

3.4	Dealing with Measurement Errors and their Effect on the Parameter Recovery	27
3.5	The Impact of Deviations of the Data from the Idealized qDF	33
3.6	The Implications of Assuming a Potential Model which Differs from the Real Potential	38
4	Discussion and Summary	41
4.1	Improved Computational Speed for Application to Larger Data Sets	41
4.2	Modelling Sensitivity to Properties and Unaccounted Imperfections of the Data Set	46
4.3	Data Deviations from the Modelling Assumptions about the Distribution Function and the Paper	
4.4	Different Modelling Approaches using Action-based Distribution Functions	49
4.5	On the Assumption of Axisymmetry	50
5	Acknowledgments	51
A	Appendix	52
A.1	Influence of wrong assumptions about incompleteness of the data parallel to the Galactic plane	
2	Stuff that still needs to be done or thought about	57

1. Introduction

Stellar dynamical modelling is the fundamental tool to infer the gravitational potential of the Milky Way from the positions and motions of its stars (Rix & Bovy 2013; Binney 2011b; Binney & Tremaine 2008) [TO DO: other / better references???]. The observational information on the phase-space coordinates of stars are currently growing at a rapid pace, and will be taken to a whole new level by the upcoming Gaia data. Yet, rigorous and practical modelling tools that turn this information into constraints both on the gravitational potential and on the distribution function (DF) of stellar orbits, are scarce (Rix & Bovy 2013) [TO DO: more references] [TO DO: References that explain that the modelling is scarce, or previous modelling approaches???

Accurately determining the Galactic gravitational potential is fundamental for understanding its dark matter and baryonic structure [TO DO: REF]. Accurately determining the stellar-population dependent orbit distribution function is a fundamental constraint on the

Summary of Comments on DynamicsPaper1_draft_v2.02.pdf

Page: 3

Author: rix Subject: Polygon	Date: 30.08.15 13:37:18
This reads like the intro to a book, or thesis. The basic elements are there, and should be there. We should just perhaps be a bit more sparing with "far-reaching", "fundamenta;" etc..	
Author: rix Subject: Polygonal Line	Date: 30.08.15 13:33:10
Author: rix Subject: Inserted Text	Date: 30.08.15 13:34:30
Author: rix Subject: Inserted Text	Date: 30.08.15 13:33:25
Author: rix Subject: Inserted Text	Date: 30.08.15 13:33:40
Author: rix Subject: Cross-Out	Date: 30.08.15 13:33:42
Author: rix Subject: Inserted Text	Date: 30.08.15 13:33:59
Author: rix Subject: Inserted Text	Date: 30.08.15 13:35:00
Author: rix Subject: Inserted Text	Date: 30.08.15 13:34:15
Author: rix Subject: Cross-Out	Date: 30.08.15 13:35:52
Author: rix Subject: Inserted Text	Date: 30.08.15 13:36:29
Author: rix Subject: Inserted Text	Date: 30.08.15 13:37:29
Author: rix Subject: Inserted Text	Date: 30.08.15 13:38:10

Galaxy's formation history.

Open questions about the MW's potential and structure, on which future modelling attempts will hopefully give more definite answers are: What is the local dark matter density (Zhang et al. 2013; Bovy & Tremaine 2012)? Is the Milky Way's dark matter halo flattened ([TO DO: REF])? Is the MW disk maximal (Sackett 1997) and, to be able to disentangle halo and disk contribution (Dehnen & Binney 1998), what is the disk's overall mass scale length (Bovy & Rix 2013)?

Open questions about the star's distribution within the MW, which dynamical modelling can help to constrain, are: How are stellar kinematics and their chemical abundances are related (Sanders & Binney 2015) [TO DO: REF])? In particular, does the disk have a thin/thick disk dichotomy (Gilmore & Reid 1983) or is it a continuum of many exponential disks (Bovy et al. 2012d)? How does radial migration affect the orbit distribution (Sellwood & Binney 2002; Roškar et al. 2008a,b; Schönrich & Binney 2008; Minchev et al. 2011) [TO DO: These are References from Rix & Bovy 2013 - should I use all of them?]? To address these questions, observed stellar positions and motions need to be turned into full orbits - which stresses again the importance of having a reliable model for the MW's gravitational potential.

In the era of big Galactic surveys all of this could soon be within our reach. Not only will there be full 6D stellar phase-space coordinates for a thousand million of stars measured by Gaia (Perryman et al. 2001) to unprecedented precision by the end of 2016. But already with existing surveys (e.g., SEGUE (Beers et al. 2006), RAVE (Steinmetz et al. 2006), LAMOST (Newberg et al. 2012), APOGEE (Majewski 2012), Gaia-ESO (Gilmore et al. 2012), GALAH (Freeman 2012) [TO DO: I just copied this from Melissa Cannon paper. Should I reference all of them??? Not in reference list yet.]) and sophisticated machine learning tools (e.g. *The Cannon* by Ness et al. (2015)) to combine them, we will soon have huge data sets at our disposal.

In this work we present a rigorous, robust and reliable dynamical modelling machinery, strongly building on previous work by Binney & McMillan (2011); Binney (2012); Bovy & Rix (2013); Bovy (2015) and explicitly developed to exploit and deal with these large data sets in the future.

There is a variety of practical approaches to dynamical modelling of discrete collisionless tracers, such as the stars in the Milky Way (e.g. Jeans modelling; Büdenbender et al. (2015);

Author: rix Subject: Polygonal Line Date: 30.08.15 13:38:10

Author: rix Subject: Cross-Out Date: 30.08.15 13:40:58

Author: rix Subject: Line Date: 30.08.15 13:42:30
This paragraph is good, but 'too soon'

Author: rix Subject: Sticky Note Date: 30.08.15 13:41:51
There is a very strong emphasis (so far) on "local" references.

Loebman et al. (2012); action-based DF modelling: Bovy & Rix (2013); Piffl et al. (2014); Sanders & Binney (2015); torus modelling: Made-to-measure modelling: McMillan & Binney (2012, 2013), De Lorenzi et al. (2007); Syer & Tremaine (1996); Bissantz et al. (2004) or Hunt & Kawata (2014); ~~[TO DO: What kind of modelling is Xiangxiang doing?]; Xue et al. (2015))~~. Most of them – explicitly or implicitly – describe the stellar distribution through a distribution function. Actions are good ways to describe orbits, because they are canonical variables with their corresponding angles, have immediate physical meaning, and obey adiabatic invariance (Binney & Tremaine 2008; McMillan & Binney 2008; Binney 2010; Binney & McMillan 2011; Binney 2011b).

Recently, Binney (2012) and Bovy & Rix (2013) [TO DO: are these the correct references??] proposed to combine parametrized axisymmetric potentials with DF's that are simple analytic functions of the three orbital actions to model discrete data. Binney (2010) and Binney & McMillan (2011) had proposed a set of simple action-based (quasi-isothermal) distribution functions (qDF). Ting et al. (2013) and Bovy & Rix (2013) showed that these qDF's may be good descriptions of the Galactic disk, when one only considers so-called mono-abundance populations (*MAP*), i.e. sub-sets of stars with similar $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ (Bovy et al. 2012b,c,d).

Bovy & Rix (2013) implemented a modelling approach that put action-based DF modelling of the Galactic disk in an axisymmetric potential in practice. Given an assumed potential and an assumed DF, they directly calculated the likelihood of the observed (\vec{x}, \vec{v}) for each sub-set of *MAP* among SEGUE Gdwarf (Yanny et al. 2009). This modelling also accounted for the complex, but known selection function of the kinematic tracers. For each *MAP*, the modelling resulted in a constraint of its DF, and an independent constraint on the gravitational potential, which members of all *MAP*s feel the same way. Taken as an ensemble, the individual *MAP* models constrained the disk surface mass density over a wide range of radii ($\sim 4 - 9$ kpc), and proved a powerful constraint on the disk mass scale length (~ 2 kpc) and on the disk to dark matter ratio at the Solar radius [TO DO: quote number???].

Yet, these recent models still leave us poorly prepared with the wealth and quality of the existing and upcoming data sets. This is because Bovy & Rix (2013) made a number of quite severe and idealizing assumptions about the potential, the DF and the knowledge of observational effects (such as the selection function). All these idealizations are likely to translate into systematic error on the inferred potential or DF, well above the formal error

bars of the upcoming data sets.

In this work we present *RoadMapping* ("Recovery of the Orbit Action Distribution of Mono-Abundance Populations and Potential Inference for our Galaxy") an improved and refined version of the original modelling machinery by Bovy & Rix (2013), making extensive use of the *galpy* python package (Bovy 2015). *RoadMapping* relaxes some of the restraining assumptions Bovy & Rix (2013) ~~had to~~ made, is more flexible and more adept in dealing with large data sets. In this paper we set out to explore the robustness of *RoadMapping* against the breakdowns of some of the most important assumptions of DF-based dynamical modelling. **What is it about the data, the model and the machinery itself that limits our recovery of the true gravitational potential?**

In the light of *Gaia* we explicitly analyze how well **the modelling machinery** behaves in the limit of large data. For a huge number of stars three ~~statistical~~ aspects become important, that ~~are~~ hidden behind Poisson noise for smaller data sets: (i) We have to make sure that our modelling is an un-biased ~~and asymptotically normal estimator~~ (§3.1). (ii) Numerical inaccuracies in the actual modelling machinery ~~start to matter and need to be avoided~~ (§2.5). (iii) Parameter estimates become so precise ~~that we start to be able to distinguish between similar models. We therefore~~ want more flexibility and more free fit parameters in the potential and DF model. The modelling machinery itself needs to be flexible and fast in effectively finding the best fit parameters for a large set of parameters. The improvements made to the machinery used in Bovy & Rix (2013) are presented in §2.6.

~~Different characteristics of the data might influence the success of the parameter recovery.~~ (i) In an era where we can choose data from different MW surveys, it might be worth to explore if different regions within the MW (i.e. differently shaped or positioned survey volumes) are especially diagnostic to recover the potential (§3.2). (ii) What happens if our knowledge about the selection function, specifically the completeness of the data set within the survey volume, is ~~not perfect~~ (§3.3)? (iii) How to account for measurement errors in the modelling (§3.4)?

One of the strongest assumptions is to restrict the dynamical modelling to a certain family of parametrized models. We investigate how well we can hope to recover the true potential, when our potential and DF models ~~deviate from~~ the true potential and DF. For the DF we specifically investigate two of our assumptions in §3.5: First, what would happen if the stars within MAPs do intrinsically not follow a single qDF as assumed by Ting et al.

Page: 6

	Author: rix Subject: Inserted Text	Date: 31.08.15 08:15:20
	explores and	
	Author: rix Subject: Cross-Out	Date: 31.08.15 08:15:24
	Author: rix Subject: Highlight	Date: 31.08.15 08:16:08
	This sentence, just stands by itself, and seems a bit out of context	
	Author: rix Subject: Cross-Out	Date: 31.08.15 08:16:52
	Author: rix Subject: Highlight	Date: 31.08.15 08:17:25
	If you introduce the term RoadMapping, use it, whenever applicable.	
	Author: rix Subject: Inserted Text	Date: 31.08.15 08:16:36
	the whisker	
	Author: rix Subject: Inserted Text	Date: 31.08.15 08:16:49
	data,	
	Author: rix Subject: Cross-Out	Date: 31.08.15 08:17:45
	Author: rix Subject: Inserted Text	Date: 31.08.15 08:18:02
	may be	
	Author: rix Subject: Cross-Out	Date: 31.08.15 08:18:40
	Unbiased, yes, but why does it have to be asymptotically normal?	
	Author: rix Subject: Inserted Text	Date: 31.08.15 08:19:23
	must not be an important source of systematics	
	Author: rix Subject: Inserted Text	Date: 31.08.15 08:19:36
	As p	
	Author: rix Subject: Inserted Text	Date: 31.08.15 08:20:17
	we need	
	Author: rix Subject: Cross-Out	Date: 31.08.15 08:22:07
	Author: rix Subject: Inserted Text	Date: 31.08.15 08:22:05
	We also explore how different aspects of the observational experiment design impact	
	Author: rix Subject: Inserted Text	Date: 31.08.15 08:22:45
	systematically imperfect	
	Author: rix Subject: Inserted Text	Date: 31.08.15 08:22:50
	best	
	Author: rix Subject: Inserted Text	Date: 31.08.15 08:36:25
	individual	
	Author: rix Subject: Inserted Text	Date: 31.08.15 08:38:09
	do not encompass	

(2013); Bovy & Rix (2013). ~~Second, and assuming MAPs do indeed follow the χ^2 DF, what would be the effect of pollution of MAPs through stars from neighbouring MAPs in the $([Fe/H], [a/Fe])$ plane due to too big abundance errors or bin sizes.~~ And last but not least, we test in §3.6 how well the modelling works, if our assumed potential family deviates from the true potential.

For all of these aspects we show some plausible and illustrative examples on the basis of investigating mock data. The mock data is generated from galaxy models presented in §2.1-2.3 following the procedure in §2.4, analysed according to the description of the machinery in §2.5-2.6 and the results are presented in §3 and discussed in §4.

The strongest assumption that goes into this kind of dynamical modelling might be the idealization of the Galaxy to be axi-symmetric and being in steady state. We do not investigate this within the scope of this paper but strongly suggest a systematic investigation of this for future work.

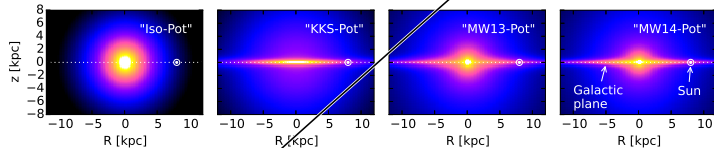


Fig. 1.— Density distribution of the four reference galaxy potentials in Table 1, for illustration purposes. These potentials are used throughout this work for mock data creation and potential recovery. [TO DO: Halo sichtbarer machen, evtl. mit isodensity contours]

Author: rix Subject: Cross-Out Date: 31.08.15 08:38:56
that can come in the main part.

Author: rix Subject: Cross-Out Date: 31.08.15 08:38:58

Author: rix Subject: Inserted Text Date: 31.08.15 08:39:03
Second,

Author: rix Subject: Sticky Note Date: 31.08.15 08:54:56
You might want to look at other paper's introduction: a standard part is the roadmap for the paper for the paper as a whole: In Section 2 we ..., while in Section 3 we ... Section 4 finally ...

2. Dynamical Modelling

2.1. Actions and Potential Models

Actions. Orbits in axisymmetric potentials are best described and fully specified by the three actions J_R, J_z and $J_\phi = L_z$. They are integrals of motion and generally defined as

$$J_i = \frac{1}{2\pi} \int_{\text{orbit}} p_i(t) dx_i(t) \quad (1)$$

and depend on the potential via the connection between position x_i and momentum p_i along the orbit. Actions have a clear physical meaning: They quantify the amount of oscillation in each coordinate direction of the full orbit [R14]. The position of a star along the orbit is denoted by a set of angles, which form together with the angles a set of canonical conjugate phase-space coordinates [Binney & Tremaine 2008]. Even though actions are the optimal choice as orbit labels and arguments for stellar distribution functions, their computation is very expensive.

Action calculation. The action calculation depends on the choice of potential in which the star moves. The spherical isochrone (Binney & Tremaine 2008) is the only potential for which Equation (1) takes an analytic form. For axisymmetric Stäckel potentials actions can be calculated exactly by the (numerical) evaluation of a single integral. In all other potentials numerically calculated actions will always be approximations, unless Equation (1) is integrated up to infinity. A computational fast way to get actions for arbitrary axisymmetric potentials is the "Stäckel fudge" by Binney (2012), which locally approximates the potential by a Stäckel potential. To speed up the calculation even more, an interpolation grid for J_R and J_z in energy E , angular momentum L_z and [TO DO: what else??] can be build out of these Stäckel fudge actions, as described in Bovy (2015).¹

Potential models. In our modelling we assume a family of parametrized potentials with a fixed number of free parameters. We use different kinds of potentials: Besides the Milky Way like potential from Bovy & Rix (2013) ("MW13-Pot") with bulge, disk and halo, we also extensively use the spherical isochrone potential ("Iso-Pot") in our test suites to make use of the analytic (and therefore exact and fast) way to calculate actions. In addition we use the 2-component Kuzmin-Kutuzov Stäckel potential by Batsleer & Dejonghe (1994) ("KKS-Pot"), which displays a disk and halo structure and also provides exact actions. Table 1

¹[TO DO: Write which numerical accuracy I needed for the grid, as the default values were not good enough.]

Author: rix Subject: Sticky Note Date: 02.09.15 09:11:03

Give a 1-2 sentence intro to the Section. E.g.

In this Section, we summarize the basic elements of the dynamical modelling presented here, which in many respects follows BR14.

Author: rix Subject: Cross-Out Date: 02.09.15 09:13:42

Author: rix Subject: Inserted Text Date: 02.09.15 09:11:30

Author: rix Subject: Inserted Text Date: 02.09.15 09:51:26

Author: rix Subject: Inserted Text Date: 02.09.15 09:11:43

Author: rix Subject: Inserted Text Date: 02.09.15 09:11:48

Author: rix Subject: Polygonal Line Date: 02.09.15 09:14:55

Author: rix Subject: Polygonal Line Date: 02.09.15 09:14:09

Author: rix Subject: Inserted Text Date: 02.09.15 09:34:52

Author: rix Subject: Cross-Out Date: 02.09.15 09:48:54

Author: rix Subject: Inserted Text Date: 02.09.15 09:36:21

Author: rix Subject: Inserted Text Date: 02.09.15 09:36:35

Author: rix Subject: Inserted Text Date: 02.09.15 09:36:43

Author: rix Subject: Inserted Text Date: 02.09.15 09:37:10

Author: rix Subject: Inserted Text Date: 02.09.15 09:38:27

Author: rix Subject: Inserted Text Date: 02.09.15 09:39:22

Author: rix Subject: Cross-Out Date: 02.09.15 09:41:29

Author: rix Subject: Inserted Text Date: 02.09.15 09:41:43

Author: rix Subject: Inserted Text Date: 02.09.15 09:42:20

summarizes all reference potentials together used in this work with their free parameters $p\Phi$. The density distribution of these potentials is illustrated in Figure 1.

2.2. Distribution Function

Distribution Function. Motivated by the findings of Bovy et al. (2012b,c,d) and Ting et al. (2013) about the simple phase space structure of MAPs, and following Bovy & Rix (2013) and their successful application, we also assume that each MAP follows a single qDF of the form given by Binney & McMillan (2011). This qDF is a function of the actions $\mathbf{J} = (J_R, J_z, L_z)$ and has the form

$$\begin{aligned} \text{qDF}(\mathbf{J} \mid p_{\text{DF}}) \\ = f_{\sigma_R}(J_R, L_z \mid p_{\text{DF}}) \times f_{\sigma_z}(J_z, L_z \mid p_{\text{DF}}) \end{aligned} \quad (2)$$

with

$$\begin{aligned} f_{\sigma_R}(J_R, L_z \mid p_{\text{DF}}) &= n \times \frac{\Omega}{\pi \sigma_R^2(R_g) \kappa} \exp\left(-\frac{\kappa J_R}{\sigma_R^2(R_g)}\right) \\ &\times [1 + \tanh(L_z/L_0)] \end{aligned} \quad (3)$$

$$f_{\sigma_z}(J_z, L_z \mid p_{\text{DF}}) = \frac{\nu}{2\pi \sigma_z^2(R_g)} \exp\left(-\frac{\nu J_z}{\sigma_z^2(R_g)}\right) \quad (4)$$

$$(5)$$

Here $R_g \equiv R_g(L_z)$ and $\Omega \equiv \Omega(L_z)$ are the (guidig-center) radius and the circular frequency of the circular orbit with angular momentum L_z in a given potential. $\kappa \equiv \kappa(L_z)$ and $\nu \equiv \nu(L_z)$ are the radial/epicycle (κ) and vertical (ν) frequencies with which the star would oscillate around the circular orbit in R - and z -direction when slightly perturbed (Binney & Tremaine 2008). The term $[1 + \tanh(L_z/L_0)]$ suppresses counter-rotation for orbits in the disk with $L \gg L_0$ which we set to a random small value ($L_0 = 10 \times R_\odot / 8 \times v_{\text{circ}}(R_\odot) / 220$).

For this qDF to be able to incorporate the findings by Bovy et al. (2012b,b,c) about the phase space structure of MAPs summarized in §1, we set the functions n , σ_R and σ_z , which indirectly set the stellar number density and radial and vertical velocity dispersion profiles,

$$n(R_g \mid p_{\text{DF}}) \propto \exp\left(-\frac{R_g}{h_R}\right) \quad (6)$$

$$\sigma_R(R_g \mid p_{\text{DF}}) = \sigma_{R,0} \times \exp\left(-\frac{R_g - R_\odot}{h_{\sigma,R}}\right) \quad (7)$$

$$\sigma_z(R_g \mid p_{\text{DF}}) = \sigma_{z,0} \times \exp\left(-\frac{R_g - R_\odot}{h_{\sigma,z}}\right). \quad (8)$$

Author: rix Subject: Inserted Text	Date: 02.09.15 09:45:33
Author: rix Subject: Inserted Text	Date: 02.09.15 09:45:34
Author: rix Subject: Inserted Text	Date: 02.09.15 09:50:35
Author: rix Subject: Polygonal Line	Date: 02.09.15 09:50:25
Author: rix Subject: Inserted Text	Date: 02.09.15 09:49:28
Author: rix Subject: Inserted Text	Date: 02.09.15 09:49:51
Author: rix Subject: Inserted Text	Date: 02.09.15 09:50:02
Author: rix Subject: Cross-Out	Date: 02.09.15 10:00:51
Author: rix Subject: Sticky Note	Date: 02.09.15 09:53:27
Author: rix Subject: Sticky Note	Date: 02.09.15 09:59:26
Author: rix Subject: Cross-Out	Date: 02.09.15 10:02:24
Author: rix Subject: Inserted Text	Date: 02.09.15 10:00:46
Author: rix Subject: Cross-Out	Date: 02.09.15 10:03:18
Author: rix Subject: Inserted Text	Date: 02.09.15 10:02:34
Author: rix Subject: Inserted Text	Date: 02.09.15 10:02:59

The qDF for each MAP has therefore a set of five free parameters p_{DF} : the density scale length of the tracers h_R , the radial and vertical velocity dispersion at the solar position R_\odot , $\sigma_{R,0}$ and $\sigma_{z,0}$, and the scale lengths $h_{\sigma,R}$ and $h_{\sigma,z}$, that describe the radial decrease of the velocity dispersion. The MAPs we use for illustration through out this work are summarized in Table 2.



Tracer Density. One crucial point in our dynamical modelling technique (§2.5), as well as in creating mock data (§2.4), is to calculate the (axisymmetric) spatial tracer density $\rho_{\text{DF}}(\mathbf{x} | p_\Phi, p_{\text{DF}})$ for a given qDF and potential. We do this by integrating the qDF at a given (R, z) over all three velocity components, using a N_{velocity} -th order Gauss-Legendre quadrature for each integral:

$$\begin{aligned} \rho_{\text{DF}}(R, |z| | p_\Phi, p_{\text{DF}}) &= \int_{-\infty}^{\infty} \text{qDF}(J[R, z, \mathbf{v} | p_\Phi] | p_{\text{DF}}) d^3\mathbf{v} \end{aligned} \quad (9)$$

$$\approx \int_{-N_{\text{sigma}}\sigma_R(R|p_{\text{DF}})}^{N_{\text{sigma}}\sigma_R(R|p_{\text{DF}})} \int_{-N_{\text{sigma}}\sigma_z(R|p_{\text{DF}})}^{N_{\text{sigma}}\sigma_z(R|p_{\text{DF}})} \int_0^{1.5v_{\text{circ}}(R_\odot)} \text{qDF}(J[R, z, \mathbf{v} | p_\Phi] | p_{\text{DF}}) dv_T dv_z dv_R \quad (10)$$

where $\sigma_R(R | p_{\text{DF}})$ and $\sigma_z(R | p_{\text{DF}})$ are given by Equations (7) and (8) and the integration ranges are motivated by Figure 2. The integration range $[0, 1.5v_{\text{circ}}(R_\odot)]$ over v_T is in general sufficient (only for observation volumes at smaller Galactocentric radii with larger velocities this upper limit needs to be increased). For a given p_Φ and p_{DF} we explicitly calculate the density on $N_{\text{spatial}} \times N_{\text{spatial}}$ regular grid points in the (R, z) plane; in between grid points the density is evaluated with a bivariate spline interpolation. The grid is chosen to cover the extent of the observations for $z > 0$. The total number of actions that need to be calculated to set up the density interpolation grid is $N_{\text{spatial}}^2 \cdot N_{\text{velocity}}^3$. §2.5 and Figure 3 show the importance of choosing N_{spatial} , N_{velocity} and N_{sigma} sufficiently large in order to get the density with an acceptable numerical accuracy.

2.3. Selection Function

Galactic Coordinate System. Our modelling takes place in the Galactocentric rest-frame with cylindrical coordinates $\mathbf{x} \equiv (R, \phi, z)$ and corresponding velocity components $\mathbf{v} \equiv (v_R, v_\phi, v_z)$. If the stellar phase-space data is given in observed coordinates, position $\mathbf{x} \equiv (\alpha, \delta, m - M)$ in right ascension α , declination δ and distance modulus $(m - M)$, and velocity $\mathbf{v} \equiv (\mu_\alpha, \mu_\delta, v_{\text{los}})$ as proper motions $\boldsymbol{\mu} \equiv (\mu_\alpha, \mu_\delta)$ [TO DO: cos somewhere??] and

Author: rix Subject: Sticky Note Date: 02.09.15 10:04:13

In this section, you have so indicate somehow, where you recapitulate BR13, and what is added new.

Author: rix Subject: Line Date: 02.09.15 10:05:55

Author: rix Subject: Line Date: 02.09.15 10:06:00

Author: rix Subject: Line Date: 02.09.15 10:05:49

Author: rix Subject: Polygonal Line Date: 02.09.15 10:05:30

Author: rix Subject: Highlight Date: 02.09.15 10:09:04

Why z>0? Is it not symmetric in z, by construction?

Author: rix Subject: Highlight Date: 02.09.15 10:09:32

This seems out of place here: move to the beginning of the paragraph.

line-of-sight velocity v_{los} , the data (\mathbf{x}, \mathbf{v}) has to be converted first into the Galactocentric rest-frame coordinates (\mathbf{x}, \mathbf{v}) using the sun's position and velocity. For simplicity we assume for the sun

$$\begin{aligned} (R_{\odot}, \phi_{\odot}, z_{\odot}) &= (8 \text{ kpc}, 0^{\circ}, 0 \text{ kpc}) \\ (v_{R,\odot}, v_{T,\odot}, v_{z,\odot}) &= (0, 230, 0) \text{ km s}^{-1}. \end{aligned}$$

Selection Function. A survey's selection function can be understood as a subvolume in the space of observables: e.g. position on the plane of the sky (limited by the pointing of the survey), distance from the sun (limited by the brightness of the stars and the sensitivity of the detector), colors and metallicity of the stars (limited by survey mode and targeting). Within the framework of this paper, using only mock data for testing purposes, we ignore target cuts in colors and metallicity and simply use spatial selection functions, which we define as

$$\text{sf}(\mathbf{x}) \equiv \begin{cases} \text{completeness}(\mathbf{x}) & \text{if } \mathbf{x} \text{ within observed volume} \\ 0 & \text{outside} \end{cases}$$

Its value describes the probability to observe a star at \mathbf{x} .

For the observed volume we use simple geometrical shapes: Either a sphere of radius r_{max} with the sun at its center, or a "wedge", which we define as the angular segment of an cylindrical annuli, i.e. the volume with $R \in [R_{\text{min}}, R_{\text{max}}]$, $\phi \in [\phi_{\text{min}}, \phi_{\text{max}}]$, $z \in [z_{\text{min}}, z_{\text{max}}]$ within the model galaxy. The sharp outer cut of the survey volume could be understood as the detection limit in apparent brightness in the case, where all stars have the same luminosity.

The completeness is, in our framework, a function of position with $0 \leq \text{completeness}(\mathbf{x}) \leq 1$ everywhere inside the observed volume. It could be understood as a position-dependent detection probability. Unless explicitly stated otherwise, we use everywhere

$$\text{completeness}(\mathbf{x}) = 1.$$

2.4. Mock Data

One goal of this work is to test how the loss of information in the process of measuring stellar phase space coordinates can affect the outcome of the modelling. To investigate this, we assume first that our measured stars do indeed come from our assumed families of

Author: rix Subject: Highlight	Date: 02.09.15 10:09:09
Author: rix Subject: Line	Date: 02.09.15 10:10:10
Author: rix Subject: Inserted Text	Date: 02.09.15 10:11:29
Author: rix Subject: Inserted Text	Date: 02.09.15 10:16:45
Author: rix Subject: Cross-Out	Date: 02.09.15 10:20:47
Author: rix Subject: Inserted Text	Date: 02.09.15 10:20:53
Author: rix Subject: Cross-Out	Date: 02.09.15 10:25:12
Author: rix Subject: Inserted Text	Date: 02.09.15 10:22:36
Author: rix Subject: Polygonal Line	Date: 02.09.15 10:26:03
Author: rix Subject: Cross-Out	Date: 02.09.15 10:26:15
Author: rix Subject: Cross-Out	Date: 02.09.15 10:25:18
Author: rix Subject: Inserted Text	Date: 02.09.15 11:30:47
Author: rix Subject: Inserted Text	Date: 02.09.15 11:31:22
Author: rix Subject: Cross-Out	Date: 02.09.15 11:32:21
Author: rix Subject: Inserted Text	Date: 02.09.15 11:34:36
Author: rix Subject: Cross-Out	Date: 02.09.15 11:35:01
Author: rix Subject: Inserted Text	Date: 02.09.15 11:34:57
Author: rix Subject: Cross-Out	Date: 02.09.15 11:35:26
Author: rix Subject: Polygonal Line	Date: 02.09.15 11:35:44
Author: rix Subject: Inserted Text	Date: 02.09.15 11:35:22
Author: rix Subject: Inserted Text	Date: 02.09.15 14:53:35

potentials and distribution functions and draw mock data from a given true distribution. In further steps we can manipulate and modify these mock data sets to mimick observational effects.

The distribution function is given in terms of actions and angles. The transformation $(J_i, \theta_i) \rightarrow (x_i, v_i)$ is however difficult to perform and computationally much more expensive than the transformation $(x_i, v_i) \rightarrow (J_i, \theta_i)$. We propose a fast and simple two step method for drawing mock data from an action distribution function, which also accounts effectively for a given survey selection function.

Preparation: Tracer density. We first setup the interpolation grid for the tracer density $\rho(R, |z| | p_\Phi, p_{DF})$ generated by the given qDF and according to §2.2 and Equation (10). For the creation of the mock data we use $N_{\text{spatial}} = 20$, $N_{\text{velocity}} = 40$ and $N_{\text{sigma}} = 5$.

Step 1: Drawing positions from the selection function. To get positions \mathbf{x} , for our mock data stars, we first sample random positions (R_i, z_i, ϕ_i) uniformly from the observed volume. Then we apply a rejection Monte Carlo method to these positions using the pre-calculated $\rho_{DF}(R, |z| | p_\Phi, p_{DF})$. In an optional third step, if we want to apply a non-uniform selection function, $\text{sf}(\mathbf{x}) \neq \text{const.}$ within the observed volume, we use the rejection method a second time. The sample then follows

$$\mathbf{x}_i \rightarrow p(\mathbf{x}) \propto \rho_{DF}(R, z | p_\Phi, p_{DF}) \times \text{sf}(\mathbf{x}).$$

Step 2: Drawing velocities according to the distribution function. The velocities are independent of the selection function and observed volume. For each of the positions (R_i, z_i) we now sample velocities directly from the qDF $(R_i, z_i, \mathbf{v} | p_{Phi}, p_{DF})$ using a rejection method. To reduce the number of rejected velocities, we use a Gaussian in velocity space as an envelope function, from which we first randomly sample velocities and then apply the rejection method to shape the Gaussian velocity distribution towards the velocity distribution predicted by the qDF. We now have a mock data set according to the required:

$$(x_i, v_i) \rightarrow p(\mathbf{x}, \mathbf{v}) \propto \text{qDF}(\mathbf{x}, \mathbf{v} | p_\Phi, p_{DF}) \times \text{sf}(\mathbf{x}).$$

Example: Figure 2 shows examples of mock data sets in configuration space (\mathbf{x}, \mathbf{v}) and action space. The qDF represents realistic stellar distributions in position-velocity space. More stars are found at smaller R and $|z|$, and are distributed uniformly in ϕ according to our assumption of axisymmetry. The distribution in radial and vertical velocities, v_R and v_z , is approximately Gaussian with the (total projected) velocity dispersion being $\sim \sigma_{R,0}$

Author: rix Subject: Inserted Text	Date: 02.09.15 14:54:06
Subsequently,	
Author: rix Subject: Polygonal Line	Date: 02.09.15 14:57:47
Author: rix Subject: Cross-Out	Date: 02.09.15 14:55:44
Author: rix Subject: Inserted Text	Date: 02.09.15 14:56:23
space	
Author: rix Subject: Inserted Text	Date: 02.09.15 14:58:41
within the entire observable	
Author: rix Subject: Inserted Text	Date: 02.09.15 14:59:06
to	
Author: rix Subject: Polygonal Line	Date: 02.09.15 15:00:11
Author: rix Subject: Inserted Text	Date: 02.09.15 14:59:42
resulting	
Author: rix Subject: Inserted Text	Date: 02.09.15 15:00:34
within the	
Author: rix Subject: Inserted Text	Date: 02.09.15 15:01:44
satisfying	
Author: rix Subject: Polygonal Line	Date: 02.09.15 15:01:57
Author: rix Subject: Cross-Out	Date: 02.09.15 15:02:25
Author: rix Subject: Inserted Text	Date: 02.09.15 15:05:48
The mock data from the	
Author: rix Subject: Inserted Text	Date: 02.09.15 15:06:25
lead to the expected distributions in configuration space	

and $\sim \sigma_{z,0}$ (see Table 2). The distribution of tangential velocities v_T is skewed because of asymmetric drift [TO DO: Find out, if we need an explanation for asymmetric drift here]. The distribution in action space demonstrates the intuitive physical meaning of actions: The stars of the "cool" MAP have in general lower radial and vertical actions, as they are on more circular orbits. The different relative distributions of the radial and vertical actions J_R and J_z of the "hot" and "cool" MAP is due to them having different velocity anisotropy $\sigma_{R,0}/\sigma_{z,0}$. The different ranges of angular momentum L_z in the two volumes reflect $L_z \sim R v_{\text{circ}}$ and the different radial extent of both volumes. The volume above the plane contains more stars with higher J_z , because stars with small J_z can't reach that far above the plane. Circular orbits with $J_R = 0$ and $J_z = 0$ can only be observed in the Galactic mid-plane. An orbit with L_z much smaller or larger than $L_z(R_\odot)$ can only reach into a volume located around R_\odot , if it is more eccentric and has therefore larger J_R . This together with the effect of asymmetric drift can be seen in the asymmetric distribution of J_R in the top central panel of Figure 2. [TO DO: Part of this could also be mentioned in the figure caption.]

Introducing measurement errors. If we want to add measurement errors to the mock data, we need to apply two modifications to the above procedure. First, measurement errors are best described in the phase-space of observables. We use the heliocentric coordinate system right ascension and declination (α, δ) and distance modulus $(m - M)$ as proxy for the distance from the sun, the proper motion in both α and δ direction (μ_α, μ_δ) and the line-of-sight velocity v_{los} . For the conversion between these observables and the Galactocentric cylindrical coordinate system in which the analysis takes place, we need the position and velocity of the sun, which we set for simplicity in this study to be $(R_\odot, z_\odot) = (8, 0)$ kpc and $(v_R, v_T, v_z) = (0, 230, 0)$ km s⁻¹. We assume Gaussian measurement errors in the observables $\tilde{\mathbf{x}} = (\alpha, \delta, (m - M))$, $\tilde{\mathbf{v}} = (\mu_\alpha, \mu_\delta, v_{\text{los}})$. Second, in the case of distance errors, stars can virtually scatter in and out of the observed volume. To account for this, we first draw "true" positions from a volume that is larger than the actual observation volume, perturb the stars positions according to the distance errors and then reject all stars that lie now outside of the observed volume. This procedure mirrors the Poisson scatter around the detection threshold for stars whose distances are determined from the apparent brightness and the distance modulus. [TO DO: Can I say it like this??] We then sample velocities (given the "true" positions of the stars) as described above and perturb them according to the measurement errors as well.

Author: rix Subject: Cross-Out	Date: 02.09.15 15:07:31
no, we don't	
Author: rix Subject: Inserted Text	Date: 02.09.15 15:07:45
illustrates	
Author: rix Subject: Highlight	Date: 02.09.15 15:08:11
Did you ever talk about "cool" and "hot" before?	
Author: rix Subject: Cross-Out	Date: 02.09.15 15:09:40
Author: rix Subject: Inserted Text	Date: 02.09.15 15:13:02
cannot	
Author: rix Subject: Cross-Out	Date: 02.09.15 15:16:30
Author: rix Subject: Cross-Out	Date: 02.09.15 15:16:19
You've said that before.	
Author: rix Subject: Inserted Text	Date: 02.09.15 15:16:49
, where we have taken m-M as a proxy for distance.	
Author: rix Subject: Highlight	Date: 02.09.15 15:15:31
Looks good to me.	

2.5. Likelihood

Form of the likelihood. As data we use the positions and velocities of stars coming from a given MAP and survey selection function $\text{sf}(\mathbf{x})$,

$$D = \{ \mathbf{x}_i, \mathbf{v}_i \mid \begin{array}{l} \text{(star } i \text{ belonging to same MAP)} \\ \wedge \text{ (sf}(\mathbf{x}_i) > 0) \end{array} \}$$

The model that we fit to the data is a parametrized potential and a single qDF with a given number of fixed and free parameters,

$$p_M = \{p_{\text{DF}}, p_\Phi\},$$

We fit the qDF parameters (see §2.2) with a logarithmically flat prior, i.e. flat priors in

$$p_{\text{DF}} := \{ \begin{array}{l} \ln(h_R/8\text{kpc}), \\ \ln(\sigma_{R,0}/220\text{km s}^{-1}), \ln(\sigma_{z,0}/220\text{km s}^{-1}), \\ \ln(h_{\sigma,R}/8\text{kpc}), \ln(h_{\sigma,z}/8\text{kpc}) \end{array} \}.$$

The orbit of the i -th star in a potential with p_Φ is labeled by the actions $\mathbf{J}_i := \mathbf{J}[\mathbf{x}_i, \mathbf{v}_i \mid p_\Phi]$ and the qDF evaluated for the i -th star is then $\text{qDF}(\mathbf{J}_i \mid p_M) := \text{qDF}(\mathbf{J}[\mathbf{x}_i, \mathbf{v}_i \mid p_\Phi] \mid p_{\text{DF}})$.

The likelihood of the data given the model $\mathcal{L} = (D \mid p_M)$ is the product of the probabilities for each star to move in the potential with p_Φ , being within the survey's selection function and it's orbit to be drawn from the qDF with p_{DF} , i.e.

$$\begin{aligned} \mathcal{L}(p_M \mid D) &\equiv \prod_i^N P(\mathbf{x}_i, \mathbf{v}_i \mid p_M) \\ &= \prod_i^N \frac{1}{(r_o v_o)^3} \cdot \frac{\text{qDF}(\mathbf{J}_i \mid p_M) \cdot \text{sf}(\mathbf{x}_i)}{\int d^3x d^3v \text{qDF}(\mathbf{J} \mid p_M) \cdot \text{sf}(\mathbf{x})} \\ &\propto \prod_i^N \frac{1}{(r_o v_o)^3} \cdot \frac{\text{qDF}(\mathbf{J}_i \mid p_M)}{\int d^3x \rho_{\text{DF}}(R, |z| \mid p_M) \cdot \text{sf}(\mathbf{x})}, \end{aligned} \quad (11)$$

where N is the number of stars in the data set D . In the last step we used Equation (9). The factor $\prod_i \text{sf}(\mathbf{x}_i)$ is independent of the model parameters, so we simply evaluate Equation (11) in the likelihood calculation. We find the best set of model parameters by maximising the likelihood.

Author: rix Subject: Inserted Text	Date: 02.09.15 15:21:36
Data	
Author: rix Subject: Cross-Out	Date: 02.09.15 15:21:39
Author: rix Subject: Inserted Text	Date: 02.09.15 15:22:37
consider here	
Author: rix Subject: Inserted Text	Date: 02.09.15 15:25:25
is specified by a	
Author: rix Subject: Inserted Text	Date: 02.09.15 15:25:30
for the	
Author: rix Subject: Inserted Text	Date: 02.09.15 15:25:50
we assume a prior that is flat in	
Author: rix Subject: Inserted Text	Date: 02.09.15 15:26:26
, and in	
Author: rix Subject: Inserted Text	Date: 02.09.15 15:27:39