show the control case where we see the effect of the potential alone: at large radii the flattening of the potential supports high $J_r$-orbits and the system is radially biased; closer to the centre, the density distribution is more puffed up ($b/a = 0.45$), this leaves more space for the high $J_z$-orbits and as a result the distribution is tangentially biased. With a $J_z \to J_r$

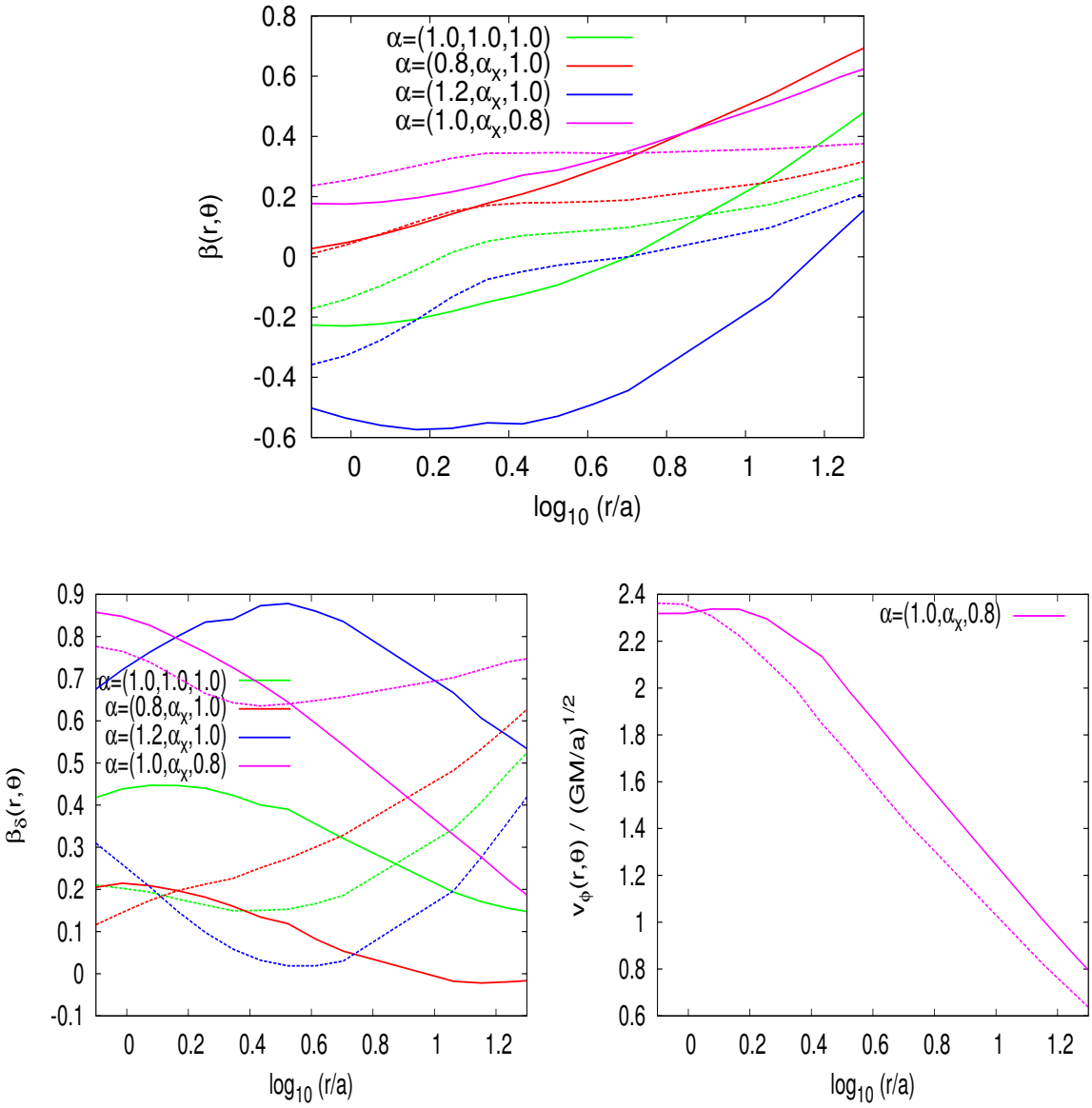


Figure 6.14: Same as Fig. 6.11, but the DFs here were computed in a Miyamoto-Nagai potential of parameters $a = 1.0\,\mathrm{kpc}$, $b = 0.45\,\mathrm{kpc}$, $M = 10^{12} M_{\odot}$.

transfer the radial pressure increases and the above trend just shifts upwards in the $(r, \beta)$ plane (red line, same Figure), while the inverse transfer fosters the tangential pressure making the anisotropy shift systematically towards the negative side (blue line). Still in Fig.6.14, the bottom-left panel shows that $\beta_\delta$ is positive for any model and in particular it is so for the control case, hinting at the fact that the flattening happened in a fashion associable with a $J_\phi \to J_r$ transfer rather than with a $J_z \to J_r$ one. When $J_z$ is impaired by the DF, $\beta_\delta$ drops (red line), while when it is favoured the spread between $\sigma_\phi$ and $\sigma_z$ increases further (blue line). The bottom-right panel of Fig. 6.14 shows the mean-streaming motion generated with the $J_z \to J_\phi > 0$ transfer: the effect of the flattened potential is to steepen its decay with distance.

The $f(\mathbf{J})$ description not only provides a simple framework to understand how a DF interacts with the potential in shaping the morphology and velocity distribution of a stellar system, but also allows to make indirect inferences on the potential given that the same probability transfer (and hence the same DF) will react differently in a different potential. In Fig. 6.15 we plot the anisotropy as a function of radius for the generalized DF associated with parameters $\alpha = (1.0, 1.0, 1.0)$ and $\alpha = (0.8, \alpha_x, 1.0)$ in three different potentials: the spherical Isochrone potential, the flattened Isochrone and the Miyamoto-Nagai potential; the spherical Isochrone (green lines) is taken as the reference case. In the flattened Isochrone, the isotropic DF barely changes w.r.t. the spherical case (top row, red lines), while the $J_z \to J_r$ transfer associated with the second DF appears enhanced, especially at small radii (bottom row, red lines). On the other hand both DFs react in qualitatively the same way when evaluated in a Miyamoto-Nagai potential (blue lines). The idea is therefore that, since we understand how each DF changes the distribution of orbits in phase-space, we can vary its parameters and infer from how the system reacts, the shape of the potential.

In conclusion, we have shown that the observed features of a stellar system can be transparently interpreted as the composition of probability transfers in action space and biases induced by the shape of the potential. The effect of the potential dominates over the intrinsic distribution of the stellar system though: the first step in modelling a stellar system has therefore to be the determination of the underlying potential; thankfully, the state of the art of potential fitting techniques has now reached a consistent solution to the problem (Magorrian 2013).

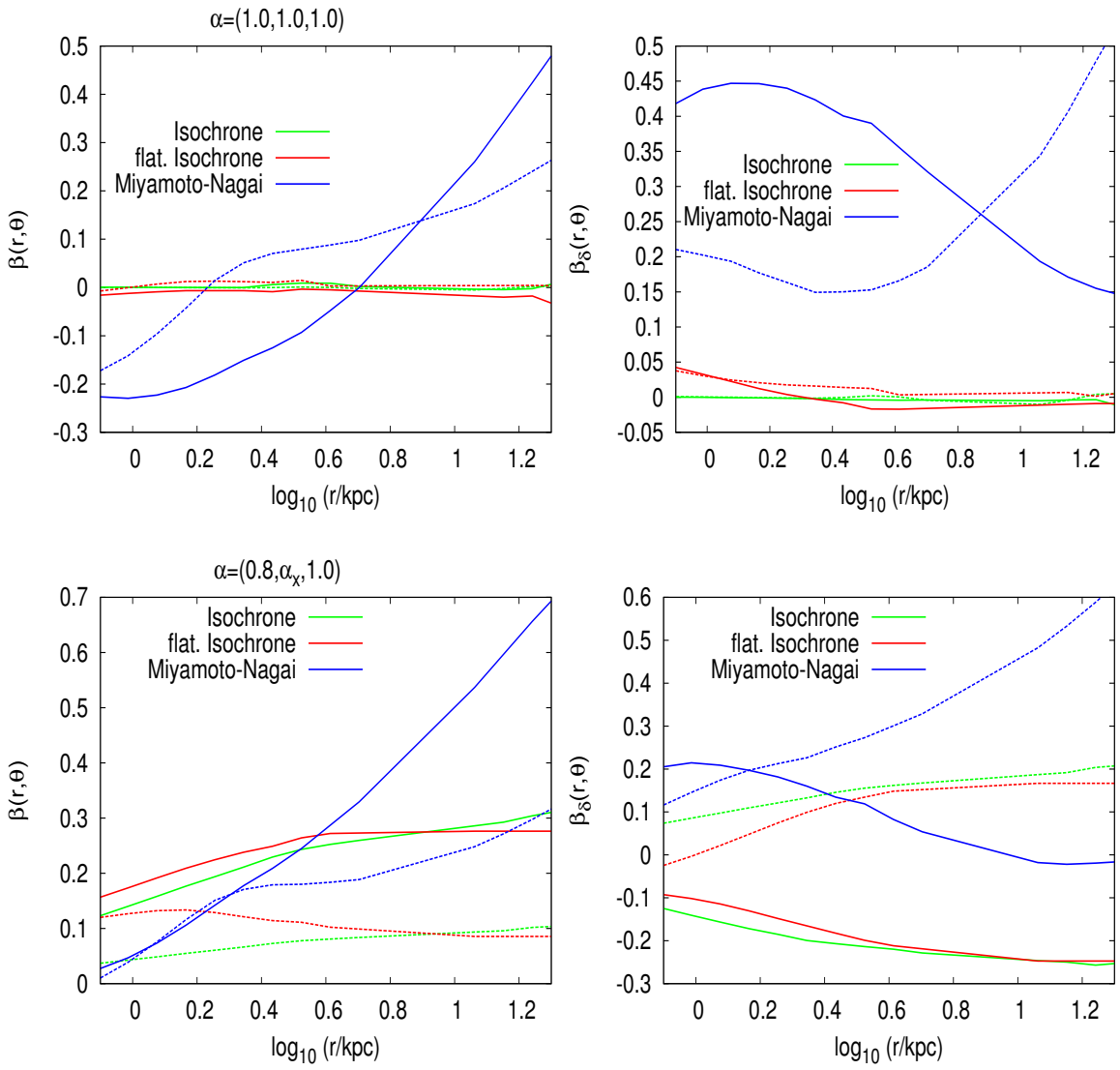


Figure 6.15: Anisotropy as a function of radius along two lines of sight ($\theta = 0$, solid lines and $\theta = \pi/4$, dashed lines) for the isotropic DF and the generalised one of parameters $\alpha = (0.8, \alpha_x, 1.0)$. Colours indicate the different potentials in which a same DF was computed: the spherical Isochrone potential ($b = 1\,\mathrm{kpc}$, $M = 10^{12}\,M_\odot$), green lines; the flattened Isochrone potential ($q = 0.9$), red lines; and the Miyamoto-Nagai potential ($a = 1.0\,\mathrm{kpc}$, $b = 0.45\,\mathrm{kpc}$, $M = 10^{12}\,M_\odot$), blue lines.

## 6.3 Alternative $f(\mathbf{J})$ models for spheroids

Binney & Tremaine (2008), §4.6 propose a procedure that separates the energy dependence in the distribution function from the density of actions on surfaces of constant energy:

$$f(\mathbf{J}) = f_0(E)s(\mathbf{J}). \tag{6.18}$$

$f_0(E)$ is an ergodic DF that controls the real space density. Any departure from the isotropic case is handled by $s(\mathbf{J})$, which *shifts* the density of orbits on surfaces of constant energy and produces the observables required.

As for the generalized models presented in §6.2, we want to keep the number of stars at each energy fixed while varying $s(\mathbf{J})$ to produce the required phase-space features. Eq. (6.10) in this case simplifies to:

$$\int \mathrm{d}^3 J s(\mathbf{J})\delta(E - H) = g(E), \tag{6.19}$$

which, substituting eq. (6.11), becomes

$$\int_{H(\mathbf{J})=E} \frac{\mathrm{d}J_\phi \mathrm{d}J_z}{\Omega_r} = \int_{H(\mathbf{J})=E} \frac{\mathrm{d}J_\phi \mathrm{d}J_z}{\Omega_r} s(\mathbf{J}). \tag{6.20}$$

In the spherical isotropic case $s(\mathbf{J}) = 1$ and the condition is trivially met.

As a result of (6.20), $s(\mathbf{J})$ can change the density distribution of actions on a given surface and make the system flattened (shifts away from $J_z$ axis), radially or tangentially biased (shifts towards and away from the $J_r$ axis respectively) or rotating (oddness in $J_\phi$) without significantly changing the radial profile of the system.

### 6.3.1 A possible shape for the shift function

Pontzen & Governato (2013) propose the following dynamical model for the dark-matter halo of our Galaxy:

$$f(\mathbf{J}) \propto \exp\left(-\vec{\alpha} \cdot \mathbf{J} - \beta_E H(\mathbf{J})\right), \tag{6.21}$$

which they derive via arguments of maximum entropy and imposing constrains on both conservation of mean energy and mean values of the actions. They follow Jaynes & Bretthorst (2003) in considering that *maximising entropy subject to given constraints is equivalent to testing whether those constraints encapsulate the physics of the situation.* They consider their procedure validated by the fact that the model (6.21) is a good fit to three high-resolution dark-matter simulations of varying mass.

We can obtain an $f(\mathbf{J})$-model for a spheroidal component by adopting a generalization of their functional form:

$$f(\mathbf{J}) \propto f_0(H) \exp\left(-\vec{\alpha} \cdot (\mathbf{J}^m(H) - \mathbf{J})\right) \tag{6.22}$$

where $\vec{\alpha} = (\alpha_r, \alpha_z, \alpha_\phi)$, $f_0(H)$ is obtained by Eddington (1916) inversion of the radially averaged density-potential pair and the constant of proportionality is determined by (6.20). Each parameter $\alpha_i$ correlates directly to the main observables by enhancing a different action $J_i$: the rotation of the system will depend on $\alpha_\phi$, its axis-ratio on both $\alpha_z$, $\alpha_\phi$ and the anisotropy will depend on all the three parameters. Note that the only constraint on the shift function is effectively condition (6.20): the choice of an exponential functional form makes it possible both to satisfy such condition and to model a wide range of features.

Even though this scheme is in essence similar to the one proposed in the previous section (phase-space features are still produced by probability shifts in action-space), it comes with the numerical complication of requiring the solution of the integral condition (6.20) for every choice of the model parameters.

## 6.4 Conclusions

We have constructed an approximation of the Hamiltonian as a function of the actions for a general axisymmetric potential by constructing a library of quadratic fits to constant energy surfaces in action space. Remarkably very few (of order of 15) special orbits are needed to constrain each surface and these orbits can be accessed via direct integration. We quantify the accuracy of the overall approximation on two examples (an Isochrone profile and a Miyamoto-Nagai profile) and this turns out to be better than 0.1% in predicting the $E(\mathbf{J})$ and its inverses. Our approximation predicts the frequencies ratios also to a similar accuracy.

Further we presented a method that generalizes a spherical model $f(E)$ to an $f(\mathbf{J})$ model for a spheroidal component in a general axisymmetric potential: we used the new approximation to make explicit the dependence of $E$ on the three actions and consistently rescaled each action to produce phase-space features such as flattening, anisotropy and rotation. We remarked that the flattening of the underlying potential enhances or even induces phase-space features such as flattening and anisotropy. This positive correlation could be used to determine the potential itself as an implicit problem.

We reviewed the alternative strategy of producing $f(\mathbf{J})$ models outlined in §4.6 of Binney & Tremaine (2008). We note that this is in essence similar to the strategy we propose, but requires the additional step of computing a consistency condition for the preservation of the radial profile.

In conclusion, our algorithm generalizes any isotropic spherical model to a more realistic $f(\mathbf{J})$ model by simply writing the Hamiltonian in terms of the actions and then transferring probability in action space. This scheme both offers the flexibility to model realistic stellar systems, but also allow a direct insight of the physics implied by the model.

# Chapter 7

# Conclusions and future work

## 7.1 Overview

In this thesis we have presented a tool-kit for the analysis of the smooth component of the Milky Way stellar halo kinematics and dynamics. We proposed an algorithm to construct and validate the distance calibration of a sample of tracers, critiqued four different methods to estimate the mean-streaming motion and a series of tests to assess the consistency of their measurements, and presented a new class of action-based dynamical models for spheroidal components.

In this last chapter we summarise the methodological and observational results we obtained and trace the next steps in the exploration of the Galaxy formation history and evolution.

### 7.1.1 Methodological results

We have proposed a first algorithm that uses synthetic catalogues to determine the optimal survey size. The idea is that given the discrepancy $\sigma_0$ between the two closest models consistent with the data, the ideal sample size to discriminate between them is the one such that the associated statistical error is $\sim \sigma_0$ as any further increase in the sample size ends up fitting the noise. This simple idea can contribute to optimize the surveys economical costs as well as the computational ones in model-fitting.

Following Schönrich, Binney & Asplund (2012) we tracked distance errors through the correlations they introduce between the velocity components and estimated the mean distance bias on a sample using velocity information only on average. On a set of control cases, we checked that the method measures the correct bias within $(2 \pm 2)\%$. The scheme makes it possible to fit a calibration to a dataset: one has to guess a functional form for the calibration and then varies the parameters that define it until

the average distance error is zero. However, if the data sample is likely to be affected by systematics in the kinematics, it is unwise to let the calibration parameters be decided by the data because their dependence and sensitivity to the data systematics is unknown. All the same, the validation step can still be a viable option for it relies on a smaller information content, which can be still systematics-free.

The mean streaming motion of the stellar halo encodes valuable information on the formation history of the whole Galaxy. Unfortunately its determination is complicated by the difficulty to isolate a complete and uncontaminated set of tracers, by the observations' low signal-to-noise due to the large distances involved and due to the uncertainty in the tracers tangential motion. We presented four different methods to measure rotation: two rely on l.o.s. velocity only and two use the full 3D motion. When only l.o.s. velocity is used it is crucial to account for the spherical geometry of the halo for at large distances the input information (l.o.s. velocity) tends to become singular in relation to the output (the azimuthal motion) we are trying to infer. On the other side, the proper-motions of halo stars are poorly constrained and beyond 20 kpc the signal becomes as large as the uncertainty (see SDSS for example). All the same, we have shown that the average tangential-motion contains valuable information and can be used to infer rotation once all significant systematics have been removed. Because each method manipulates velocity information in a different way, conflicting estimates are symptoms that the assumptions the model or estimator is based upon fail due to non-negligible systematics in the data or inappropriateness of the measuring tool. Moreover, because we are modelling the smooth component of the stellar halo, we expect the rotational signature also to be smooth and in particular, spatially coherent. If either of the consistency requirements is unsatisfied, the inferred average rotation cannot be trusted as it is affected by either a stream-like structure, distance errors or pipe-line systematics.

The two most powerful tools to disentangle the morphology and history of a galaxy are its chemistry and its dynamics. We have presented a new class of dynamical models for spheroidal components that provide a flexible tool to understand halo-like structures. We take any ergodic distribution function and make explicit the dependence of the energy on the three actions and then rescale the weight of each action to induce phase-space features such as anisotropy, flattening and rotation. The Hamiltonian as a function of the actions is obtained by interpolation on a grid of surfaces of constant energy in action space, previously computed and approximated with quadratics. Each surface requires the direct integration of few special orbits, namely of order of 15 points in action space that lie on the edges of the surface,

where at least one of the actions is zero. Remarkably, the quadratic approximation is accurate enough to predict the frequency ratios with an error of $< 0.1\%$ even though the fit happened at the edges and using only actions and no frequency information. The compelling feature of our models is that with just three numbers we can explore a wide variety of possible phase-space distributions and because they are defined in terms of action we have an immediate and intuitive understanding of the physics implied by the model. Thanks to the scheme of Binney (2012a) that computes angle-action coordinates from $(\mathbf{x}, \mathbf{v})$ in any axisymmetric potential, the application to real data is now possible.

### 7.1.2 Observational results

We have selected two samples of Blue Horizontal Branch stars from the Sloan Digital Sky Survey: one using a photo-spectroscopic selection and one kindly provided to us by Xianxiang Xue (see Xue et al. 2011), who used Blamer-line shape measurements to improve the identification of BHB candidates. Looking at theoretical stellar evolutionary models, we constructed a distance calibration for these stars and validated it with the method presented in §3.3. Comparing with previously available calibrations, our formula is the most accurate, with the fractional average distance error consistent with zero. It accounts for both metallicity and colour dependence, hence avoiding any metallicity dependent bias which might be fatal in morphological and kinematic studies. Crucially our calibration is the only one to pass the "sky falling" test of Schönrich, Binney & Asplund (2012). Given the resemblance with the BASTI isochrones, we used our distance calibration to asses these stellar models and conclude that they successfully predict *ugriz* colours for the blue horizontal branch; and in particular they are superior to the Dartmouth isorchones, at least in this colour system.

With the same samples of SDSS BHB stars, but imposing hard geometric cuts to exclude the disk and the Sagittarius stream, we measured the stellar halo rotational signature with the four methods presented in §3.4.1, 3.4.2. We favour a non-rotating stellar halo and in particular observe no rotation gradient either in metallicity or in radius, hence rejecting the double-halo hypothesis. We trace back previously claimed trends of counter-rotation of the more metal-poor SDSS BHB stars (Deason, Belokurov & Evans 2011a) to a contamination of the sample by both non-BHB stars and disk stars. The rotation gradient observed by Kafle et al. (2012) and Beers et al. (2012) on the sample of Xue et al. (2011) is generated by $\sim 500$ hot metal-poor stars whose kinematics we find to be correlated with spectral line indices, while the other

stars in the sample, both metal-poor and metal-rich stars and both stars closer than 15 kpc and further than that, show no dichotomy in their rotational signature. We also flag the analysis of Hattori et al. (2013) as the neglecting the geometrical effects associated with deprojecting the l.o.s. velocity to infer the dispersion tensor in a spherical system.

## 7.2 Road Ahead

The kinematics of the smooth component of the stellar halo proves to be a controversial topic of debate. We made our best attempt coherently to trace the rotation profile of this component and future analyses shall falsify our picture with larger and more accurate data samples than the ones available to date. Good candidates for this cross-check will be RR Lyrae stars, which are bright standard candles and whose catalogues are steadily increasing in size. Further, multi-epoch surveys and cross-matching among different catalogues have started to make it possible to obtain proper motion for field halo stars with uncertainties below $< 0.1\,\mathrm{masyr}^{-1}$ (e.g. Sohn et al. 2013), opening to a direct probe of the 3D velocity profile of the stellar halo. Different data samples and accurate tangential motions will hopefully also identify the nature of the $\sim 500$ hot metal-poor stars in the sample of Xue et al. (2011), whose kinematics correlates with line indices.

A rich content of information on the galaxy history lies in the debris of accretion events that survive as substructures in the halo. Relaxation and tidal-stripping make it difficult cleanly to identify them in observable space and chemical and dynamical tagging seems to be the key to disentangle the hank (Freeman & Bland-Hawthorn 2002; Binney 2010). In this thesis we have provided a new class of dynamical models for the smooth component of the stellar halo, which will prove useful to set the background to subtract when searching for structures in phase-space. The ultimate fitting of our models to the data, awaits a large enough catalogue of field halo stars with full 6D phase-space information though. Meanwhile, we outline below a general method to fit a self-consistent $f(\mathbf{J})$ model and propose some other applications and extensions of the schemes introduced in this work, further to reinforce their usefulness.

### 7.2.1 Observations modelling

The availability of a class of analytical and physically meaningful models for spheroids allows a transparent, fully dynamical and intuitive **interpretation of stellar halo data**. Once the potential has been determined (Magorrian 2013), the distribution

function of the stellar system can be determined with a straight-forward three-step algorithm:

1. Given the potential $\Phi(R, z)$, construct the approximation $H(\mathbf{J}) \simeq \tilde{E}(E, \mathbf{J})$ (§6.1).

2. Calculate the monopole of both the potential and the space density of the tracer population and by Eddington (1916) inversion determine $f_0(E)$ .

3. Fit the parameters of the generalized models (§6.2, 6.3) by finding the maximum likelihood of the data given the model.

Once a model of the stellar component is at hand, one can iterate the original guess for the **dark matter profile** until consistency with the potential is achieved:

1. Given the DF for the stellar component, choose an initial DF for the dark matter (e.g. Pontzen & Governato 2013).

2. Compute the real space density associated with the composition of the two DFs and find the associated potential via Poisson's equation.

3. Compare with the actual potential (Magorrian 2013) and adjust the dark matter DF if the discrepancy exceeds the acceptance threshold. Iterate until self-consistency.

The potential together with the global DF constitute a complete equilibrium model for the galaxy under study. Even though the presence of streams in galaxy halos is clear evidence that the system has yet to reach equilibrium, an equilibrium model is the best initial guess for the realistic treatment of the galaxy dynamics. The unrelaxed nature of a given substructure can be taken into account by considering time dependence in the Boltzmann equation, while the distortions due to encounters and accretion events can be tackled via the Fokker-Plank diffusion theory (§7.4, Binney & Tremaine 2008). Then the **evolution history of the Galaxy can be fully disentangled**.

### 7.2.2 Methodology

We have presented a scheme to approximate the Hamiltonian as a function of the actions for an axisymmetric potential, but what if the **potential** is **triaxial** (e.g. Law & Majewski 2010)? It is natural to expect that surfaces of constant energy in

action space are smooth also for more general potentials: then the scheme presented in §6.1 can be applied slavishly once the minimum amount of orbits necessary to constrain a surface of constant energy in such a potential has been redetermined (as well as possibly the degree of the polynomial fit to the surface).

From the knowledge of $H(\mathbf{J})$, we can reconstruct the physical problem $(\vec{x}, \vec{v})$ modulus a coordinate transformation that links angle-action space to phase-space canonical coordinates. The generating function of such transformation is the solution of the Hamilton-Jacobi equation, which in general is difficult if not impossible to solve explicitly (§D.4.6, Binney & Tremaine 2008, or §3.5 for an example of explicit computation). An alternative approach might be to derive **the generating function for a general potential** from the one of a fully worked-out example. The idea is to construct a transformation from the Hamiltonian as function of the actions of a general potential to the Hamiltonian relative to a potential with known generating function (e.g. the one relative to a spherical potential).[1] The generating function associated with the general potential is then the composition of such a transformation with the generating function of the known case. Note that the procedure does not require the Hamiltonian or the generating function to be separable, which is otherwise a necessary yet not trivial condition one has to impose to manipulate the two functions. The idea is essentially the one behind the torus machine (Kaasalainen & Binney 1994), but the form in which we approximate $H(\mathbf{J})$, now opens to the possibility of writing this mapping in a closed form.

[1] A geometrically intuitive way to construct the transformation between two Hamiltonians is again to define it at the level of surfaces of constant energy. One can first rescale the surfaces so that their corners match; then each point of the second surface is mapped onto the intercept between the normal to the second surface at that point and the first surface. For surfaces with constant curvature sign, this is a invertible mapping.

# Appendix A

# SQL query for SDSS

The following SQL query was used to drawn data from the SDSS archive:

```
SELECT
sp.PLATE,sp.FIBERID,sp.specObjID,ph.rerun,
,sp.TARGETSTRING,sp.mjd,sp.flag,sp.fehadop,sp.loggadop,
sp.fehadopunc,sp.FEHWBG,sp.TEFFWBG,sp.LOGGWBG,
ph.psfmag_u,ph.extinction_u,ph.psfmag_g,ph.extinction_g,
ph.psfmag_r,ph.extinction_r,ph.psfmag_i,ph.extinction_i,
ph.psfmag_z,ph.extinction_z,ph.l,ph.b,ph.ra,ph.dec,
m.pmL,m.pmB,m.pmRa,m.pmDec,m.pmRaErr,m.pmDecErr,
m.sigRa,m.sigDec,m.nFit,m.dist22,m.match,
sp.ELODIERVFINAL,sp.ELODIERVFINALERR,sp.SNR,
sp.TEFFADOP,sp.TEFFADOPUNC
FROM sppParams sp
JOIN PhotoObjAll ph on sp.bestobjid = ph.objid
JOIN ProperMotions m on m.objid = ph.objid
WHERE sp.sciencePrimary = 1
AND ph.CLEAN=1
AND sp.TARGETSTRING = 'BHB'
```

# Appendix B

# The Sagittarius Stream

The masking of Sgr is performed in right ascension-declination space $(\alpha, \delta)$ according to the following polygon kindly provided by the authors of Deason, Belokurov & Evans (2011a) on request:

$$
\begin{aligned}
\alpha_1 &= [245, 225, 211.5, 210, 208, 200.5, 199.3, 197.5, 197, \\
&\quad 180, 160, 140, 110, 113, 140, 160, 180, 202, 225, 245], \\
\delta_1 &= [-3, 8, 14.2, 12.5, 16, 19.5, 16.5, 16.9, 21.5, 28, 33, \\
&\quad 35, 38, 15, 13.5, 11, 6, -3, -3, -3]; \\
\cup\, \alpha_2 &= [10, 50, 70, 30, 10],\ \delta_2 = [-10, 20, 10, -20, -10].
\end{aligned}
$$

# Appendix C

# Necessary condition for the $L^{-2\beta}E^s$ DF's consistency

**Theorem:** Let the sample of stars be $X = \{x_{(1)}, \ldots, x_{(N)}\}$ divided in the sub-samples $X = X_1 \cup X^1 = \{x_1, \ldots, x_m\} \cup \{x^1, \ldots, x^n\}$. If $P(\eta|X)$ is the probability of the model relative to the parameter $\eta$ given the data and $\eta_{best,X}$ is the value that maximize it for the type of distribution function under consideration:

$$\eta_{best,X} = \frac{card(X_1)}{card(X)}\eta_{best,X_1} + \frac{card(X^1)}{card(X)}\eta_{best,X^1}.$$

**Proof:**

Since $f \propto 1 + (1-\eta)\tanh(\frac{L_z}{\Delta})$, assuming $\Delta \ll L_z$ and at fixed $\beta$ (such that the normalization constant, independent of $\eta$, is $n_c$):

$$P(\eta|X) = \prod_{i=1}^{N} P(\eta|x_{(i)}) = \frac{\prod[1 + (1-\eta)\tilde{H}(x_{(i)})]}{n_c^N}$$

where $\tilde{H}(x) = 2H(x) - 1$ and $H(x)$ is the Heaviside function. Then one can split the above in the positive and negative contributions, obtaining:

$$P(\eta|X) = \prod_{i=1}^{N^-} \frac{\eta}{n_c} \cdot \prod_{i=1}^{N^+} \frac{2-\eta}{n_c}$$

where $N^+$ is the number of stars with $L_z > 0$ and the opposite for $N^-$. The probability is maximum for $\frac{\mathrm{d}P(\eta|X)}{\mathrm{d}\eta} = 0$, which happen for $1 - \eta = \frac{N^+ - N^-}{N}$. The same holds for the two sub-samples of $X$: calling $m^+$ the number of stars with positive $L_z$ that belong to $X_1$ and $n^+$ the ones belonging to $X^1$ and similarly for stars with negative $L_z$, we have:

$$\eta_{best,X_1} = \frac{2m^-}{m} \qquad \text{and} \qquad \eta_{best,X^1} = \frac{2n^-}{n}$$

with $N^+ = m^+ + n^+$ and $N^- = m^- + n^-$. We conclude, showing that $\eta_{best,X}$ can be expressed as follows:

$$\eta_{best,X} = \frac{2(m^- + n^-)}{N} = \frac{2m^-}{m} \cdot \frac{m}{N} + \frac{2n^-}{n} \cdot \frac{n}{N}$$

This completes the proof.

# Appendix D

# Floquet Analysis

Computing the azimuthal frequency of a radial orbit is a well defined problem only in the limit $L \to 0$ or similarly the radial frequency of a shell orbit makes sense only in the limit $J_\phi \to 0$. This means that while the other frequencies can be computed by direct integration of the orbit in question, the third frequency needs to be tackled by means of perturbation theory. To extract the frequency of a perturbed orbit in the direction of the perturbation we can use Floquet analysis.

Floquet Theorem states that the differential equation

$$\ddot{x}(t) = -K(t)x(t) \tag{D.1}$$

with $K(t)$ periodic of period $T$ has a complete set of solutions of the form:

$$x(t) = \mathrm{Re}\left(Ae^{i\nu t}P(t)\right) \tag{D.2}$$

with $P(t)$ periodic of period $T$ and $A \in \mathbb{C}$. We can write the problem of extracting the frequency in the direction $\zeta$ of a perturbed orbit in a form similar to (D.1) to exploit this result. Given the Hamiltonian $H_{\text{eff}} = 1/2(p_R^2 + p_z^2) + \Phi_{\text{eff}}(R, z)$, the equations of motions are:

$$\dot{\zeta} = \frac{\partial H_{\text{eff}}}{\partial p_\zeta} = p_\zeta \ ; \qquad \dot{p_\zeta} = -\frac{\partial H_{\text{eff}}}{\partial \zeta} \tag{D.3}$$

Perturbing the orbit along $\zeta$,[1] $\zeta = \zeta_0 + \Delta$ (and therefore $\dot{\zeta_0} + \dot{\Delta} = p_\zeta$), the second equation becomes

$$\dot{p_\zeta} = -\frac{\partial \Phi_{\text{eff}}}{\partial \zeta} = -\frac{\partial \Phi_{\text{eff}}}{\partial \zeta}(\zeta_0) - \frac{\partial^2 \Phi_{\text{eff}}}{\partial \zeta^2}(\zeta_0)\Delta. \tag{D.4}$$

[1]For a perturbation of a radial orbit $\mathbf{J} = (J_r, 0, 0)$, $\zeta = z$ as the perturbation is in the vertical direction. Vice-versa for a shell orbit $\mathbf{J} = (0, J_z, 0)$, the perturbation is radial and $\zeta = R$.

Given that $\ddot{\zeta}_0 + \ddot{\Delta} = \dot{p_\zeta}$, we have that the perturbation evolves according to the following equation:

$$\ddot{\Delta} = -\frac{\partial^2 \Phi_{\text{eff}}}{\partial \zeta^2} \Delta \tag{D.5}$$

where $\partial^2 \Phi_{\text{eff}} / \partial \zeta^2$ is periodic of period $T_\zeta$. Following Floquet's Theorem the solution will have the form (D.2) and $\Omega_\zeta = \nu$.

In practise to extract $\nu$ we map a set of initial conditions at time $t$ into their evolution after a period $T_\zeta$: this mapping depends on $\nu$, but can be easily computed via orbit integration of the equation of motion (D.5), with $\Phi_{\text{eff}}$ computed along the unperturbed orbit. Equation D.1 implies that if $\Delta(t)$ is a solution then $\Delta(t+T)$ is also a solution, which in general in not true. We can write (D.1) as follow:

$$\frac{\mathrm{d}\vec{\Delta}}{\mathrm{d}t} = A(t)\vec{\Delta}, \quad \vec{\Delta} := (\Delta, \dot{\Delta}) \text{ and } A(t) := \begin{pmatrix} 0 & 1 \\ -K(t) & 0 \end{pmatrix} \tag{D.6}$$

where $A(t)$ is still periodic of period $T_\zeta$. We define $X(t)$ to be the fundamental solution matrix, where each of its columns is a linearly independent solution of the linearised differential problem. The columns of $X(t)$ form therefore a basis for the n-dimensional solution space of (D.6) and any other fundamental solution matrix $\tilde{X}(t)$ may be written as $\tilde{X}(t) = MX(t)$, where $M$ is a $n \times n$ non-singular constant matrix (commonly referred to as monodromy matrix). In our case n=2. In particular $X(t + T_\zeta)$ is a fundamental solution of (D.6) and can be written as

$$X(t + T_\zeta) = MX(t). \tag{D.7}$$

The eigenvalues $(\rho_j)_{j=1,\ldots,n}$ of $M$ are called Floquet multipliers and are non-zero. More important, they are the same whatever the choice of the fundamental solution $X(t)$ as the monodromy matrices associated with each fundamental solution are all similar to each other. To ease the notation we can imagine $M$ to be diagonal: then (D.7) implies

$$x_{i,j}(t + T_\zeta) = \rho_j x_{i,j}(t).$$

A possible solution is of the form $x_{i,j}(t) = \rho_j^{kt} p_{i,j}(t)$ with $k = 1/T_\zeta$ and $p_{i,j}(t)$ a periodic function of period $T_\zeta$. If we compare this solution with (D.2), we get that:

$$\nu_j = \frac{1}{T} \ln \rho_j \tag{D.8}$$

A consequence of Liouville's Theorem implies $\sum_j \nu_j = 0$, see e.g. Bonetto (lecture notes). Therefore the solution of (D.1) is reduced to extract the eigenvalues of M. To determine it, we can choose without loss of generality, each column of $X(t)$ to be a

solution of (D.6) with initial condition $\mathbf{e}_i$ (a column vector with all components zero except the $i-th$ one). Therefore $X(0) = \mathbb{I}_n$ is the identity matrix and $M = X(T_\zeta)$. The method outlined above is quite common in literature: see, for example, Rand (lecture notes).

# Bibliography


Abadi M. G., Navarro J. F., Steinmetz M., Eke V. R., 2003, ApJ, 597, 21

Abazajian K. N. et al., 2009, ApJ Special, 182, 543

Adelman-McCarthy J. K. et al., 2008, ApJ Special, 175, 297

Ahn C. P. et al., 2012, ApJ Special, 203, 21

Aihara H. et al., 2011a, ApJ Special, 193, 29

Aihara H. et al., 2011b, ApJ Special, 195, 26

Allende Prieto C. et al., 2008, AJ, 136, 2070

An D. et al., 2008, ApJ Special, 179, 326

Armandroff T. E., 1989, AJ, 97, 375

Athanassoula E., 2002, in Astronomical Society of the Pacific Conference Series, Vol. 275, Disks of Galaxies: Kinematics, Dynamics and Peturbations, Athanassoula E., Bosma A., Mujica R., eds., pp. 141–152

Athanassoula E., Romero-Gómez M., Masdemont J. J., 2009, MNRAS, 394, 67

Baev P. V., Markov H., Spassova N., 2001, MNRAS, 328, 944

Bahcall J. N., Soneira R. M., 1980, ApJ Special, 44, 73

Barbuy B. et al., 2008, Physica Scripta Volume T, 133, 014032

Baugh C. M., Cole S., Frenk C. S., 1996, MNRAS, 283, 1361

Beers T. C. et al., 2012, ApJ, 746, 34

Beers T. C., Chiba M., Yoshii Y., Platais I., Hanson R. B., Fuchs B., Rossi S., 2000, AJ, 119, 2866

Bell E. F. et al., 2008, ApJ, 680, 295

Belokurov V. et al., 2006, ApJ Letters, 642, L137

Bertin G., Lin C. C., 1996, Spiral structure in galaxies a density wave theory

Binney J., 2010, MNRAS, 401, 2318

Binney J., 2012a, MNRAS, 426, 1324

Binney J., 2012b, MNRAS, 426, 1328

Binney J., May A., 1986, MNRAS, 218, 743

Binney J., McMillan P., 2011, MNRAS, 413, 1889

Binney J., Merrifield M., 1998, Galactic Astronomy

Binney J., Spergel D., 1984, MNRAS, 206, 159

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

Blitz L., Spergel D. N., 1991, ApJ, 379, 631

Bonetto F., lecture notes, Lecture notes on *Ordinary Differential Equations*

Bovy J., Rix H.-W., Hogg D. W., Beers T. C., Lee Y. S., Zhang L., 2012, ApJ, 755, 115

Brown T. M., Sweigart A. V., Lanz T., Landsman W. B., Hubeny I., 2001, ApJ, 562, 368

Bullock J. S., Johnston K. V., 2005, ApJ, 635, 931

Buonanno R., 1993, in Astronomical Society of the Pacific Conference Series, Vol. 48, The Globular Cluster-Galaxy Connection, Smith G. H., Brodie J. P., eds., p. 131

Byrd G. G., Saarinen S., Valtonen M. J., 1986, MNRAS, 220, 619

Caldwell J. A. R., Ostriker J. P., 1981, ApJ, 251, 61

Cappellari M., 2008, MNRAS, 390, 71

Cappellari M. et al., 2011, MNRAS, 413, 813

Carollo D. et al., 2010, ApJ, 712, 692

Carollo D. et al., 2007, Nature, 450, 1020

Chalonge D., Divan L., 1973, A&A, 23, 69

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

Chiba M., Beers T. C., 2000, AJ, 119, 2843

Clewley L., Warren S. J., Hewett P. C., Norris J. E., Peterson R. C., Evans N. W., 2002, MNRAS, 337, 87

Cole S., Hatton S., Weinberg D. H., Frenk C. S., 1998, MNRAS, 300, 945

Combes F., Sanders R. H., 1981, A&A, 96, 164

Contopoulos G., 1963, AJ, 68, 1

Cooper A. P. et al., 2010, MNRAS, 406, 744

Courteau S., de Jong R. S., Broeils A. H., 1996, ApJ Letters, 457, L73

Davies R. L. et al., 2001, ApJ Letters, 548, L33

D'Cruz N. L. et al., 2000, ApJ, 530, 352

De Propris R., Harrison C. D., Mares P. J., 2010, ApJ, 719, 1582

de Vaucouleurs G., 1977, AJ, 82, 456

Deason A. J., Belokurov V., Evans N. W., 2011a, MNRAS, 411, 1480

Deason A. J., Belokurov V., Evans N. W., 2011b, MNRAS, 416, 2903

Dorman B., Rood R. T., O'Connell R. W., 1993, ApJ, 419, 596

Dotter A., Chaboyer B., Jevremović D., Baron E., Ferguson J. W., Sarajedini A., Anderson J., 2007, AJ, 134, 376

Dotter A., Chaboyer B., Jevremović D., Kostov V., Baron E., Ferguson J. W., 2008, ApJ Special, 178, 89

Eddington A. S., 1916, MNRAS, 76, 572

Eggen O. J., Lynden-Bell D., Sandage A. R., 1962, ApJ, 136, 748

Eisenstein D. J. et al., 2011, AJ, 142, 72

Emsellem E., Monnet G., Bacon R., 1994, A&A, 285, 723

Evans N. W., Hafner R. M., de Zeeuw P. T., 1997, MNRAS, 286, 315

Eyre A., Binney J., 2011, MNRAS, 413, 1852

Faber S. M., Gallagher J. S., 1979, Ann. Rev. A&A, 17, 135

Faulkner J., 1967, Nature, 215, 44

Fermani F., Schönrich R., 2013a, MNRAS, 430, 1294

Fermani F., Schönrich R., 2013b, MNRAS, 432, 2402

Freeman K., 2012, Structure and Evolution of the Milky Way, Miglio A., Montalbán J., Noels A., eds., p. 137

Freeman K., Bland-Hawthorn J., 2002, Ann. Rev. A&A, 40, 487

Frenk C. S., White S. D. M., 1980, MNRAS, 193, 295

Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, AJ, 111, 1748

Gebhardt K. et al., 2000, AJ, 119, 1157

Gilmore G., Reid N., 1983, MNRAS, 202, 1025

Girolami M., Calderhead B., 2011, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 73, 123

Gratton R. G., Fusi Pecci F., Carretta E., Clementini G., Corsi C. E., Lattanzi M., 1997, ApJ, 491, 749

Grillmair C. J., 2006, ApJ Letters, 645, L37

Han Z., 2008, A&A, 484, L31

Hartwick F. D. A., 1987, in NATO ASIC Proc. 207: The Galaxy, Gilmore G., Carswell B., eds., pp. 281–290

Hattori K., Yoshii Y., Beers T. C., Carollo D., Lee Y. S., 2013, ApJ Letters, 763, L17

Heavens A., 2009, ArXiv e-prints

Helmi A., 2004, ApJ Letters, 610, L97

Helmi A., White S., 2000, in Astronomische Gesellschaft Meeting Abstracts, Vol. 16, Astronomische Gesellschaft Meeting Abstracts, Schielicke R. E., ed., pp. 20–24

Helmi A., White S. D. M., de Zeeuw P. T., Zhao H., 1999, Nature, 402, 53

Howard C. D. et al., 2009, ApJ Letters, 702, L153

Hunter C., Qian E., 1993, MNRAS, 262, 401

Ibata R., Lewis G. F., Irwin M., Totten E., Quinn T., 2001, ApJ, 551, 294

Ivezić **v**Z., Beers T. C., Jurić M., 2012, Ann. Rev. A&A, 50, 251

Jaynes E. T., Bretthorst G. L., 2003, Probability Theory

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

Johnston K. V., Law D. R., Majewski S. R., 2005, ApJ, 619, 800

Johnston K. V., Majewski S. R., Siegel M. H., Reid I. N., Kunkel W. E., 1999, AJ, 118, 1719

Jurić M. et al., 2008, ApJ, 673, 864

Kaasalainen M., Binney J., 1994, MNRAS, 268, 1033

Kaczmarczik M. C., Richards G. T., Mehta S. S., Schlegel D. J., 2009, AJ, 138, 19

Kafle P. R., Sharma S., Lewis G. F., Bland-Hawthorn J., 2012, ApJ, 761, 98

Kapteyn J. C., 1922, ApJ, 55, 302

Kauffmann G., Guiderdoni B., White S. D. M., 1994, MNRAS, 267, 981

Kinman T. D., 1959, MNRAS, 119, 559

Kinman T. D., 1995, in IAU Symposium, Vol. 164, Stellar Populations, van der Kruit P. C., Gilmore G., eds., p. 75

Kinman T. D., Cacciari C., Bragaglia A., Buzzoni A., Spagna A., 2007, MNRAS, 375, 1381

Kinman T. D., Cacciari C., Bragaglia A., Smart R., Spagna A., 2012, MNRAS, 422, 2116

Klement R. J., 2010, A&A Rev., 18, 567

Kormendy J., 1993, in IAU Symposium, Vol. 153, Galactic Bulges, Dejonghe H., Habing H. J., eds., p. 209

Law D. R., Majewski S. R., 2010, ApJ, 714, 229

Law D. R., Majewski S. R., Johnston K. V., 2009, ApJ Letters, 703, L67

Law D. R., Majewski S. R., Skrutskie M. F., Johnston K. V., 2004, in Astronomical Society of the Pacific Conference Series, Vol. 327, Satellites and Tidal Streams, Prada F., Martinez Delgado D., Mahoney T. J., eds., p. 239

Layden A. C., Hanson R. B., Hawley S. L., Klemola A. R., Hanley C. J., 1996, AJ, 112, 2110

Lee Y. S. et al., 2011, ApJ, 738, 187

Lee Y. S. et al., 2008a, AJ, 136, 2022

Lee Y. S. et al., 2008b, AJ, 136, 2050

Lindblad B., 1927, MNRAS, 87, 553

Liszt H. S., Burton W. B., 1980, ApJ, 236, 779

Lynden-Bell D., 1962, MNRAS, 123, 447

Magorrian J., 2006, MNRAS, 373, 425

Magorrian J., 2013, submitted to MNRAS

Majewski S. R., 1992, ApJ Special, 78, 87

Majewski S. R., Skrutskie M. F., Weinberg M. D., Ostheimer J. C., 2003, ApJ, 599, 1082

Maller A. H., 2005, in IAU Colloq. 199: Probing Galaxies through Quasar Absorption Lines, Williams P., Shu C.-G., Menard B., eds., pp. 237–242

May A., Binney J., 1986, MNRAS, 221, 857

McMillan P. J., Binney J., 2012, MNRAS, 419, 2251

McMillan P. J., Binney J. J., 2008, MNRAS, 390, 429

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

Munn J. A. et al., 2004, AJ, 127, 3034

Murante G., Poglio E., Curir A., Villalobos A., 2010, ApJ Letters, 716, L115

Ness M. et al., 2013, ArXiv e-prints

Newberg H. J. et al., 2002, ApJ, 569, 245

Newberg H. J., Yanny B., Willett B. A., 2009, ApJ Letters, 700, L61

Niederste-Ostholt M., Belokurov V., Evans N. W., Peñarrubia J., 2010, ApJ, 712, 516

Nordström B. et al., 2004, A&A, 418, 989

Norris J. E., Ryan S. G., 1989, ApJ Letters, 336, L17

Ollongren A., 1965, Ann. Rev. A&A, 3, 113

Oort J. H., 1927, Bull. Astron. Inst. Netherlands, 3, 275

Oort J. H., Kerr F. J., Westerhout G., 1958, MNRAS, 118, 379

Oort J. H., Plaut L., 1975, A&A, 41, 71

Philip A. G. D., Adelman S. J., 1993, in Databases for Galactic Structure, Philip A. G. D., Hauck B., Upgren A. R., eds., pp. 245–248

Pietrinferni A., Cassisi S., Salaris M., Castelli F., 2004, ApJ, 612, 168

Pietrinferni A., Cassisi S., Salaris M., Castelli F., 2006, ApJ, 642, 797

Pontzen A., Governato F., 2013, MNRAS, 430, 121

Preston G. W., Shectman S. A., Beers T. C., 1991, ApJ, 375, 121

Quinn P. J., Goodman J., 1986, ApJ, 309, 472

Rand, R. H., lecture notes, Lecture notes on *Non linear Vibrations*

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

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

Ryan S. G., 1992, AJ, 104, 1144

Salaris M., Chieffi A., Straniero O., 1993, ApJ, 414, 580

Samland M., Gerhard O. E., 2003, A&A, 399, 961

Sandage A., 1970, ApJ, 162, 841

Schneider D. P. et al., 2010, AJ, 139, 2360

Schönrich R., 2012, MNRAS, 427, 274

Schönrich R., Asplund M., Casagrande L., 2011, MNRAS, 415, 3807

Schönrich R., Binney J., Asplund M., 2012, MNRAS, 420, 1281

Schönrich R., Binney J., Dehnen W., 2010, MNRAS, 403, 1829

Schwarzschild M., 1979, ApJ, 232, 236

Searle L., Rodgers A. W., 1966, ApJ, 143, 809

Searle L., Zinn R., 1978, ApJ, 225, 357

Sellwood J. A., 2010a, MNRAS, 409, 145

Sellwood J. A., 2010b, ArXiv e-prints

Sellwood J. A., Carlberg R. G., 1984, ApJ, 282, 61

Sesar B. et al., 2010, ApJ, 708, 717

Shapely H., 1918, ApJ, 48, 154

Sharma S., Bland-Hawthorn J., Johnston K. V., Binney J., 2011, ApJ, 730, 3

Shen J., Rich R. M., Kormendy J., Howard C. D., De Propris R., Kunder A., 2010, ApJ Letters, 720, L72

Sirko E. et al., 2004, AJ, 127, 914

Smith M. C. et al., 2009, MNRAS, 399, 1223

Smith M. C. et al., 2007, MNRAS, 379, 755

Smith M. C., Wyn Evans N., An J. H., 2009, ApJ, 698, 1110

Sohn S. T., Besla G., van der Marel R. P., Boylan-Kolchin M., Majewski S. R., Bullock J. S., 2013, ApJ, 768, 139

Stanek K. Z., Mateo M., Udalski A., Szymanski M., Kaluzny J., Kubiak M., 1994, ApJ Letters, 429, L73

Steinmetz M., 2003, in Astronomical Society of the Pacific Conference Series, Vol. 298, GAIA Spectroscopy: Science and Technology, Munari U., ed., p. 381

Steinmetz M., 2012, Astronomische Nachrichten, 333, 523

Suntzeff N. B., Kinman T. D., Kraft R. P., 1991, ApJ, 367, 528

Toomre A., Toomre J., 1972, ApJ, 178, 623

Tremaine S., 1999, MNRAS, 307, 877

Tsoutsis P., Kalapotharakos C., Efthymiopoulos C., Contopoulos G., 2009, A&A, 495, 743

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

Varghese A., Ibata R., Lewis G. F., 2011, MNRAS, 417, 198

Verde L., 2007, ArXiv e-prints

Watkins L. L., Evans N. W., van de Ven G., 2013, MNRAS, 430, 971

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

Whitney J. H. et al., 1998, ApJ, 495, 284

Wilhelm R., Beers T. C., Gray R. O., 1999, AJ, 117, 2308

Xue X.-X. et al., 2011, ApJ, 738, 79

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

Yanny B. et al., 2000, ApJ, 540, 825

Yanny B. et al., 2009, AJ, 137, 4377

Zibetti S., Charlot S., Rix H.-W., 2009, MNRAS, 400, 1181

Zinn R., 1985, ApJ, 293, 424