

# Action-based Modelling of the Milky Way Disk with *RoadMapping* and our Imperfect Knowledge of the “Real World”

Wilma H. Trick<sup>1,2</sup>, Jo Bovy<sup>3</sup>, and Hans-Walter Rix<sup>1</sup>

trick@mpia.de

## ABSTRACT

We present *RoadMapping*, a dynamical modelling machinery that aims to recover the Milky Way’s (MW) gravitational potential and the orbit distribution of stellar mono-abundance populations (MAP) in the Galactic disk (Bovy & Rix 2013; Binney & McMillan 2011; Binney 2012). *RoadMapping* is a full likelihood analysis that models the observed positions and velocities of MAP stars with an equilibrium, three-integral quasi-isothermal distribution function (qDF) in an axisymmetric potential. In preparation for the application to the large data sets of modern Galactic surveys like Gaia, we create and analyze a large suite of mock data sets and develop qualitative “rules of thumb” which characteristics and limitations of data, model and machinery affect constraints on the potential and qDF most. We find that, while the precision of the recovery increases with the number of stars, the numerical accuracy of the likelihood normalisation becomes increasingly important and dominates the computational efforts. The modelling has to account for the survey’s selection function, but *RoadMapping* seems to be very robust against small misjudgments of the data completeness. Large radial and vertical coverage of the survey volume gives in general the tightest constraints. But no observation volume of special shape or position and stellar population should be clearly preferred, as there seem to be no stars that are on manifestly more diagnostic orbits. We propose a simple approximation to include measurement errors at comparably low computational cost that works well if the distance error is  $\lesssim 10\%$ . The model parameter recovery is also still possible, if the proper motion errors are known to within 10% and are  $\lesssim 2 \text{ mas yr}^{-1}$ . We also investigate how small deviations of the stars’ distribution from the assumed qDF

---

<sup>1</sup>Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany

<sup>2</sup>Correspondence should be addressed to trick@mpia.de.

<sup>3</sup>University of Toronto [TO DO: What is Jo’s current address??]

affect the modelling: Having less stars at small Galactocentric radii  $R$  and with lower velocities as predicted by the qDF recovers the potential more reliably than having too many stars at small  $R$  and with higher velocities. The deviations of the true orbit distribution from the qDF introduced by binning stars into MAPs in the  $([\text{Fe}/\text{H}], [\alpha/\text{Fe}])$  plane and due to abundance errors do not matter much, as long as the qDF parameters of two neighbouring MAPs do not vary more than 20% [TO DO: CKECK]. As the modelling has to assume a parametric form for the gravitational potential, deviations from the true potential have to be expected. We find, that in the axisymmetric case we can still hope to find a potential that is indeed a reliable best fit within the limitations of the assumed potential. Overall *RoadMapping* works as a reliable and unbiased estimator, and is robust against small deviations between model and the “real world”.

*Subject headings:* Galaxy: disk — Galaxy: fundamental parameters — Galaxy: kinematics and dynamics — Galaxy: structure

## Contents

|          |  |          |
|----------|--|----------|
| <b>1</b> | <b>Introduction</b>                      | <b>3</b> |
| <b>2</b> | <b>Dynamical Modelling</b>               | <b>6</b> |
| 2.1      | Coordinate System . . . . .              | 6        |
| 2.2      | Actions and Potential Models . . . . .   | 6        |
| 2.3      | Stellar Distribution Functions . . . . . | 8        |
| 2.4      | Selection Functions . . . . .            | 10       |
| 2.5      | Mock Data . . . . .                      | 10       |
| 2.6      | Data Likelihood . . . . .                | 13       |

[TO DO: REMOVE TABLE OF CONTENTS FOR FINAL VERSION]

## 1. Introduction

[Note: The first corrections of Jo and HW are included. TO DO: REMOVE THIS LATER.]

Stellar dynamical modelling can be employed to infer the Milky Way’s gravitational potential from the positions and motions of individual stars (Binney & Tremaine 2008; Binney 2011b; Rix & Bovy 2013). Observational information on the 6D phase-space coordinates of stars is currently growing at a rapid pace, and will be taken to a whole new level in number and precision by the upcoming data from the Gaia mission (Perryman et al. 2001). Yet, rigorous and practical modelling tools that turn position-velocity data of individual stars 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??] [TO DO: Hans-Walter suggested a Sanders & Binney reference, but I’m still not sure to what kind of paper: modelling approach or review of scarce modelling tools...]

The Galactic gravitational potential is fundamental for understanding the Milky Way’s dark matter and baryonic structure (Rix & Bovy 2013; McMillan 2012; Strigari 2013; Read 2014) and the stellar-population dependent orbit distribution function is a basic constraint on the Galaxy’s formation history (Binney 2013; Rix & Bovy 2013; Sanders & Binney 2015) [TO DO: more references].

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: Kuijken & Gilmore (1989), Bovy & Tremaine (2012), Garbari et al. (2012), Zhang et al. (2013), Büdenbender et al. (2015); action-based DF modelling: Bovy & Rix (2013), Piffl et al. (2014), Sanders & Binney (2015); torus modelling: McMillan & Binney (2012, 2013); Made-to-measure modelling: Syer & Tremaine (1996), De Lorenzi et al. (2007) or Hunt & Kawata (2014). 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 rigorous 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 G-dwarf stars (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 MAPs 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 and on the disk-to-dark-matter ratio at the Solar radius.

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 dynamical modelling machinery by Bovy & Rix (2013), making extensive use of the *galpy* Python package (Bovy 2015) and the *Stäckel Fudge* for fast action calculations by Binney (2012). *RoadMapping* is robust and well-tested and explicitly developed to exploit and deal with the large data sets of the future. *RoadMapping* explores and relaxes some of the restraining assumptions that Bovy & Rix (2013) made and 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. Our goal is to examine which aspects of the data, the model and the machinery itself limit our recovery of the true gravitational potential.

In the light of the imminent Gaia data, we analyze how well *RoadMapping* behaves in the limit of large data. For a huge number of stars three aspects become important,

that may be hidden behind Poisson noise for smaller data sets: (i) We have to make sure that *RoadMapping* is an un-biased estimator (Section ??). (ii) Numerical inaccuracies in the actual modelling machinery must not be an important source of systematics (Section 2.6). (iii) As parameter estimates become so precise, we need more flexibility and more free 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 Section ??.

We also explore how different aspects of the observational experiment design impact 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 (Section ??). (ii) What happens if our knowledge about the selection function, specifically the completeness of the data set within the survey volume, is systematically imperfect (Section ??)? (iii) How to best account for individual measurement errors in the modelling (Section ??)?

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 do not encompass the true potential and DF. First, we examine in Section ?? what would happen if the stars within MAPs do intrinsically not follow a single qDF as assumed by Ting et al. (2013) and Bovy & Rix (2013). Second, we test in Section ?? how well the modelling works, if our assumed potential family deviates from the true potential.

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.

For all of the above 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 Sections 2.1-2.4 following the procedure in Section 2.5, analysed according to the description of the machinery in Section 2.6-?? and the results are presented in Section ?? and discussed in Section ??.

[TO DO: Comment from Hans-Walter: Make sure, any topic/issue appears only once]

[TO DO: Is now one quarter shorter than before. But maybe shorten it even more...] [TO DO: ]

## 2. Dynamical Modelling

[TO DO: HW: In this section you have to indicate somehow, where you recapitulate BR13 and what is added new.]

In this section we summarize the basic elements of *RoadMapping*, the dynamical modelling machinery presented in this work, which in many respects follows Bovy & Rix (2013).

### 2.1. 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 heliocentric coordinates, position  $\tilde{\mathbf{x}} \equiv (\text{RA}, \text{DEC}, m - M)$  in right ascension RA, declination DEC and distance modulus  $(m - M)$  as proxy for the distance from the sun, and velocity  $\tilde{\mathbf{v}} \equiv (\mu_{\text{RA}}, \mu_{\text{DEC}}, v_{\text{los}})$  as proper motions  $\boldsymbol{\mu} = (\mu_{\text{RA}}, \mu_{\text{DEC}})$  [TO DO: cos somewhere???] in both RA and DEC direction and line-of-sight velocity  $v_{\text{los}}$ , the data  $(\tilde{\mathbf{x}}, \tilde{\mathbf{v}})$  has to be converted first into the Galactocentric rest-frame coordinates  $(\mathbf{x}, \mathbf{v})$  using the sun’s position and velocity. We assume for the sun

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

### 2.2. Actions and Potential Models

Orbits in axisymmetric potentials are best described and fully specified by the three actions  $\mathbf{J} \equiv (J_R, J_z, J_\phi = L_z)$ , defined as

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

and which 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 [TO DO: REF: HW suggested Binney & Tremaine (2008), but I can’t find a corresponding statement in the book]. 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, §3.5.1).

Even though actions are excellent orbit labels and arguments for stellar distribution functions, their computation is typically very expensive and depends on the choice of potential in which the star moves. The spherical isochrone potential (Henon 1959) is the only [TO DO: Jo suggested "most general Galactic" instead of "only", but the isochrone is actually not Galactic... Ask him.] potential for which Equation 1 takes an analytic form (Binney & Tremaine 2008, §3.5.2). For 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 along the whole (often not periodic) orbit. A computational fast way to get actions for arbitrary axisymmetric potentials is the *Stäckel fudge* by ?, 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).<sup>1</sup>

For the gravitational potential in our modelling we assume a family of parametrized potential models with a fixed number of free parameters. We use different kinds of potentials: The Milky Way like potential from Bovy & Rix (2013) (MW13-Pot) with bulge,

---

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

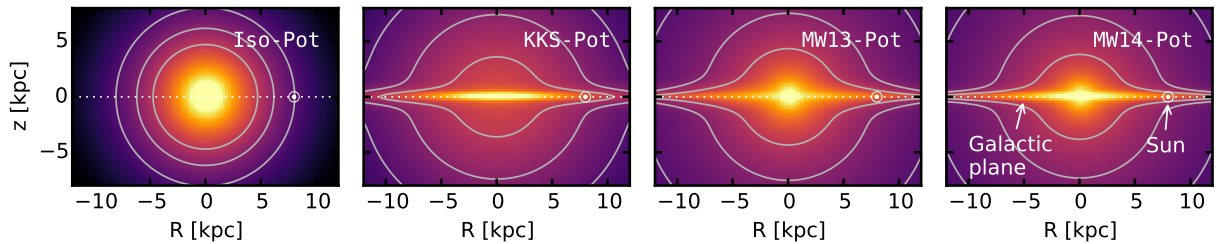


Fig. 1.— Density distribution of the four reference galaxy potentials in Table ??, for illustration purposes. These potentials are used throughout this work for mock data creation and potential recovery. [TO DO: Potential and/or population names in typewriter]

disk and halo; 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; and the 2-component Kuzmin-Kutuzov Stäckel potential (Batsleer & Dejonghe 1994; **KKS-Pot**), which displays a disk and halo structure and also provides exact actions. Table ?? 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.

### 2.3. Stellar Distribution Functions

Throughout, we assume that the orbits of each MAP can be described by a single qDF of the form given by Binney & McMillan (2011). This is 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. This qDF 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

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

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

Here  $R_g \equiv R_g(L_z)$  and  $\Omega \equiv \Omega(L_z)$  are the (guiding-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 ( $\kappa$ ) and vertical ( $\nu$ ) frequencies with which the star would oscillate around the circular orbit in  $R$ - and  $z$ -direction when slightly perturbed (Binney & Tremaine 2008, §3.2.3) [TO DO: ask someone, if I'm messing up different definitions of  $\kappa$ ] [TO DO: In all Bin08 references mention subsection]. 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 small value ( $L_0 = 10 \times R_\odot/8 \times v_{\text{circ}}(R_\odot)/220$  [TO DO: Jo said, galpy default is 10 km/s kpc. But I got the value actually from the code...]).

To match the observed properties of MAPs (see Bovy et al. 2012b,b,c), we chose the func-



tional forms

$$n(R_g | p_{\text{DF}}) \propto \exp\left(-\frac{R_g}{h_R}\right) \quad (5)$$

$$\sigma_R(R_g | p_{\text{DF}}) = \sigma_{R,0} \times \exp\left(-\frac{R_g - R_\odot}{h_{\sigma,R}}\right) \quad (6)$$

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

which indirectly set the stellar number density and radial and vertical velocity dispersion profiles. The qDF for each MAP has therefore a set of five free parameters  $p_{\text{DF}}$ : 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. Throughout this work we use for illustration purposes a few example stellar populations, each following a single qDF, whose parameters are given in in Table ?? . Most tests use the `hot` and `cool` qDFs from Table ??, which correspond to kinematically hot and cool populations, respectively.

One crucial point in our dynamical modelling technique (§2.6), as well as in creating mock data (§2.5), 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  $N_v$ -th order Gauss-Legendre quadrature for each integral:

$$\begin{aligned} & \rho_{\text{DF}}(R, |z| | p_\Phi, p_{\text{DF}}) \\ &= \int_{-\infty}^{\infty} \text{qDF}(\mathbf{J}[R, z, \mathbf{v} | p_\Phi] | p_{\text{DF}}) d^3\mathbf{v} \end{aligned} \quad (8)$$

$$\begin{aligned} & \approx \int_{-n_\sigma \sigma_R(R | p_{\text{DF}})}^{n_\sigma \sigma_R(R | p_{\text{DF}})} \int_{-n_\sigma \sigma_z(R | p_{\text{DF}})}^{n_\sigma \sigma_z(R | p_{\text{DF}})} \int_0^{1.5 v_{\text{circ}}(R_\odot)} \\ & \quad \text{qDF}(J[R, z, \mathbf{v} | p_\Phi] | p_{\text{DF}}) dv_T dv_z dv_R, \end{aligned} \quad (9)$$

where  $\sigma_R(R | p_{\text{DF}})$  and  $\sigma_z(R | p_{\text{DF}})$  are given by Equations 6 and 7 and the integration ranges are motivated by Figure 2. The integration range  $[0, 1.5 v_{\text{circ}}(R_\odot)]$  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  $p_\Phi$  and  $p_{\text{DF}}$  we explicitly calculate the density on  $N_x \times N_x$  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  $|z| \leq 0$ , because the model is symmetric in  $z$  by construction). The total number of actions that need to be calculated to set up the density interpolation grid is  $N_x^2 \cdot N_v^3$ . §2.6 and Figure 3 show the importance of choosing  $N_x$ ,  $N_v$  and  $n_\sigma$  sufficiently

large in order to get the density with an acceptable numerical accuracy [TO DO: Jo thinks that this statement is difficult to understand here, because you have not yet talked about the normalization].

## 2.4. Selection Functions

Any survey’s selection function can be understood as defining an effective sample sub-volume in the space of observables: e.g. position on the plane of the sky (the survey area), 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).

We simply use spatial selection functions, which describe the probability to observe a star at  $\mathbf{x}$ ,

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

For the observed volume we use simple geometrical shapes. Either a sphere of radius  $r_{\text{max}}$  with the sun at its center, or an angular segment of an cylindrical annulus (**wedge**), 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. Here  $0 \leq \text{completeness}(\mathbf{x}) \leq 1$  everywhere inside the observed volume, so it can be understood as a position-dependent detection probability. Unless explicitly stated otherwise, we simplify to  $\text{completeness}(\mathbf{x}) = 1$ .

## 2.5. Mock Data

We will rely on mock data as input to explore the limitations of the modelling. We investigate this, we assume first that our measured stars do indeed come from our assumed families of potentials and distribution functions and draw mock data from a given true distribution. Subsequently, we manipulate and modify these mock data sets to mimic observational effects.

The distribution function is given in terms of actions and angles. The transformation  $(\mathbf{J}_i, \boldsymbol{\theta}_i) \longrightarrow (\mathbf{x}_i, \mathbf{v}_i)$  is however difficult to perform and computationally much more expensive than the transformation  $(\mathbf{x}_i, \mathbf{v}_i) \longrightarrow (\mathbf{J}_i, \boldsymbol{\theta}_i)$ . We employ a fast and simple two-step method for drawing mock data from an action distribution function, which also accounts

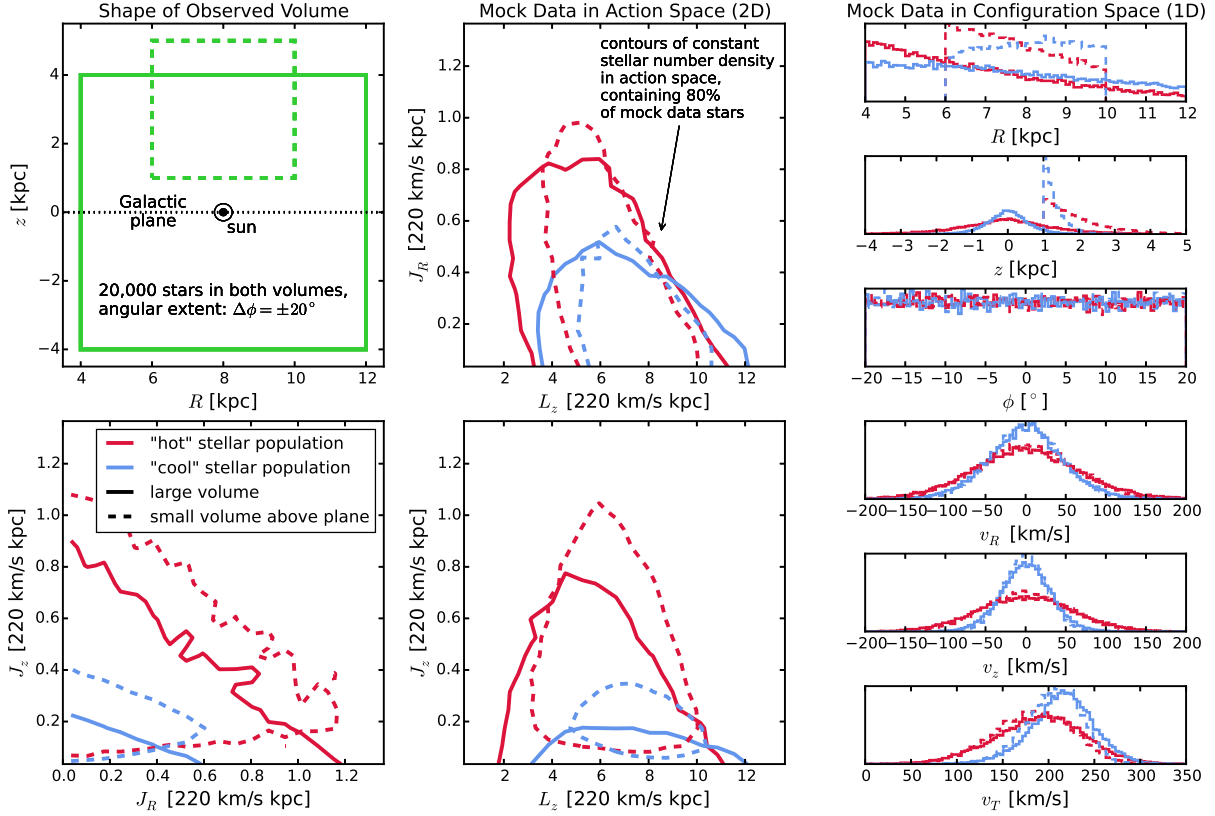


Fig. 2.— Distribution of mock data in action space (2D iso-density contours, enclosing 80% of the stars, the two central and the lower left panel) and configuration space (1D histograms, right panels), depending on shape and position of the survey observation volume and temperature of the stellar population. The parameters of the mock data model is given as Test ?? in Table ?. In the upper left panel we demonstrate the shape of the two different wedge-like observation volumes within which we were creating each a **hot** (red) and **cool** (blue) mock data set: a large volume centred on the Galactic plane (solid lines) and a smaller one above the plane (dashed lines). The distribution in action space visualizes how orbits with different actions also reach into different regions within the Galaxy. The 1D histograms on the right illustrate that qDFs generate realistic stellar distributions in galactocentric coordinates ( $R, z, \phi, v_R, v_z, v_T$ ). [TO DO: fancybox Legend] [TO DO: Potential and/or population names in typewriter font] [TO DO: Jo suggests to make two or three separate figures out of this. I'm not yet convinced, as I think it is nice and tidy like this.]

effectively for a given survey selection function.

In the first step we draw positions  $\mathbf{x}_i$  for our mock data stars from the selection function and tracer density. We start by setting up the interpolation grid for the tracer density  $\rho(R, |z| \mid p_\Phi, p_{\text{DF}})$  generated by the given qDF and according to §2.3 and Equation 9. For the creation of the mock data we use  $N_x = 20$ ,  $N_v = 40$  and  $n_\sigma = 5$ . Next, we sample random positions  $(R_i, z_i, \phi_i)$  uniformly within the entire observable volume. Then we apply a rejection Monte Carlo method to these positions using the pre-calculated  $\rho_{\text{DF}}(R, |z| \mid p_\Phi, p_{\text{DF}})$ . 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 resulting sample then follows  $\mathbf{x}_i \longrightarrow p(\mathbf{x}) \propto \rho_{\text{DF}}(R, z \mid p_\Phi, p_{\text{DF}}) \times \text{sf}(\mathbf{x})$ .

In the second step we draw velocities according to the distribution function. The velocities are independent of the selection function within the observed volume. For each of the positions  $(R_i, z_i)$  we sample velocities directly from the  $\text{qDF}(R_i, z_i, \mathbf{v} \mid p_\Phi, p_{\text{DF}})$  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 satisfying  $(\mathbf{x}_i, \mathbf{v}_i) \longrightarrow p(\mathbf{x}, \mathbf{v}) \propto \text{qDF}(\mathbf{x}, \mathbf{v} \mid p_\Phi, p_{\text{DF}}) \times \text{sf}(\mathbf{x})$ .

Figure 2 shows examples of mock data sets in configuration space  $(\mathbf{x}, \mathbf{v})$  and action space. The mock data from the qDF lead to the expected distributions in configuration space: More stars are found at smaller  $R$  and  $|z|$ , and are distributed uniformly in  $\phi$  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  $\sim \sigma_{R,0}$  and  $\sim \sigma_{z,0}$  (see Table ??). The distribution of tangential velocities  $v_T$  is skewed because of asymmetric drift. The distribution in action space illustrates the intuitive physical meaning of actions: The stars of the `cool` population 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` population 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  $L_z \sim Rv_{\text{circ}}$  and the different radial extent of both volumes. The volume above the plane contains stars with higher  $J_z$ , because stars with small  $J_z$  cannot 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.

If we want to add measurement errors to the mock data, we need to apply the following modifications to the above procedure. First, measurement errors are best described in heliocentric observables (see Section 2.1), we therefore assume and apply Gaussian errors to the *true* phase-space coordinates  $\tilde{\mathbf{x}} = (\text{RA}, \text{DEC}, (m - M))$ ,  $\tilde{\mathbf{v}} = (\mu_{\text{RA}}, \mu_{\text{DEC}}, v_{\text{los}})$ , where we have taken  $(m - M)$  as a proxy for distance. Second, in the case of distance errors, stars can virtually scatter in and out of the observed volume. To account for this, we draw the *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. We then sample velocities (given the *true* positions of the stars) as described above and perturb them according to the measurement errors as well.

## 2.6. Data Likelihood

As data we consider here 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 is specified by a number of fixed and free parameters,

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

For the qDF parameters (see Section 2.3) we assume a prior that is flat 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  $p_{\Phi}$  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  $p_{\Phi}$ , being within the survey's selection

function and it's orbit to be drawn from the qDF with  $p_{\text{DF}}$ , i.e.

$$\begin{aligned}
& \mathcal{L}(p_M \mid D) \\
& \equiv \prod_i^N p(\mathbf{x}_i, \mathbf{v}_i \mid p_M) \\
& = \prod_i^N \frac{\text{qDF}(\mathbf{J}_i \mid p_M) \cdot \text{sf}(\mathbf{x}_i)}{\int d^3x d^3v \text{qDF}(\mathbf{J} \mid p_M) \cdot \text{sf}(\mathbf{x})} \\
& \propto \prod_i^N \frac{\text{qDF}(\mathbf{J}_i \mid p_M)}{\int d^3x \rho_{\text{DF}}(R, |z| \mid p_M) \cdot \text{sf}(\mathbf{x})}, \tag{10}
\end{aligned}$$

where  $N$  is the number of stars in the data set  $D$ , and in the last step we used Equation 8. The factor  $\prod_i \text{sf}(\mathbf{x}_i)$  is independent of the model parameters, so we simply evaluate Equation 10 in the likelihood calculation. We find the best set of model parameters by maximizing the likelihood.

The normalisation in Equation 10 is a measure for the total number of tracers inside the survey volume,

$$M_{\text{tot}} \equiv \int d^3x \rho_{\text{DF}}(R, |z| \mid p_M) \cdot \text{sf}(\mathbf{x}). \tag{11}$$

In the case of an axisymmetric galaxy model and  $\text{sf}(\mathbf{x}) = 1$  everywhere inside the observed volume (i.e. a complete sample as assumed in most tests in this work), the normalisation is essentially a two-dimensional integral in  $R$  and  $z$  of the interpolated tracer density  $\rho_{\text{DF}}$  (see Equation 9 and surrounding text) over the survey volume times the observation volume's geometric angular contribution at each  $(R, z)$ . We perform this integral as a Gauss Legendre quadrature of order 40 in each  $R$  and  $z$  direction.

Unfortunately the evaluation of the likelihood for only one set of model parameters is computationally expensive. The computation speed is set by the number of action calculations required, i.e. the number of stars and the numerical accuracy of the integrals in Equation 9 needed for the normalisation, which requires  $N_x^2 \times N_v^3$  action calculations. The accuracy has to be chosen high enough, such that a resulting numerical error

$$\begin{aligned}
& \delta_{M_{\text{tot}}} \\
& \equiv \frac{M_{\text{tot}}(N_x, N_v, N_\sigma) - M_{\text{tot, true}}}{M_{\text{tot, true}}} \tag{12}
\end{aligned}$$

does not dominate the likelihood, i.e.

$$\begin{aligned}
& \log \mathcal{L}(p_M | D) \\
&= \sum_i^N \log \text{qDF}(\mathbf{J}_i | p_M) - 3N \log(r_o v_o) \\
& - N \log(M_{\text{tot,true}}) - N \log(1 + \delta_{M_{\text{tot}}}),
\end{aligned} \tag{13}$$

with

$$N \log(1 + \delta_{M_{\text{tot}}}) \lesssim 1.$$

In other words, this error is only small enough, if it does not affect the comparison of two adjacent models whose likelihoods differ, to be clearly distinguishable, by a factor of 10. Otherwise numerical inaccuracies could lead to systematic biases in the potential and DF fitting. For data sets as large as  $N = 20,000$  stars in one MAP, which in the age of Gaia could very well be the case [TO DO: Really???], we would need a numerical accuracy of 0.005% in the normalisation. Figure 3 demonstrates that the numerical accuracy we use in the analysis,  $N_x = 16$ ,  $N_v = 24$  and  $N_{\text{sigma}} = 5$ , does satisfy this requirement.

If the data is affected by measurement errors, they have to be accounted for in the likelihood. We assume Gaussian errors in the observable space  $\mathbf{y} \equiv (\tilde{\mathbf{x}}, \tilde{\mathbf{v}}) = (\text{RA}, \text{DEC}, (m - M), \mu_{\text{RA}}, \mu_{\text{DEC}}, v_{\text{los}})$ ,

$$\begin{aligned}
N[\mathbf{y}_i, \delta \mathbf{y}_i](\mathbf{y}') &= N[\mathbf{y}', \delta \mathbf{y}_i](\mathbf{y}_i) \\
&\equiv \prod_k \frac{1}{\sqrt{2\pi(\delta y_{i,k})^2}} \exp\left(-\frac{(y_{i,k} - y'_k)^2}{2(\delta y_{i,k})^2}\right),
\end{aligned}$$

where  $y_{i,k}$  is the  $k$ -th coordinate in  $\mathbf{y}_i$  of the  $i$ -th star. Observed stars follow the (quasi-isothermal) distribution function ( $\text{DF}(\mathbf{y}) \equiv \text{qDF}(\mathbf{J}[\mathbf{y} | p_\Phi] | p_{\text{DF}})$  for short), convolved with the error distribution  $N[0, \delta \mathbf{y}](\mathbf{y})$ . The selection function  $\text{sf}(\mathbf{y})$  acts on the space of (error affected) observables. Then the probability of one star coming from potential  $p_\Phi$ , distribution function  $p_{\text{DF}}$  and being affected by the measurement errors  $\delta \mathbf{y}$  becomes

$$\begin{aligned}
& \tilde{p}(\mathbf{y}_i | p_\Phi, p_{\text{DF}}, \delta \mathbf{y}_i) \\
&\equiv \frac{\text{sf}(\mathbf{y}_i) \cdot \int d^6 y' \text{DF}(\mathbf{y}') \cdot N[\mathbf{y}_i, \delta \mathbf{y}_i](\mathbf{y}')}{\int d^6 y \text{DF}(\mathbf{y}) \cdot \int d^6 y' \text{sf}(\mathbf{y}') \cdot N[\mathbf{y}, \delta \mathbf{y}_i](\mathbf{y}')}.
\end{aligned}$$

In the case of errors in distance or position, the evaluation of this is computational expensive - especially if the stars' have heteroscedastic errors  $\delta \mathbf{y}_i$ , for which the normalisation would have to be calculated for each star separately. In practice we apply the following approximation:

$$\begin{aligned}
& \tilde{p}(\mathbf{y}_i | p_\Phi, p_{\text{DF}}, \delta \mathbf{y}_i) \\
&\approx \frac{\text{sf}(\mathbf{x}_i)}{\int d^6 y \text{DF}(\mathbf{y}) \cdot \text{sf}(\mathbf{x})} \cdot \frac{1}{N_{\text{error}}} \sum_n^{N_{\text{error}}} \text{DF}(\mathbf{x}_i, \mathbf{v}[\mathbf{y}'_{i,n}])
\end{aligned} \tag{14}$$

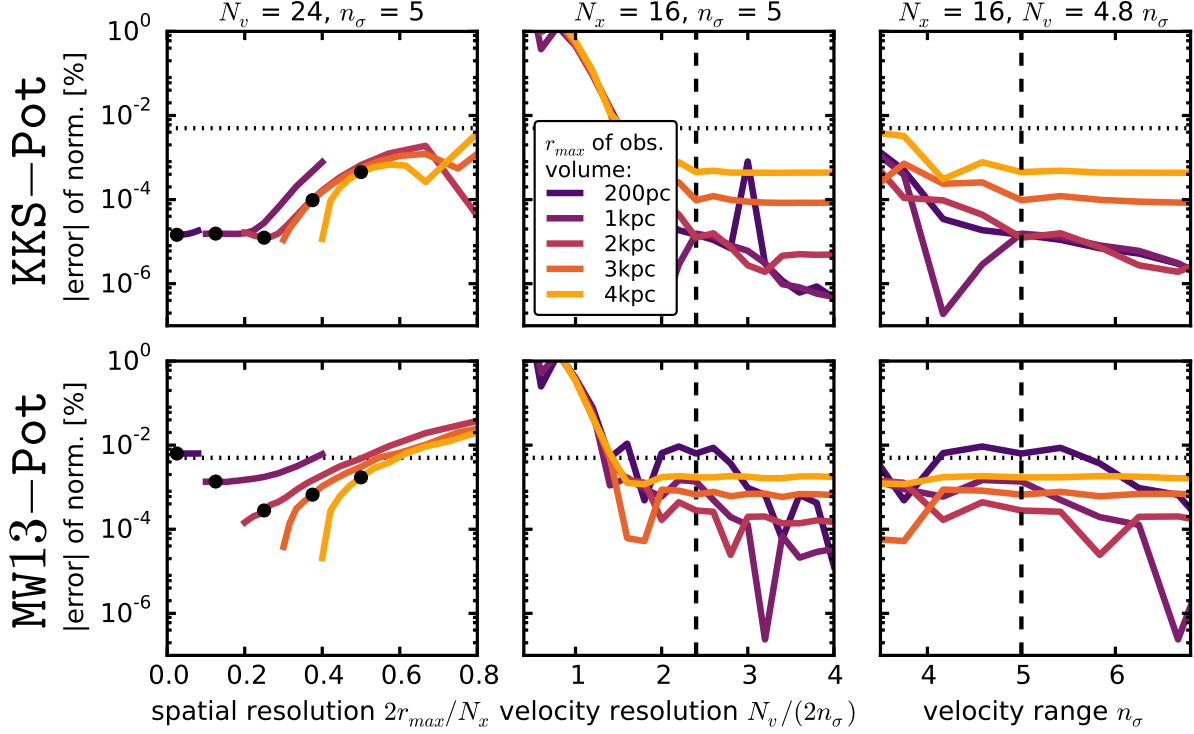


Fig. 3.— Relative error of the likelihood normalization in Equation 12 depending on the accuracy of the density calculation in Equation 9 (and surrounding text). The different colors represent calculations for different radii of the spherical observation volume around the sun, as indicated in the legend.  $N_x$  is the number of regular grid points in each  $R$  and  $z > 0$  within the observed volume on which the tracer density is evaluated according to Equation 9. At each  $(R, z)$  a Gauss-Legendre integration of order  $N_v$  is performed over an integration range of  $\pm N_x$  times the dispersion in  $v_R$  and  $v_z$  and  $[0, 1.5v_{\text{circ}}(R_\odot)]$  in  $v_T$ . To integrate the interpolated density over the observed volume to arrive at the likelihood normalization in Equation 11, we perform a 40th-order Gauss-Legendre integration in each  $R$  and  $z$  direction. We compare the convergence of the normalisation for the hot qDF in three potentials, Iso-Pot, MW13-Pot and KKS-Pot (see also Test ?? in Table ?? for all other model details). In each column of plots we keep two of the accuracy parameters fixed (indicated on top), while the third parameter is varied. (Caption continues on next page.)



Fig. 3.— (Continued.) We calculate the true normalization with high accuracy as  $M_{\text{tot,true}} \approx M_{\text{tot}}(N_x = 20, N_v = 56, N_\sigma = 7)$ . The dashed lines indicate the accuracy used in our analyses: it is better than 0.002% for all three potential types. Only for the smallest volume in the **MW13-Pot** (yellow line) the error is only  $\sim 0.005\%$ . This could be due to the fact, that, while we have analytical formulas to calculate the actions for the isochrone and the Staeckel potential exactly, we have to resort to an approximate action calculation for the MW-like potential (see Section 2.2). [TO DO: Try to redo yellow curve in MW. Weird, that it does not depend on  $N_{\text{spatial}}$ .???] [TO DO: fancybox Legend] [TO DO: Potential and/or population names in typewriter font] [TO DO: I completely changed this plot. Also change the whole caption, test table, text etc.]

with

$$\mathbf{y}'_{i,n} \sim N[\mathbf{y}_i, \delta \mathbf{y}_i](\mathbf{y}')$$

In doing so, we ignore errors in the star’s position  $\mathbf{x}_i$  altogether. This simplifies the normalisation drastically and makes it independent of measurement errors, including the velocity errors. Distance errors however are included, but only implicitly in the convolution over the stars’ velocity errors in the Galactocentric restframe. We calculate the convolution using Monte Carlo integration with  $N_{\text{error}}$  samples drawn from the full error Gaussian in observable space,  $\mathbf{y}'_{i,n}$ .

[TO DO: Check if all references are actually used in paper. ???]

## REFERENCES

- Batsleer, P., & Dejonghe, H. 1994, A& A [TO DO], 287, 43
- Binney, J. J. 2010, MNRAS, 401, 2318
- Binney, J. J., & McMillan, P. 2011, MNRAS, 413, 1889
- Binney, J. 2011, Pramana, 77, 39
- Binney, J. 2012, MNRAS, 426, 1324
- Binney, J. 2012, MNRAS, 426, 1328
- Binney, J. 2013, New A Rev., 57, 29
- Binney, J. & Tremaine, S. 2008, Galactic Dynamics: Second Edition

- Bovy, J., & Tremaine, S. 2012, *ApJ*, 756, 89
- Bovy, J., Rix, H.-W., & Hogg, D. W. 2012b, *ApJ*, 751, 131
- Bovy, J., Rix, H.-W., Hogg, D. W. et al., 2012c, *ApJ*, 755, 115
- Bovy, J., Rix, H.-W., Liu, C. et al., 2012d, *ApJ*, 753, 148
- Bovy, J., & Rix, H.-W. 2013, *ApJ*, 779, 115
- Bovy, J. 2015, *ApJS*, 216, 29 [TO DO]
- Büdenbender, A., van de Ven, G., & Watkins, L. L. 2015, *MNRAS*, 452, 956
- Dehnen, W., & Binney, J. 1998, *MNRAS*, 294, 429
- Dehnen, W. 1998, *AJ*, 115, 2384
- De Lorenzi F., Debattista V.P., Gerhard O., Sambhus N. 2007, *MNRAS*, 376, 7
- Famaey, B., & Dejonghe, H. 2003, *MNRAS*, 340, 752
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP* [TO DO], 125, 306
- Garbari, S., Liu, C., Read, J. I., & Lake, G. 2012, *MNRAS*, 425, 1445
- Henon, M. 1959, *Annales d’Astrophysique*, 22, 126
- Holmberg, J., Nordström, B., & Andersen, J. 2009, *A&A*, 501, 941
- Hunt, J. A. S., & Kawata, D. 2014, *MNRAS*, 443, 2112
- Jurić, M., Ivezić, Ž., Brooks, A., et al. 2008, *ApJ*, 673, 864
- Kawata, D., Hunt, J. A. S., Grand, R. J. J., Pasetto, S., & Cropper, M. 2014, *MNRAS*, 443, 2757
- Klement, R., Fuchs, B., & Rix, H.-W. 2008, *ApJ*, 685, 261
- Kuijken, K., & Gilmore, G. 1989, *MNRAS*, 239, 605
- McMillan, P. 2011, *MNRAS*, 414, 2446
- McMillan, P. J. 2012, *European Physical Journal Web of Conferences*, 19, 10002
- McMillan, P. J., & Binney, J. J. 2008, *MNRAS*, 390, 429

- McMillan, P. J., & Binney, J. 2012, MNRAS, 419, 2251
- McMillan, P. J., & Binney, J. J. 2013, MNRAS, 433, 1411
- Navarro, J. F., Helmi, A., & Freeman, K. C. 2004, ApJ, 601, L43
- Ness, M., Hogg, D. W., Rix, H.-W. et al., 2015, ApJ, 808, 16
- Nordström, B., Mayor, M., Andersen, J., et al. 2004, A&A, 418, 989
- Perryman, M. A. C., de Boer, K. S., Gilmore, G., et al. 2001, A&A, 369, 339
- Piffl, T., Binney, J., & McMillan, P. J. et al., 2014, MNRAS, 455, 3133
- Read, J. I. 2014, Journal of Physics G Nuclear Physics, 41, 063101
- Rix, H.-W., & Bovy, J. 2013, [TO DO] A& ARv, 21, 61
- Sackett, P. 1997, ApJ, 483, 103
- Sanders J. L., Binney J. 2015, MNRAS, 449, 3479
- Sellwood, J. A. 2010, MNRAS, 409, 145
- Steinmetz, M. et al., 2006, AJ, 132, 1645
- Strigari, L. E. 2013, Phys. Rep., 531, 1
- Syer D., Tremaine S. 1996, MNRAS, 282, 223
- Ting, Y.-S., Rix, H.-W., Bovy, J., & van de Ven, G. 2013, MNRAS, 434, 652
- Yanny, B., Newberg, H.-J., Johnson, J. A., et al. 2009, AJ, 137, 4377
- Zhang, L., Rix, H.-W., van de Ven, G., et al. 2013, ApJ, 772, 108

[TO DO: Mit wie vielen J. wird Binney geschrieben?] [TO DO: Kommas nach letztem Namen oder nicht?] [TO DO: In welcher Reihenfolge soll ich sortieren?] [TO DO: Wie viele Autoren nennen, bevor et al.???