Dissertation

submitted to the

Combined Faculties of the Natural Sciences and Mathematics

of the Ruperto-Carola-University of Heidelberg, Germany

for the degree of

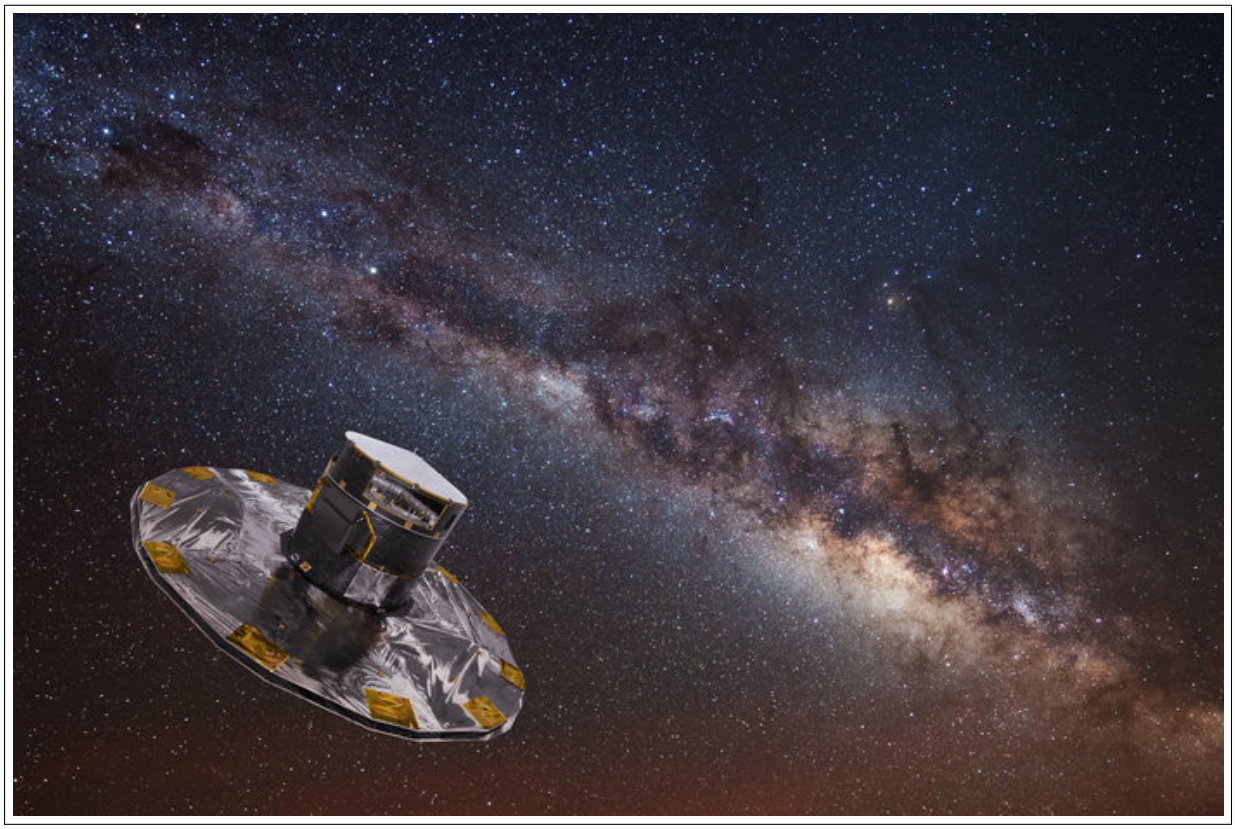
Doctor of Natural Sciences

Put forward by

Wilma Henriette Trick

Born in: Stuttgart, Germany

Oral examination: November 27, 2017



Action-based Dynamical Modeling for the Milky Way Disk

Dissertation

submitted to the Combined Faculties of the Natural Sciences and Mathematics
of the Ruperto-Carola-University of Heidelberg, Germany,
for the degree of *Doctor of Natural Sciences*

Put forward by

Wilma Henriette Trick

Born in: Stuttgart, Germany

Oral examination: November 27, 2017

Referees:
Prof. Dr. Hans-Walter Rix
Prof. Dr. Volker Springel

Cover image:
Description: Artist's impression of Gaia mapping the stars of the Milky Way.
Copyright: ESA/ATG medialab; background: ESO/S. Brunier

Abstract

Action-based dynamical modeling for the Milky Way disk. Understanding the Milky Way's present structure and assembly history constitutes a crucial constraint on galaxy formation and evolution theory. Galactic surveys like the *Gaia* mission provide high-precision measurements of positions, velocities, and chemical abundances of soon millions of stars in the Milky Way disk. Exploiting these high quality data requires sophisticated modeling tools.

This PhD thesis is dedicated to the development, characterization, and application of RoadMapping, a dynamical modeling machinery aiming to constrain the Galactic gravitational potential and chemo-orbital distribution function (DF) of the stellar disk. RoadMapping proceeds by modeling the observed discrete 6D phase-space positions of stellar mono-abundance populations (MAPs) by an axisymmetric parameterized potential model and an axisymmetric action-based orbit DF in a full-likelihood Bayesian framework. RoadMapping takes into account the survey's selection function (SF) and measurement uncertainties. RoadMapping builds on previous work by Bovy & Rix (2013), Binney & McMillan (2011), and Binney (2012a).

The first part of this work was published as Trick et al. (2016a) and gives an overview of the RoadMapping machinery. Its characteristics are studied by analyzing a large suite of axisymmetric mock data sets. It is found that RoadMapping constraints on the gravitational potential are robust against minor imperfections in the knowledge of the optimal potential or DF model family, selection effects, or velocity measurement uncertainties, as long as the distance uncertainties of the stars are better than 10%.

The second part is based on Trick et al. (2017) and investigates RoadMapping in the presence of spiral arms by modeling data drawn from an N-body simulation snapshot of a disk-dominated galaxy with strong spiral arms by D'Onghia et al. (2013). This provides a realistic test scenario for RoadMapping to model non-axisymmetric data with axisymmetric models. It is found that RoadMapping always recovers a good average model for the gravitational forces at the location of the stars that entered the analysis.

The third part applies RoadMapping to real data in the Milky Way. It combines measurements by *Gaia*-TGAS (Lindegren et al. 2016), RAVE (Kunder et al. 2017), and RAVE-on (Casey et al. 2017). Red clump stars are selected and photometric distances are assigned to them following Bovy et al. (2014). A strategy is devised to setup an SF for this sample that can be used in RoadMapping. The sample consists of 16 MAPs in the low-α disk. All MAPs provide independent and consistent constraints on the Milky Way's gravitational potential, measuring the disk scale length and circular velocity at the Sun to high precision, $R_{\mathrm{s,disk}} = 3.01 \pm 0.05$ kpc and $v_{\mathrm{circ}}(R_\odot) = 231.4 \pm 0.7$ km s^{-1}. The total surface mass density at the Sun that is recovered is, with $\Sigma_{\mathrm{tot,1.1kpc}} = 98 \pm 3$ M$_\odot$ pc^{-2}, larger than previous estimates in the literature, which is attributed, however, to the data.

Overall, RoadMapping is a well-tested and robust dynamical modeling machinery, whose preliminary and successful application to *Gaia* data promises new, precise, and reliable constraints on the Galactic gravitational potential in the near future.

Zusammenfassung

Action-basierte dynamische Modellierung für die Milchstraßenscheibe. Die Struktur der Milchstraße ist ein wichtiges Indiz für die Theorie der Galaxienentstehung und -entwicklung. Durchmusterungen der Milchstraße wie die *Gaia*-Mission liefern Hochpräzisionsmessungen von Positionen, Geschwindigkeiten und atmosphärischen Parametern von bald Millionen von Sternen in der Galaktischen Scheibe. Diese Daten verlangen fortgeschrittene Modellierungstechniken.
Diese Doktorarbeit widmet sich der Entwicklung, Charakterisierung und Anwendung von RoadMapping, einer Dynamischen-Modellierungs-Methode zur Messung des Galaktischen Gravitationspotentials und der chemo-orbitalen Verteilungsfunktion (DF) der stellaren Scheibe. RoadMapping modelliert die individuellen 6D Phasenraum-Positionen von Subpopulationen—bestehend aus Sternen mit ähnlicher chemischer Zusammensetzung, sogenannte MAPs—mithilfe von achsensymmetrischen, parametrisierten Potentialmodellen und Action-basierten Orbit-DFs auf Basis von Bayes'scher Statistik. RoadMapping berücksichtigt außerdem Selektionseffekte und Messunsicherheiten. RoadMapping ist eine Weiterentwicklung basierend auf früheren Arbeiten von Bovy & Rix (2013), Binney & McMillan (2011) und Binney (2012a).
Der erste Teil dieser Arbeit wurde unter Trick et al. (2016a) publiziert und präsentiert die RoadMapping-Modellierungstechnik. Er charakterisiert die Methode mithilfe der Analyse eines großen Satzes von achsensymmetrischen Mockdaten. RoadMapping liefert Messungen des Gravitationspotentials, die robust sind, selbst wenn das Wissen bezüglich der optimalen Potential- oder DF-Familien, Selektionseffekten oder Messfehlern in den Geschwindigkeiten geringfügig fehlerhaft ist und solange die Entfernungsfehler der Sterne kleiner als 10% sind.
Der zweite Teil basiert auf Trick et al. (2017) und untersucht den Effekt von Spiralarmen auf die RoadMapping-Modellierung anhand der N-Body-Simulation einer Scheibengalaxie mit ausgeprägten Spiralarmen von D'Onghia et al. (2013). Dies ermöglicht ein realistisches Testszenario, bei dem achsensymmetrische Modelle an nicht-achsensymmetrische Daten angepasst werden. RoadMapping findet stets ein gutes Durchschnittsmodell für die Gravitationskräfte an den Positionen jener Sterne, die Teil der Analyse waren.
Der dritte Teil wendet RoadMapping auf echte Daten in der Milchstraße an. Wir kombinieren Messungen von *Gaia*-TGAS (Lindegren et al. 2016), RAVE (Kunder et al. 2017) und RAVE-on (Casey et al. 2017). Rote-Klumpen-Sterne werden selektiert und ihre photometrischen Entfernungen analog zu Bovy et al. (2014) bestimmt. Eine Strategie wird entwickelt, um eine Selektionsfunktion für diesen Datensatz zu erstellen, die in RoadMapping verwendet werden kann. Der Datensatz besteht aus 16 MAPs in der α-defizienten Milchstraßen-Scheibe. Alle MAPs erlauben unabhängige und miteinander konsistente Messungen des Gravitationspotentials, und bestimmen die Skalenlänge der Scheibe und die Kreisgeschwindigkeit an der Sonne mit hoher Genauigkeit zu $R_{\mathrm{s,disk}} = 3.01 \pm 0.05$ kpc und $v_{\mathrm{circ}}(R_\odot) = 231.4 \pm 0.7$ kms s^{-1}. Die gemessene Gesamtmassenflächendichte an der Sonne ist mit $\Sigma_{\mathrm{tot,1.1kpc}} = 98 \pm 3$ $\mathrm{M}_\odot$ pc^{-2} wesentlich größer als frühere Messungen, was jedoch auf die Daten selbst zurückzuführen ist.
RoadMapping ist eine wohlerprobte und robuste Dynamische-Modellierungs-Methode, deren vorläufige und erfolgreiche Anwendung auf *Gaia*-Daten neue, genaue und verlässliche Messungen des Galaktischen Gravitationspotentials für die nahe Zukunft verspricht.

Contents

Abstract (English) v

Zusammenfassung (Deutsch) vi

Table of Contents vii

Preface xi

1 Introduction 1
- 1.1 Galaxy Formation in a Nutshell . . . 1
  - 1.1.1 From the Big Bang to Structure Formation . . . 1
  - 1.1.2 Open Questions about the Nature of Dark Matter . . . 4
  - 1.1.3 Formation and Population of Galaxies . . . 7
  - 1.1.4 The Evolution of Disk Galaxies . . . 10
  - 1.1.5 Open Questions about Disk Galaxies . . . 12
- 1.2 The Structure of the Milky Way . . . 13
  - 1.2.1 Stellar Spheroids in the Galaxy . . . 14
  - 1.2.2 The Disk . . . 15
  - 1.2.3 Stellar Populations in the Disk . . . 17
  - 1.2.4 (Sub-)Structure in the Stellar Disk . . . 20
  - 1.2.5 Open Questions about the Milky Way's Structure . . . 22
- 1.3 Galactic Archaeology in the Milky Way . . . 23
  - 1.3.1 Introduction to Galactic Archaeology . . . 23
  - 1.3.2 The Milky Way as Ideal Site for Galactic Archaeology . . . 25
  - 1.3.3 Importance of Constraining the Galaxy's Potential and Disk DF . . . 26
  - 1.3.4 Galactic Surveys . . . 28
- 1.4 Galactic Dynamics . . . 30
  - 1.4.1 Galactic Dynamics in an Axisymmetric, Collisionless Framework . . . 30
  - 1.4.2 Heuristic Introduction to Actions . . . 32
  - 1.4.3 Formal Introduction to Actions . . . 36
  - 1.4.4 Action Estimation . . . 39
  - 1.4.5 Excursion: Resonances in the Disk . . . 42
- 1.5 Dynamical Modeling for the Milky Way Disk . . . 43
  - 1.5.1 Overview of Dynamical Modeling Methods . . . 44
  - 1.5.2 The Quasi-isothermal Distribution Function . . . 45
  - 1.5.3 Introducing RoadMapping Modeling . . . 47
  - 1.5.4 Other Modeling Approaches using Action-based DFs . . . 50

2 RoadMapping Modeling and the Breakdown of its Assumptions 51
- 2.1 Preface . . . 51

2.2 RoadMapping Dynamical Modeling of Mock Data . . . 53
2.2.1 Coordinate System . . . 53
2.2.2 Action Estimation . . . 53
2.2.3 Potential Models . . . 53
2.2.4 Stellar DF Models and Tracer Density . . . 55
2.2.5 Selection Functions . . . 56
2.2.6 Creating Mock Data . . . 56
2.2.7 Data Likelihood . . . 57
2.2.8 Likelihood Normalization . . . 60
2.2.9 Measurement Uncertainties . . . 61
2.2.10 Fitting Procedure . . . 63
2.3 Results . . . 65
2.3.1 Model Parameter Estimates in the Limit of Large Data Sets . . . 69
2.3.2 The Role of the Survey Volume Geometry . . . 71
2.3.3 Impact of Misjudging the Selection Function of the Data Set . . . 71
2.3.4 Measurement Uncertainties and their Effect on the Parameter Recovery . . . 73
2.3.5 The Impact of Deviations of the Data from the Idealized DF . . . 77
2.3.6 The Implications of a Gravitational Potential not from the Space of Model Potentials . . . 80
2.3.7 The Influence of the Stellar Population's Kinematic Temperature . . . 83
2.4 Summary and Discussion . . . 83
2.4.1 Computational Speed . . . 84
2.4.2 Properties of the Data Set . . . 84
2.4.3 Deviations from the DF Assumption . . . 85
2.4.4 Gravitational Potential Beyond the Parameterized Functions Considered . . . 85
2.4.5 Outlook . . . 86

3 The Influence of Spiral Arms in RoadMapping Modeling 87
3.1 Preface . . . 87
3.2 Data from a Galaxy Simulation . . . 89
3.2.1 Description of the Galaxy Simulation Snapshot . . . 89
3.2.2 Data Selection and Survey Volume . . . 91
3.2.3 True Symmetrized Potential . . . 91
3.2.4 Quantifying the Strength of Spiral Arms . . . 93
3.2.5 Comparison of the N-Body Simulation to the Milky Way . . . 95
3.3 Ingredients for the RoadMapping Modeling . . . 97
3.3.1 Likelihood . . . 97
3.3.2 Distribution Function Model . . . 98
3.3.3 Potential Model . . . 99
3.4 Results . . . 100
3.4.1 An Axisymmetric Galaxy Model from RoadMapping . . . 100
3.4.2 The Influence of Spiral Arms in RoadMapping Modeling . . . 107
3.5 Discussion and Summary . . . 119
3.5.1 On the Informativeness of an Orbit Distribution Function . . . 119
3.5.2 On the Restrictiveness of the Parametrized Potential Model . . . 120
3.5.3 *Gaia* Measurement Errors and Choosing the Survey Volume Size . . . 120
3.5.4 Spiral Arms in the Solar Neighborhood . . . 121
3.5.5 Absence of a Central Bar in the Simulation . . . 121
3.5.6 Interpreting RoadMapping Results . . . 122
3.5.7 Summary . . . 122
3.5.8 Outlook . . . 123

4 RoadMapping Modeling of the Milky Way with Stars from *Gaia* and RAVE 125
4.1 Preface . . . 125
4.2 The Data . . . 126
4.2.1 Proper Motions from *Gaia*-TGAS . . . 127
4.2.2 Radial Velocities from the RAVE Survey . . . 129
4.2.3 Stellar Parameters and Chemical Abundances from RAVE-on . . . 130
4.2.4 Red Clump Stars as Standard Candles . . . 132
4.3 Selection Function . . . 136
4.3.1 Completeness of RAVE . . . 136
4.3.2 Completeness of TGAS . . . 137
4.3.3 Selection Function and Likelihood Normalisation for Red Clump Stars in RoadMapping . . . 140
4.4 First Glance at the Data . . . 145
4.4.1 Size and Location of the Data . . . 145
4.4.2 Spatial and Kinematic Structure of the Data . . . 146
4.4.3 Estimates for the Orbital Location . . . 147
4.5 Further Modeling Ingredients for the RoadMapping Analysis . . . 148
4.5.1 Potential Model . . . 148
4.5.2 Distribution Function for the Disk and Outlier Model for the Halo . . . 149
4.5.3 Likelihood and Priors . . . 152
4.5.4 Additional Details . . . 153
4.6 Results from RoadMapping Modeling of TGAS/RAVE Data . . . 154
4.6.1 First Considerations about the Best-fit Model . . . 154
4.6.2 Our Best-fit Estimate for the Galactic Gravitational Potential . . . 157
4.6.3 Findings about the Best-fit Stellar Distribution of MAPs . . . 161
4.7 Discussion and Summary . . . 165
4.7.1 What is the Reason for the Overestimation of the Surface Mass Density? 166
4.7.2 Precision vs. Accuracy in RoadMapping Modeling . . . 168
4.7.3 Discussion of the Potential Estimates and Implications . . . 170
4.7.4 Discussion of the Phase-space Structure of MAPs in a Galaxy Formation Context . . . 172
4.7.5 Substructure in Action Space . . . 174
4.7.6 Caveats of the Modeling and Future Improvements . . . 174
4.7.7 What will Improve with *Gaia* DR2? . . . 177
4.7.8 Summary . . . 177

5 Conclusion 181
5.1 Future Prospects . . . 181
5.2 Summary . . . 183

Acknowledgments 186

Abbreviations 187

Software 189

Publications by Wilma Trick 190

Bibliography 217

Statement 218

Preface

My PhD thesis is dedicated to the development and application of a technique that uses the motions of stars to constrain the distribution of (dark) matter and stellar orbits in our home galaxy, the Milky Way (MW). My motivation for this work is to understand and contribute to the story that explains the beauty of the night sky. This story is an important building block in our knowledge of galaxy formation and evolution. The interplay of the laws of physics—from the small scales of particle interactions, to the largest cosmological scales on which gravity and dark energy dominate—leads to highly complex processes and structures within our and other galaxies. Unravelling them leads therefore to a better understanding of the laws of nature themselves.
In the introduction of this work, I aim to tell this story of galaxy formation (Section 1.1) and Galactic structure (Section 1.2) as we currently know it. Based on the questions that are still left open in this story, I motivate in Section 1.3 the goals and strategies of the field of Galactic archaeology. This sub-field of astronomy aims to deduce the origins of our Galaxy from its stellar content. In Section 1.4, I present an overview of the principles of Galactic Dynamics that I will employ in this PhD thesis to learn more about the potential and stellar orbits in the MW. This section also provides an in-depth, but easy to understand introduction to orbital actions for future students. Section 1.5 ties everything together and presents the idea behind our dynamical modeling technique: *Recovery of the Orbit Action Distribution of Mono-Abundance Populations and Potential INference for our Galaxy (RoadMapping)*. I explain how and why RoadMapping is especially well-suited to help answering some of the open questions in galaxy formation and evolution.
Chapter 2 *characterizes* RoadMapping. It gives an overview of RoadMapping's technical framework and uses mock data to characterize its general robustness to different aspects of the modeling.
Chapter 3 *tests* RoadMapping. Here, we use data from an N-body spiral galaxy simulation as a realistic test case to prove that RoadMapping can be a powerful modeling tool in the MW.
Chapter 4 *applies* RoadMapping to real data in the MW from *Gaia* and RAVE. This preliminary application to ~ 5000 stars already gives new insight into the matter and orbit distribution in the disk.
Parts of this thesis, in particular Chapters 2 and 3, have already been published as peer-reviewed articles in the *Astrophysical Journal* as Trick et al. (2016a) and Trick et al. (2017).
In this PhD thesis, I will consistently use the pronoun "we". Even though I did all the work and writing, this choice is supposed to make the style of this thesis less awkward for the reader than using passive voice or the lecturing "I". In addition, the "we" reflects the fact, that science is always a team effort, and the research that lead to this thesis was accompanied by discussions and advice, and built upon previous work by my collaborators Hans-Walter Rix and Jo Bovy.
And last but not least, science is for everyone. It is often a challenging and murky path, and I want to take you, dear reader, on the path that I took during my doctoral studies on modeling the MW. Let us begin.

CHAPTER 1

Introduction

Galaxies are our lighthouses in the darkness of the universe. Their light guides us through the distant past of time and space and teaches us about the interplay of fundamental physics from the smallest to the largest scales. Galaxies are made out of billions of stars, gas, dust, and invisible matter, bound together by gravity. They come in a vast range of sizes, shapes, and colors, ever changing and evolving over the course of cosmic time. Any complete theory of physics will need to be able to explain not only their diversity, but also why there are fundamental similarities between them. What has set the initial conditions in the early universe from which these beautiful and complex structures emerged? What are the physical laws behind their evolution, in particular, what are the criteria for stability and the regulation mechanisms that shape them?

1.1 Galaxy Formation in a Nutshell

In this section, we will give a short overview of the current picture of galaxy formation, with a particular focus on the evolution of disk galaxies, and conclude with a list of open questions that still challenge our understanding of the universe. The main references for this section are Mo et al. (2010) and Maoz (2011), and references therein, and Carroll & Ostlie (2007). Figure 1.1 summarizes the most important events in the history of the universe.

1.1.1 From the Big Bang to Structure Formation

Big Bang. The universe, space and time, was born in a *Hot Big Bang*. This is a natural conclusion from the ground-breaking discovery by Hubble (1929) that the universe is expanding today and therefore had to be smaller early on. During its first 10^{-42} seconds the universe was dominated by quantum-effects of gravity, something that cannot be grasped by our current understanding of physics. Apart from this earliest phase, the evolution of space-time is well described by Einstein's theory of general relativity and the Friedmann equations (Friedmann 1922). The early universe was dense and hot and filled with a plasma of high-energy photons and low-mass particles like electrons, protons, etc. (Gamow 1948). The radiation-dominated plasma was in thermal equilibrium and subject to photon scattering, particle pair-production, interactions, and annihilation.

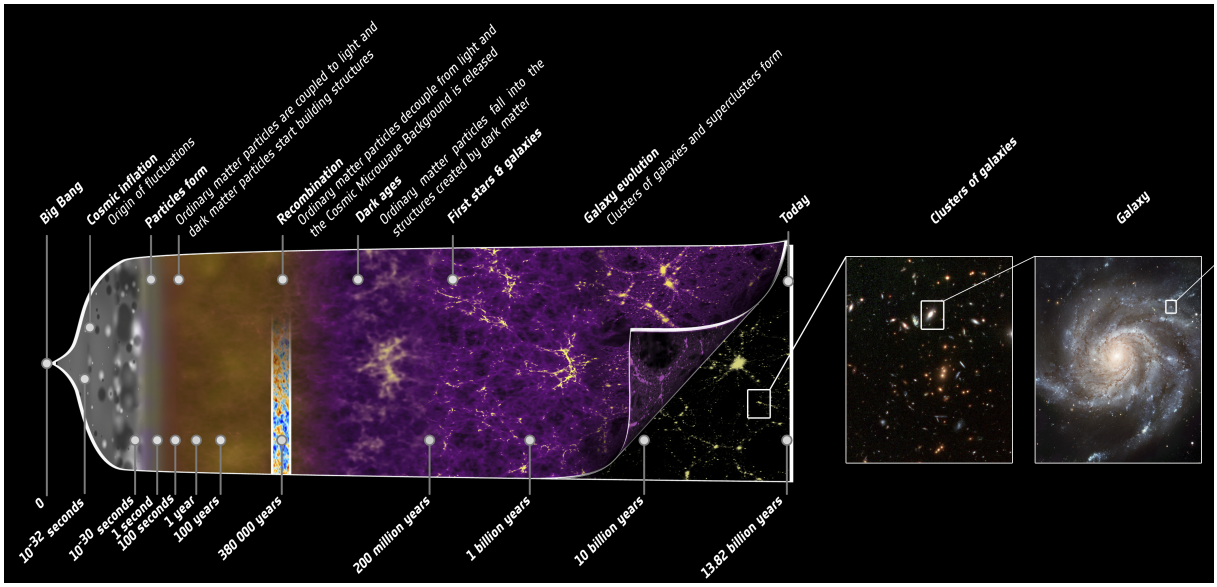


Figure 1.1: History of the universe, from the Big Bang to the formation of galaxies. (**Figure credit**: ESA, C. Carreau. `http://www.esa.int/spaceinimages/Images/2013/03/Planck_history_of_Universe`)

Inflation. *Inflation* started 10^{-36} seconds after the Big Bang, a phase of exponential expansion of space triggered by a scalar field with negative pressure. During inflation, primordial quantum fluctuations in this scalar field grew rapidly into large-scale perturbations that coupled to the radiation-matter fluid. It was also inflation that (i) separated regions that were previously in causal contact, explaining why the early universe appears isotropic and homogeneous everywhere (the *cosmological principle*), and (ii) drove the large-scale geometry of space towards being flat (Guth 1981).

Baryon asymmetry and nucleosynthesis. During subsequent phases of slower expansion, the photon content of the universe cooled down further. Whenever the photon temperature dropped below the rest-mass energy of a certain particle species, most of these particles annihilated with their respective anti-particles (*freeze-out*). An unexplained asymmetry between matter and anti-matter (known as *baryon asymmetry*) resulted in an excess of matter particles from which everything we observe today was later formed. A few minutes after the Big Bang, the *nucleosynthesis* of protons and neutrons into ionized deuterium, helium (Alpher et al. 1948; Hoyle & Tayler 1964), and a few other light elements (Wagoner et al. 1967) began, setting the ratios between the chemical elements that we observe today in primordial gas.

Primordial density perturbations. The matter budget of the universe is dominated by Dark Matter (DM), a mysterious collisionless matter component that does not interact with radiation (see Section 1.1.2). After inflation had instantiated the primordial density perturbations, the first DM overdensities—which initially expanded together with space—started to turn around and grown under their own gravitational attraction. The baryon-photon fluid began to fall into the growing potential wells of the DM. However, gas pressure halted its collapse and alternating overshooting between gravitational collapse and pressure-driven expansion caused standing density waves in the fluid (*acoustic oscillations*).

Recombination. Meanwhile, the universe cooled down and expanded further and finally photons were not energetic enough anymore to ionize atoms. Electrons recombined with nuclei to form neutral atoms. At $\sim 3.8 \times 10^5$ years after the Big Bang, when the universe had cooled down to

$\sim$ 3000 Kelvin, most of the matter was in the form of neutral atoms. Before this point in time, the universe was opaque to radiation due to the coupling to baryonic matter via scattering at free electrons. But at this so-called epoch of *recombination* and *last scattering*, radiation could for the first time move freely in space. We still observe the photons from this phase today, now cooled down to 2.7 Kelvin, as the *Cosmic Microwave Background (CMB)*. The existence of the CMB was first predicted by Gamow (1948) and Alpher & Herman (1948). In 1965 it was accidentally discovered by Penzias & Wilson (1965) as excess noise in their microwave antenna, and explained by Dicke et al. (1965). Mather et al. (1990) showed that the CMB has a blackbody spectrum.
CMB photons which came from the over- and under-dense regions of the acoustic oscillations have temperatures that differ slightly from the mean 2.7 Kelvin blackbody spectrum. Temperature variations observed today in the CMB (Smoot et al. 1992) are therefore the imprint of the first density perturbations in the universe.

Energy constituents. The energy constituents of the universe at the time of recombination set the physical conditions for the baryon acoustic oscillations and how the CMB photons cooled and expanded afterwards. Observations of the CMB and its power-spectrum—the strength in temperature fluctuations at a given angular scale—tell us therefore a lot about the universe (Spergel et al. 2007; Planck Collaboration et al. 2016). The CMB teaches us (i) the age of the universe, 13.7 ± 0.2 Gyr, (ii) that the overall space-time geometry is flat (de Bernardis et al. 2000), and (iii) that today $\sim 75\%$ of the universe's energy density is in the form of dark energy. Dark energy is even more mysterious than DM. It is assumed to have properties similar to the cosmological constant Λ introduced by Einstein in his field equations; some kind of vacuum energy with a negative equation of state, which is required to explain the accelerated expansion of space that we observe today in Supernova (SN) Type Ia standard candles (Riess et al. 1998; Perlmutter et al. 1999). The CMB tells us also (iv) about the total amount of matter in the universe, $\sim 25\%$, and (v) the ratio between dark and baryonic matter. DM sets the background gravitational potential, and the baryonic matter the medium for the acoustic oscillations, i.e., their ratio matters for the observed power spectrum. In fact, $\sim 21\%$ of the universe's energy density is DM and only $\sim 4\%$ is baryonic. In earlier times (until $\sim 10^4$ years after the Big Bang), hot radiation dominated the energy budget, but since the universe cooled down, its amount became negligible. This currently favored model of the universe is called the ΛCDM model.

Structure formation. A consequence of inflation is, that the initial random density perturbations were Gaussian and that their amplitudes were the same, independent of their wavelength (Harrison 1970; Zel'dovich 1970; Hawking 1982; Guth & Pi 1982; Bardeen et al. 1983). The density perturbations in the DM, which were decoupled from radiation, grew already before recombination according to gravitational Jeans instability (Jeans 1902; see also Section 1.1.4) and hierarchical merging. At recombination, diffusion of photons damped perturbations in the baryons smaller than typical galaxy masses (Silk 1968), making the CMB more uniform, but did not affect the substructure in the DM. Only after recombination the baryonic gas also started to quickly fall into these potential wells. It is expected that the gravitational collapse started first along one axis (Zel'dovich 1970), forming matter structures dubbed *Zel'dovich pancakes*. Torques acting between these non-spherical pancakes caused them to gain angular momentum (Hoyle 1951; Doroshkevich 1970; Peebles 1971; Efstathiou & Jones 1979) (unfortunately not enough to explain the amount of rotation observed in galaxies today, see Section 1.1.2).
The structures in the universe are subject to two different forces: Dark energy is driving space apart, and gravity pulls matter together. This caused the formation of large scale structures with thin matter filaments and empty voids in between (e.g., Gregory & Thompson 1978; Tully 1987; Gott et al. 2005; Springel et al. 2006). The matter filaments become denser, while the voids become larger and emptier with time.

1.1.2 Open Questions about the Nature of Dark Matter

DM is an essential ingredient for galaxy formation. Its behavior on cosmological scales, where gravity is the dominant force, is well-understood. On smaller, galaxy-size scales, where also the particle properties play a role, there are still a number of unresolved issues. Considering that DM accounts for $\sim 80\%$ of all matter in the universe, it is crucial to learn more about its nature. We will give an overview of the current status of research on DM, with the main references being Ostriker & Steinhardt (2003), in addition to Mo et al. (2010) and Maoz (2011).

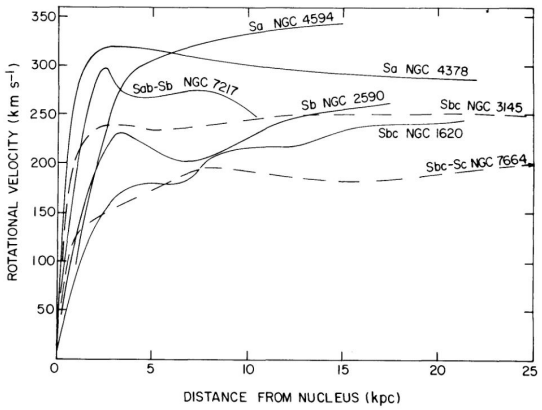
Historic discovery. The first indication that there could be a component of matter in the universe other than the visible stars, gas, and dust of galaxies, came from observations by Zwicky (1933). He investigated the motions of galaxies in the Coma cluster and realized that according to the virial theorem (see also Equation 1.2) the velocity dispersion implied a much higher amount of mass in this cluster than derived from the visible matter. Rubin et al. (1978, 1980) measured the rotation curves of many edge-on disk galaxies from optical spectra. Instead of declining with radius, r, as expected from the observed exponential distribution of stars and gas, the rotation curves, $v_{\rm circ}(r)$, stay approximately constant over a large radial range (see Figure 1.2(a)). If the laws of gravity are correct, this means that galaxies are embedded in large and massive spheroidal halos of invisible *dark matter* and the rotation curve or circular velocity curve is

$$v_{\rm circ}(r) = \sqrt{\frac{GM(r)}{r}} \sim \text{const.} \tag{1.1}$$

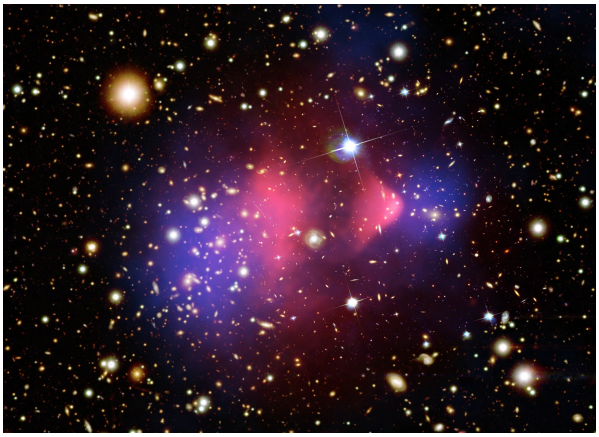
for a spherical mass distribution with $M(r)$ being the mass enclosed in r.
Since then, the existence of DM has been established as the current working theory. The need for DM in and around galaxies has also been shown via strong gravitational lensing (e.g., Trick et al. 2016b), weak gravitational lensing (Tyson et al. 1990; Kaiser & Squires 1993) and X-ray gas emission from galaxy clusters (e.g., Buote & Canizares 1992).

Need for non-baryonic DM in galaxy formation and cosmology. Ordinary baryonic matter is made from protons and neutrons—i.e. atoms—and strongly interacts with radiation via absorption and emission. There are several indications which suggest that DM is non-baryonic and interacts mostly via the gravitational force. As mentioned in Section 1.1.1, the CMB makes a strong case for non-baryonic DM: (i) Without non-baryonic DM that starts to collapse before recombination, there would be not enough time since recombination for the small baryonic density fluctuations observed in the CMB to grow into structures as large as galaxies. (ii) To model the angular scales and amplitudes in the power spectrum of the CMB fluctuations also a non-baryonic matter component is required.
Theories of Big Bang nucleosynthesis together with observed chemical abundance ratios in primordial gas predict that $\Omega_{\rm m,bary} = 4\%$ of the universe's energy density are baryons (e.g., Walker et al. 1991; Burles et al. 2001; Kirkman et al. 2003). With the estimates derived from SN Type Ia observations for the amount of dark energy ($\Omega_\Lambda = 75\%$; see Section 1.1.1), and the assumption that the space-time geometry is flat (de Bernardis et al. 2000; $\Omega_\Lambda + \Omega_{\rm m} = 1$), this leaves a discrepancy of 21% non-baryonic matter, consistent with the CMB estimate. This excludes any DM candidates that are made up from baryonic invisible matter like black holes, brown dwarfs, and white dwarfs as suggested by Carr et al. (1984) (see also below).

Collisionless cold dark matter. The currently favored model for DM is Collisionless Cold Dark Matter (CCDM), consisting of long-lived, massive (10 GeV to a few TeV), non-baryonic particles. It is expected that it has a small scattering cross-section, i.e., it is collisionless. Otherwise, its interaction with baryons or radiation should be possible to be observed other than just via



(a) Flat galaxy rotation curves (Rubin et al. 1978).



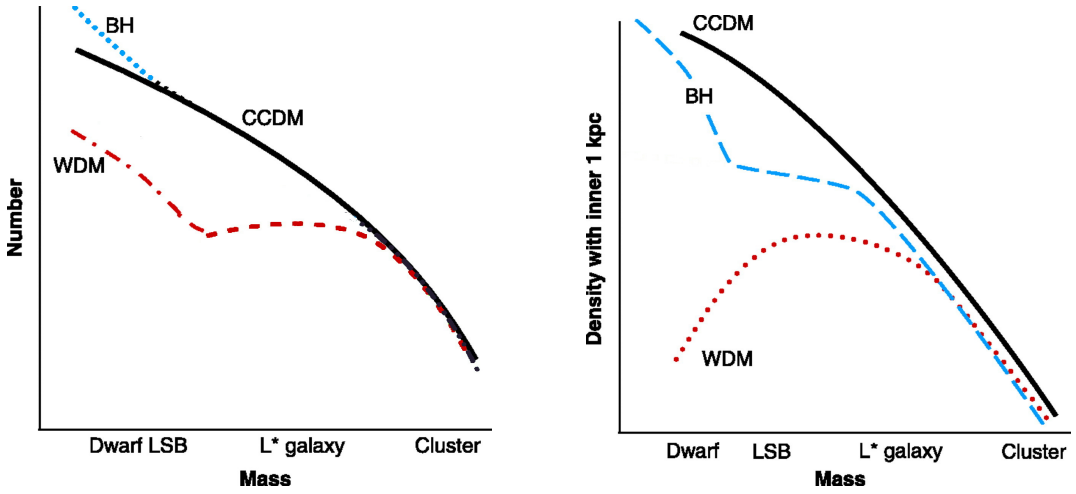
(b) Matter distribution in the Bullet cluster.

Figure 1.2: Observational evidence for the existence of DM. Panel (a) shows galactic rotation curves $v_{\mathrm{circ}}(r)$ measured by Rubin et al. (1978) from optical spectra for edge-on spiral galaxies. If only the visible matter existed in these galaxies, the rotation curves would be expected to decline with distance from nucleus r. As they are almost flat, it appears that there is an additional "dark" matter component. Panel (b) shows the Bullet cluster, two galaxy clusters that recently collided. The hot X-ray emitting gas which makes up most of the visible matter in galaxy clusters and is marked in red in the figure, got shocked during the collision and now lags behind the two moving sets of galaxies. The blue shaded regions show the total matter distribution derived from gravitational lensing of background galaxies around the cluster. This observation can only be explained if DM dominated the mass budget in galaxy clusters and the collision spatially separated the visible collisional gas and the invisible collisionless DM. **Figure credit:** Panel (a): Rubin et al. 1978; Panel (b): X-ray (red): NASA/CXC/CfA/Markevitch et al. (2002); Lensing map (blue): NASA/STScI; ESO WFI; Magellan/U.Arizona/Clowe et al. 2006; Optical: NASA/STScI; Magellan/U.Arizona/Clowe et al. 2006.

their gravitational interaction. This kind of DM particles is thought to decouple very early from the radiation-baryon fluid. Because the particles are massive, their motions should become non-relativistic already before recombination, so they can start to cluster and merge, and to form structure hierarchically bottom-up very early on. A very good candidate for CCDM are Weakly Interacting Massive Particles (WIMPs). This is a class of particles predicted by particle physics beyond the standard model, which only interact via gravity and the weak force. If the scattering cross section is set by the weak force, then theories of Big Bang nucleosynthesis would naturally explain the fraction of 21% DM in the universe (e.g., Jungman et al. 1996). The bottom-up structure formation in CCDM computer simulations agrees exceptionally well with the large-scale clustering of galaxies in surveys (Springel et al. 2006; Moster et al. 2010) assuming galaxies are tracers of the underlying structure of dark matter (Davis et al. 1985; Kaiser 1984).

Problems of CCDM on small scales. On the smaller spatial scales of galaxies, the thermodynamic properties of CCDM and interactions with baryonic matter and radiation become important and there are some difficulties with the CCDM model (e.g., Weinberg et al. 2015).

- *The missing satellite problem.* Numerical CCDM simulations predict a DM mass function with a higher number of satellite galaxies (see Figure 1.3(a)) around MW-like galaxies than is observed (e.g., Kauffmann et al. 1993; Moore et al. 1999; Klypin et al. 1999). This might not (only) be a problem with CCDM; it is also related to the question why the lowest mass halos did not form stars (e.g., White & Frenk 1991; Kauffmann et al. 1993; Cole et al. 1994; Stoehr et al. 2002; see also Section 1.1.5).



(a) Mass function of DM halos.

(b) DM density in the center of halos depending on halo mass.

Figure 1.3: Qualitative predictions on the mass function and inner halo density for different DM candidates. CCDM with its low thermal motion predicts many low-mass halos from hierarchical clustering and high-density cusps (black lines). Warm DM and DM made up from BHs, for example, would predict a quite different behavior for small galactic scales (see text in Section 1.1.2), but agree on large scales. This illustrates that resolving the true nature of DM requires the investigation of galaxies like the MW (which is a L^* galaxy) and smaller. **Figure credit:** Ostriker & Steinhardt (2003).

- *The core-cusp problem.* Simulations of CCDM predict a universal halo density profile with central density cusp (Dubinski & Carlberg 1991; Navarro et al. 1996, 1997; Moore et al. 1998; see Figure 1.3(b)), while observations are more consistent with density profiles that are flatter and cored in their center (e.g, galaxy cluster: Tyson et al. 1998; MW bulge: Binney & Evans 2001; dwarf galaxy: Kleyna et al. 2003). The simulations have mostly ignored the effect of baryons on the DM structure and observations often have their own difficulties. So this apparent conflict might be resolved in the future.

- *The cosmological angular momentum problem.* Disk galaxies in simulations have lower angular momentum and therefore smaller disk scale lengths than real galaxies (e.g., Navarro & Steinmetz 2000; Burkert & D'Onghia 2004).

DM Alternatives. A multitude of alternatives to CCDM has been proposed. Here we summarize the most common ones.

- *Massive Compact Halo Objects (MACHOs).* One theory was that DM halos in galaxies could be made up of non-luminous bodies as Black Holes (BHs), neutron stars, white dwarfs, brown dwarfs, and planets. *Advantage:* If DM consisted of massive BHs (created possibly by the SN explosions of Population III stars), friction between BHs and ordinary matter would increase the number density of BHs in galaxy centers (Figure 1.3(b)) and merging could explain the origin of Supermassive Black Holes (SMBHs) (Lacey & Ostriker 1985; see Section 1.1.3). *Counter evidence:* However, these SNe would have chemically enriched the halo gas, which is not observed. MACHOs are also not consistent with the hypothesis that DM should be non-baryonic. Finally, MACHOs were ruled out as candidate for (the majority of the) DM in the universe because of the low number of observed microlensing events in the MW halo (Paczynski 1986; Alcock et al. 1997).

- *Neutrinos as hot DM. Advantage:* Neutrinos are known non-baryonic particles and exist in large numbers in the universe. *Counter evidence:* The neutrino mass is still not known but suspected to be very tiny. The motions of neutrinos were therefore relativistic for quite a while after the Big Bang. If neutrinos accounted for all DM in the universe, their high thermal motions would have removed all small scale structure early on. Only the largest primordial fluctuations would have been possible to grow and structures would have formed via top-down fragmentation. This predicts a clustering of galaxies that is not consistent with observations (White et al. 1984).

- *Warm DM.* Warm DM is expected to consist of particles with masses of the order of 1keV, more massive than the hot neutrino DM, but less massive than cold DM. The slightly higher thermal motions compared to cold DM would allow initial structures of the sizes of galaxies to survive. Larger structures would be created via bottom-up merging, smaller structures via top-down fragmentation. *Advantages:* This would result in a DM mass function that predicts less low-mass DM halos (Bode et al. 2001; see Figure 1.3(a)). The higher thermal motions would also lead to DM halo cores that are less cuspy in their centers as compared to cold DM (Bode et al. 2001; see Figure 1.3(b)).

- *Modified gravity.* Milgrom & Bekenstein (1987) suggested that flat rotation curves and the motions of galaxies in clusters could also be explained by a modification of Newton's law of gravity (MoND; Modified Newtonian Dynamics) that becomes only observable for low accelerations. The advantage of this theory is that it does not require the postulation of new, yet undetected particles. The disadvantage is that it cannot be derived from first principles and its parameters have to be determined empirically. *Counter evidence:* Observations of the Bullet cluster by Clowe et al. 2006, in which the gravitational center of mass and the center of the visible mass are spatially separated (see Figure 1.2(b)), are very difficult to explain by MoND.

Predictions on the nature of DM testable in the MW. Each DM candidate has its own advantages and difficulties. Their characteristics make different predictions for the inner structure and mass function of DM halos. Figure 1.3 (taken from Ostriker & Steinhardt 2003) compares the predictions for cold DM with warm DM and BHs. In the long term, observations of stellar motions in the MW will become more precise and reach further into the Galactic bulge and halo. This will allow to resolve the gravitational potential in the inner regions of the Galaxy and to detect possible halo substructure. Ultimately, we hope to rule out some of the possible DM candidates and to understand the properties of the real DM better.

1.1.3 Formation and Population of Galaxies

Condensation of gas halos. Galaxies form in a two-step process: First, the DM collapses gravitationally and clusters via hierarchical merging, and second, baryonic gas condenses into the potential wells of these DM halos (White & Rees 1978; Fall & Efstathiou 1980).
Due to the collisionless nature of CCDM (see Section 1.1.2), the collapse and clustering is dissipationless and the halo settles into a more or less spheroidal shape in an equilibrium state (e.g., Navarro et al. 1996, 1997) according to the *virial theorem*,

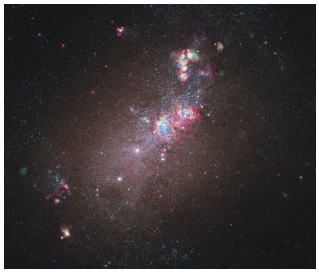
$$-V = 2K \qquad \Rightarrow \qquad \frac{GM_{\rm vir}}{r_{\rm vir}} \sim \sigma^2, \tag{1.2}$$

where V is the gravitational energy and K the internal kinetic energy of the system. $M_{\rm vir}$ is the mass enclosed in the virial radius $r_{\rm vir}$, and σ the particle velocity dispersion, which translates to

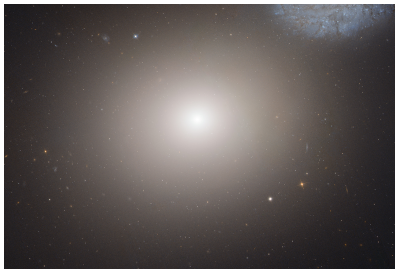
a temperature $T_{\rm vir}$ via $\sigma^2 \sim \frac{k_B T_{\rm vir}}{m}$, with m being the (average) gas particle mass, and k_B the Boltzmann constant.
Gas is not collisionless and experiences during its gravitational collapse heating and cooling processes. Gas inside the virial radius $r_{\rm vir,gas}$ has already settled into hydrostatic equilibrium between the DM halo's gravitational force and the gas pressure ((with $M_{\rm vir} = M_{\rm DM}(r_{\rm vir,gas})$ and $\sigma = \sigma_{\rm gas}$ in Equation 1.2). A gas shell that is still falling in therefore encounters a shock front at $r_{\rm shock} \sim r_{\rm vir,gas}$, which causes the gas shell to heat up to the virial temperature (set by the DM halo's circular velocity curve, $T_{\rm vir} \propto v^2_{\rm circ}(r_{\rm vir,gas}) \sim$ const., see Equation 1.1). Consequently, the pressure rises and the gas shell settles in hydrostatic equilibrium at $r_{\rm vir,gas}$ (Birnboim & Dekel 2003) with $r_{\rm vir,gas}$ growing with time as the virialized gas halo gets more massive.
The gas can cool radiatively: In hot gas ($T > 10^6 - 10^7$ K) the atoms are fully ionized and can only cool via bremsstrahlung. Below this temperature, kinetic energy can be radiated away more efficiently based on collisional ionization/recombination and excitation/de-excitation of atoms, which is even more efficient for gas with higher metallicity (*line cooling*). If the gas cools fast enough after it was shocked, it can get accreted towards smaller radii in the form of cold streams (Dekel & Birnboim 2006).
There appears to be a theoretical critical mass scale of $\sim 10^{12}$ $M_\odot$ above which gas cannot properly cool anymore (e.g., Rees & Ostriker 1977). This could explain the following observations: (i) almost no galaxies above a stellar mass of $M_* \sim 10^{12}$ $M_\odot$ are observed, and (ii) galaxy clusters are embedded in halos made up of hot, X-ray emitting gas.

Birth of the first galaxies. In DM halos where gas can be accreted via cold streams, the gas will continue to radiate away energy and form cold lumps that ultimately settle into a rotation supported disk due to angular momentum conservation (Mestel 1963; Fall & Efstathiou 1980; see also Section 1.1.4). If densities in the gas get high enough and temperatures low enough, stars can form. The first galaxies are born: gas-dominated disk-like collections of young, bluish stars, often small and of irregular shape (see Figure 1.4(a)). These first small irregular galaxies grow into a whole population of galaxies with different morphologies.

Galaxy morphology. Galaxies can be considered as a superposition of stellar components that are in different proportions supported by the random motions or by the rotation of the stars. Components dominated by random motions have spheroidal or triaxial shapes, rotation-dominated components exhibit most often a very flat and disk-like axisymmetric shape. A galaxy with negligible disk components is called an *elliptical galaxy* (see Figure 1.4(b)). A galaxy in which the disk components dominate is called a *disk galaxy* or, because spiral patterns often occur in the disk, *spiral galaxy* (see Section 1.1.4; Figures 1.4(c)-1.4(f)). Galaxies are often embedded in *stellar halos*, sparsely populated extended stellar components with random motions (Zibetti et al. 2004). Dense and kinemtically hot components in the center of disk galaxies are called galaxy *bulges*, and strongly elongated central overdensities *bars* (see Section 1.1.4).
The *Hubble sequence*, historically, sorts galaxies according to how much their disk or spheroidal components dominate their appearance.
It is more convenient to sort galaxies according to their evolutionary stage and think in terms of two growth mechanisms: (i) accretion of cold gas into the disk, and (ii) galaxy mergers. These processes have different effects on the formation of stars and their orbital distribution. Cold gas disks lead to the formation of cold collisionless stellar disk components. Galactic disks are collisionless, i.e., stellar encounters are very rare (see Section 1.4.1) and there is no energy exchange between stars. Hence, any heating of the orbits cannot be reversed. Interactions with neighboring DM halos can trigger instabilities in the disk that lead to the formation of spiral arms and bars (see Section 1.1.4), which in turn dynamically heat the disk orbits. Accretion of dwarf galaxies (*minor mergers*) supplies the halo and bulge with stars on random orbits. *Major*



(a) Dwarf irregular galaxy NGC 4214.



(b) Elliptical galaxy Messier 60.

(c) Edge-on spiral galaxy Messier 104 (Sombrero galaxy).

(d) Interacting grand-design spiral galaxy Messier 51 (Whirlpool galaxy).

(e) Barred grand-design spiral galaxy NGC 1300.

(f) Flocculent spiral galaxy NGC 2841.

Figure 1.4: Examples of galaxy types. Panels (a)-(c): Irregular, elliptical, and spiral disk galaxy. Panels (d)-(f): Various face-on spiral galaxies. **Figure credit:** NASA, ESA, and the Hubble Heritage (STScI/AURA)

mergers, i.e., mergers of galaxies of similar mass, can also completely destroy the disk and result in massive elliptical galaxies dominated by hot, random motions (Toomre & Toomre 1972). If all gas is either ejected or turned into stars during the merger, the resulting galaxy will ultimately become dominated by old, reddish, low-mass stars.

Observed trends. The above picture of galaxy growth is consistent with observed trends with galaxy morphology. In general, when going from galaxies dominated by rotation (pure disks) to galaxies overall dominated by random motions (pure spheroids/ellipticals), there are the following observed trends: increasing age, decreasing gas fraction, decreasing current star-formation, increasing bulge-to-disk ratio, increasing mass, increasing amount of reddish star light, increasing stellar mass-to-light ratio. Spirals can be found more often in low-density regions (groups and fields) or early in the universe, and ellipticals in dense regions (clusters) in the later universe where galaxy mergers are more common and have already happened more often. (Mo et al. 2010)

Supermassive black holes. It is expected that all massive galaxies host SMBHs in their centers (Kormendy & Richstone 1995; Ferrarese & Ford 2005). The origin of SMBHs is still not known, but there is strong evidence for their existence: The most luminous objects in the universe that can be observed even at high redshifts are quasars, which are a sub-class of Active Galactic Nuclei (AGNs). The radiation of AGNs is believed to be caused in the centers of galaxies by SMBHs ($M_\bullet > 10^5$ $M_\odot$) that accrete matter from a surrounding gas disk.
There appear to be tight relations between the SMBH mass, $M_\bullet$, the velocity dispersion in the galaxy spheroid around it (Ferrarese & Merritt 2000; Tremaine et al. 2002) and its stellar mass (Häring & Rix 2004; Magorrian et al. 1998). This indicates that the formation and growth of the SMBH and the stellar spheroid are tightly correlated.

1.1.4 The Evolution of Disk Galaxies

Inside-out growth. It depends on the angular momentum of accreted gas lumps in the DM halo where they end up within the galaxy. A primordial spheroidal gas shell can be thought of as being in solid body rotation (e.g., Dalcanton et al. 1997). Gas lumps originating from the poles of the gas shell have low angular momentum and will settle in the central bulge region of the galaxy during the cooling process. Gas lumps originally located close to the gas shell's plane of rotation have higher angular momentum and settle into the galaxy disk.
Gas shells that get virialized at the shock radius $r_{\rm shock}(t) \sim r_{\rm vir,gas}(t)$ later in time, have larger angular momentum[1], $L \sim r_{\rm shock} \times v_{\rm circ,DM}(r_{\rm shock})$ because $r_{\rm shock}$ grows with time (see Section 1.1.3) and the DM halo's circular velocity is approximately constant (see Equation 1.1). The gas disk grows therefore at its outer edge, where subsequent gas shells of increasing angular momentum settle, from the *inside-out* (e.g., Kepner 1999).
The scale length of the disk depends on the initial spin of the DM halo (Mo et al. 1998).

Star formation. Stars form via fragmentation and collapse of cool molecular gas clouds with radius r, temperature T, and masses higher than the Jeans mass,

$$M_{\rm Jeans} = \frac{5 k_B T}{G m} r \tag{1.3}$$

which follows from the virial theorem Equation (1.2) and the collapse condition $-V > 2K$ for a spherical gas cloud (see, e.g., Carroll & Ostlie 2007, §12.2).
The star formation efficiency scales with the gas density as $\dot{\rho}_* \propto \rho_{\rm gas}^n$ with $n \sim 1.5$ (Schmidt 1959; Kennicutt 1998). Taken together with the inside-out growth, the central bulge and central disk regions will first and most efficiently form stars. The peak of cosmic star formation occurs at a redshift of $z \sim 2$ and the majority of stars ends up on near-circular orbits within the stellar disk. Only a small fraction of stars is already formed during the initial cloud collapse before the in-spiraling gas lumps reach the plane of the disk (Eggen et al. 1962). These stars are on very random orbits, often in globular clusters, and make up the low-mass stellar halo of the galaxy.

Properties of stellar disks. Many disk galaxies are overall well-fitted by an exponential surface brightness profile (Freeman 1970) $\Sigma_*(R) \propto \exp\left(-R/R_{\rm s,disk}\right)$ with a single disk scale length $R_{\rm s,disk}$ (where R is the 2D polar radius in the disk). Only in the bright central bulge region and at large radii, $R \gtrsim (2.5 - 4.5) \cdot R_{\rm s,disk}$, the $\Sigma_*(R)$ profile deviates from the single exponential and drops steeper with R (Pohlen et al. 2000; de Grijs et al. 2001). The origin of exponential disks is still disputed (see Section 1.1.5).
Within the stellar disks, radial color gradients are observed (de Jong 1996b), most likely being the result of stellar populations of different chemical abundances and formation histories (see Section 1.2.3) and dust obscuration.
The disk scale heights are overall independent of radius. A commonly used functional form to model the vertical profile is the self-gravitating isothermal sheet, $\rho_*(z) \propto \mathrm{sech}^2\left(-|z|/z_{\rm s,disk}\right)$, whose vertical velocity dispersion $\sigma_z(z) \sim$ const (Spitzer 1942; Mo et al. 2010, §11.1.1). The typical scale height vs. length ratio is $z_{\rm s,disk}/R_{\rm s,disk} \sim 0.1$. Many galaxies appear to have two disk components, a thick and a thin disk (see Section 1.2.2).
The disks are dominated by rotation of stars and gas in the plane with an angular frequency $\Omega(R) = v_{\rm circ}(R)/R$, where the circular velocity curve $v_{\rm circ}(R) \sim$ const. (see Figure 1.2(a)) is a measure of the enclosed mass (see Section 1.1.2 and Equation 1.1).

[1] Throughout this work we will use the terms "angular momentum L_z", "energy E", and "action J_i", but refer to their specific values, i.e., we give these quantities in units of unit mass.

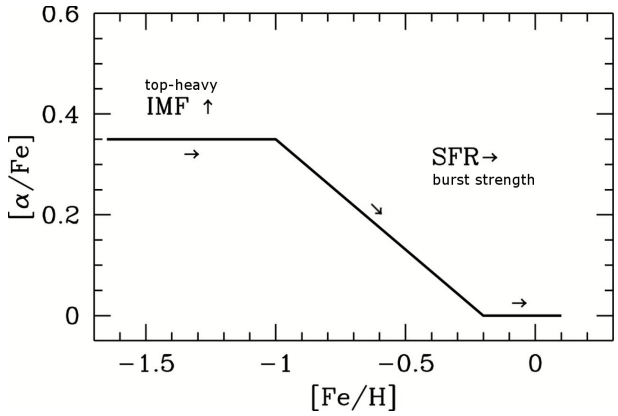


Figure 1.5: Schematic evolution of the chemical enrichment of star-forming gas in galaxies in the [α/Fe] vs. [Fe/H] plane due to SN Type II and Type Ia. See description in the text in Section 1.1.4. The path depends on the IMF and the Star Formation Rate (SFR) of the stellar population considered. **Figure credit:** McWilliam (1997).

There is a tight relation between the total luminosity L of spiral galaxies and their maximum rotation velocity $v_{\rm circ,max}$, $L \propto v_{\rm circ,max}^{\alpha}$, with $\alpha = 2.5 - 5$ (Pierce & Tully 1992). This so-called *Tully-Fisher relation* (Tully & Fisher 1977) is used as a distance estimator for disk galaxies, and indicates that there is a fundamental relation between the stellar and the total dynamical mass of the galaxies (see Section 1.1.5).

Gas enrichment through supernovae. Massive stars ($M \gtrsim 8{\rm M}_\odot$) end their life after less than 10^7 yrs in core-collapse *SNe Type II*; the central regions of the star collapse into a BH or neutron star, while the outer gas shell is blown away explosively. This gas envelope contains elements lighter than silicon (Si)—among them also several α-elements like oxygen (O) and magnesium (Mg)—that were created in stellar nucleosynthesis and which are now ejected into the Interstellar Medium (ISM) by the SN. The amount of ejected O depends strongly on the mass of the progenitor star (see, e.g., Woosley & Weaver 1995; Wyse & Gilmore 1992). The explosion also leads to the rapid nucleosynthesis of some of the Si into small amounts of heavy elements like iron (Fe) in the shocked envelope gas. The Initial Mass Function (IMF) of a stellar population determines the ratio [α/Fe] of ISM gas enriched by SN II: the more massive stars there are, the higher the [α/Fe] ratio. While over time the [Fe/H] of the ISM increases, [α/Fe] stays constant because SN II is currently the only formation channel for Fe and α (see the first stage of chemical evolution in Figure 1.5). (Mo et al. 2010, §10.4.1(b) and 10.4.4)
After ~ 1 Gyr, another channel for the enrichment of the ISM gas with Fe and α elements becomes important (e.g., Matteucci & Recchi 2001; Dahlen et al. 2008; Maoz et al. 2011). This causes the "knee" in Figure 1.5. A star with $M \sim 3$ M$_\odot$ ends up as a low-mass Carbon/Oxygen (C/O) white dwarf. If it has a close companion, it accretes gas from the companion's envelope until it reaches the Chandrasekar mass of 1.4 M$_\odot$. Above this mass limit, electron degeneracy pressure in the white dwarf cannot prevent gravitational collapse anymore. The white dwarf becomes unstable and explodes in a *SN Type Ia*. In the explosion, C and O is turned into iron-peak elements (e.g., ^{56}Ni) which then decay radioactively into Fe. During a SN Ia, ~ 0.75 M$_\odot$ of Fe are ejected into the ISM. This increases [Fe/H] and decreases [α/Fe], leading to the second stage of chemical evolution in Figure 1.5. (Mo et al. 2010, §10.4.1(c) and 10.4.4)
The position of the "knee" (in Figure 1.5) and the relation between [α/Fe] and the age of the population depends on the star-formation history (e.g., Gilmore & Wyse 1991): If all stars were created in a single burst, a lot of SNe II can explode in the first $\sim 10^8$ yrs and [Fe/H] can become already quite high before the SNe Ia set in. If the stars were created over a longer time period with low star-formation efficiency, only a few SNe II raised [Fe/H] to a moderate level, before the first SNe Ia go off.
In the long term (and if no new stars are formed), the enrichment of the ISM gas settles at a [α/Fe] ratio set by the formation channel of SN Ia.
The heavy elements in the ISM can also conglomerate into dust particles, which often trace

regions of strong star formation like spiral arms. Dust strongly shapes the appearance of spiral galaxies (see Figures 1.4 and 1.6).

Spiral arms and bars. Many disk galaxies show distinctive substructure in their disks: two-fold grand-design spiral arms (see Figures 1.4(d) and 1.4(e)) or flocculent spiral arms (see Figure 1.4(f)), and—in more than 50% of the disk galaxies—also central bars (see Figure 1.4(e)).
Bars appear as massive elongated highly flattened stellar overdensities that rotate—as opposed to the rest of the disk—as solid bodies with a certain pattern speed. Bars are expected to form when the the disk gets globally unstable and causes stars to move on more elongated orbits aligned with the bar (e.g., Dehnen 2000). Bars can have profound effects on the stellar dynamics in the disk (see Section 1.4.5).
Spiral arms are overdensities that show up especially strong in young stars and gas, but also in old stars. Hence, they seem to be true density perturbations in the overall disk. Most spiral arms are trailing, i.e., they wind in the same direction as the galactic rotation. What exactly causes these disk perturbations, if they are long-lived or transient, is still an open question (see discussion in Section 1.1.5). In any case, spiral arms appear to be the main sites for star-formation.

Formation of bulges. Many disk galaxies are observed to have bulges in their centers, dense spheroids consisting of mainly old, reddish, metal-rich stars. There are three mechanisms that are expected to play a role in the formation of bulges. (a) Some bulges in more massive spirals, called *classical bulges*, share characteristics with elliptical galaxies, i.e., they have similar relations between luminosity, color, metallicity, stellar mass, and SMBH mass (Bender et al. 1992; Balcells & Peletier 1994; Jablonka et al. 1996; Wyse et al. 1997; see also section on SMBHs in Section 1.1.3). They might, therefore, form in a similar way via gas-rich mergers. The disks either survived because they were already massive enough, or are newly formed via accretion of gas. (b) Central bars can create small bulges in secular evolution processes, so-called *pseudo-bulges*. Angular momentum transfer in the disk due to the bar (see Section 1.4.5) can fuel gas into the galactic centers. Vertical bending instabilities heat up the bar orbits vertically. This creates pseudo-bulges with a boxy peanut shape (Combes et al. 1990; Raha et al. 1991; Athanassoula 2005), like in the MW. Pseudo-bulges are still dominated by rotation. (c) Accreted globular clusters and low-mass dwarf galaxies can fall into the galactic center via tidal friction within the halo and help to build up the central bulge (Aguerri et al. 2001; Walker et al. 1996).

1.1.5 Open Questions about Disk Galaxies

Scaling. Not many galaxies with stellar mass $M_* \sim 10^{12}$ $M_\odot$ are observed. In spiral galaxies the observationally found Tully-Fisher relation (see Section 1.1.4) suggests that there is a correlation between the stellar and total galaxy mass. Galaxy formation theory, therefore, needs to explain the relation between (i) the theoretically predicted mass function of DM halos in the CCDM hierarchical merger paradigm (Press & Schechter 1974; see also Figure 1.3(a)) and (ii) the observed galaxy luminosity function of shape $\phi(L) \sim L^\alpha \exp(L/L^*)$ (Schechter 1976), with $L^* \sim 10^{11}$ $M_\odot$. The theoretical stellar-to-halo mass ratio (Moster et al. 2010, 2013) has its peak around $M_{\mathrm{vir,DM}} \sim 10^{12}$ $M_\odot$; in low-mass galaxies, star-formation is expected to be suppressed by stellar winds and SN, in high-mass galaxies due to virial shocks, inefficient gas cooling, and AGN feedback (Somerville et al. 2008; see also Section 1.1.3). Mergers can trigger star-formation or eject gas. The exact interplay of all these aspects is not yet fully understood, but best investigated in disk galaxies that are not yet "red and dead" and that still have ongoing star-formation and feedback processes.

Origin of exponential disks. The origin of the observed exponential disk profiles (Freeman 1970) is still disputed. Several aspects are suspected to play a role: (i) The angular momentum

distribution of the primordial accreted gas (see Section 1.1.3) matters for the settling of the original gas disk (Fall & Efstathiou 1980; Dalcanton et al. 1997; Mo et al. 1998; Dutton 2009). But which fraction of the gas ends up in the bulge or is ejected? (ii) Subsequent dynamical evolution matters, too, during which angular momentum is re-distributed between material. This can happen (a) through viscosity of the star-forming gas disk (Lin & Pringle 1987; Pringle 1981), or (b) through radial migration of stars due to bar and spiral arm perturbations (Sellwood 2014; see Section 1.4.5). Herpich et al. (2017) showed that radial migration as a statistical diffusion mechanism results in approximately exponential disk profiles. What is the interplay and relative contribution of the different mechanisms?

Disk maximality. Related to the problem of the true distribution of DM is the question about disk maximality. The *maximum-disk hypothesis* assumes that the disk contributes the maximum possible amount to the measured rotation curve. This is used in external galaxies to decompose the contributions of disk and halo to the rotation curve (e.g., van Albada & Sancisi 1986; Bell & de Jong 2001). It is disputed if this hypothesis is actually correct (Courteau & Rix 1999).
The maximum-disk hypothesis provides an upper limit for the stellar mass-to-light ratio of galaxy disks. Residuals between the disk contribution and the observed rotation curve would then be attributed to DM (and bulge and halo). If the contribution to v_{circ} at 2.2 times the disk scale length (the appoximate peak of the rotation curve) is $85\% \pm 10\%$, the disk is considered to be maximal (Sackett 1997; Binney & Tremaine 2008, §6.3.3).

Nature of spiral arms. The true nature and origin of spiral arms is still disputed.
The current working theory is that flocculent spiral arms are local matter overdensities that rotate with the underlying disk rotation and get sheared into elongated spiral arms by the differential rotation. They can be caused, for example, by giant molecular clouds in the disk (D'Onghia et al. 2013) and are short-lived.
Grand-design spiral arms cannot be material arms that rotate with the same velocity as stars and gas, otherwise they would quickly become tightly wound around the Galactic Center (GC) and probably break up. Lin & Shu (1964) proposed the *spiral density wave theory*. The spiral arms move as an (almost) solid body through the disk with their own pattern speed. The gas and stars cross the arms at a higher velocity, but linger slightly longer in the arm due to its gravitational attraction. In terms of stellar orbits, the spiral density wave can be considered as overdensities created by nested elliptical orbits in the co-rotating frame.
In the case of flocculent spirals, it is a local gas instability that triggers both the formation of the spiral arm and stars. In the case of a long-lived spiral density wave, it is the global spiral pattern which triggers star formation by creating gas overdensities while sweeping through the disk.
In galaxy simulations, however, transient and recurrent spiral arm features are favoured (e.g., Sellwood 2011). Also, DM halos and bulges should stabilize galaxy disks against global instabilities like bars and spiral arms. It might be, that these disk instabilities are triggered by galaxy encounters (see Figure 1.4(d)).

1.2 The Structure of the Milky Way

The MW, our home galaxy, looks like a milky-white thin band of stars stretching across the whole sky (see Figure 1.6). The Greek word for "milk", $\gamma\acute{\alpha}\lambda\alpha$ (gála), gave our Galaxy (with capital "G") its name and is nowadays also used for other galaxies. The MW is a disk galaxy with presumably $\sim 10^{10} - 10^{11}$ stars. Our Sun is located 25 pc away from the Galactic midplane (Jurić et al. 2008) and approximately 8 kpc from the GC (e.g, Ghez et al. 2008; Gillessen et al. 2009a), which corresponds to approximately 2.5 times the disk scale length. The Sun rotates with approximately 220 km s^{-1} around the GC (e.g., Kerr & Lynden-Bell 1986).

Figure 1.6: Panoramic view of the MW on the night sky. Visible is the thin plane of the stellar disk and the central bulge, obscured by interstellar dust. On the lower right, the Magellanic clouds are visible, two of the dwarf satellite galaxies of the MW. This composite photograph was taken between 2008-2009 by Serge Brunier from the Atacama desert (Southern sky) and La Palma (Northern sky). **Figure credit:** ESO/Serge Brunier (`http://www.eso.org/public/images/eso0932a`).

The illustration in Figure 1.10 demonstrates the main stellar components of the MW: a disk threaded by spiral arms, with an elongated central bulge and bar in the center. The MW disk is rich in structure and sub-structure considering the ages, chemical abundances and motions of stars. In this section, we give an overview of the present day configuration of the MW, with focus on the stellar disk.
The main references for this section are the reviews on the MW structure by Bland-Hawthorn & Gerhard (2016) and the MW disk by Rix & Bovy (2013), as well as references therein.

1.2.1 Stellar Spheroids in the Galaxy

Bar. The MW has a long bar in its center (Liszt & Burton 1980; Binney et al. 1991, 1997; Blitz & Spergel 1991; Hammersley et al. 2000; Wegg & Gerhard 2013) which extents out to $R = 5.0 \pm 0.2$ kpc (Wegg et al. 2015). It is assumed that the bar is fast-rotating like a solid body and that the disk substructure in the velocity space of the Solar neighborhood (see Section 1.2.4 and Figure 1.11) is due to resonances with the bar (see Section 1.4.5). This leads to the indirect measurement of a pattern speed $\Omega_b = 43 \pm 9$ km s^{-1} kpc^{-1} (see Bland-Hawthorn & Gerhard 2016, §4.4, for a review). The bar has a scale height of only 180 pc, suggesting that it is part of the disk and was probably formed from disk instabilities (see Section 1.1.4). Its orientation with respect to the line-of-sight from the Sun to the GC is $28° - 33°$ (Wegg et al. 2015). In projection, as seen from the vantage point of the Sun, this causes the peanut-shaped appearance of the central regions of the MW.

Bulge. The MW's central bulge is a pseudo-bulge (see Section 1.1.4) that has a boxy-peanut shape (Weiland et al. 1994; Binney et al. 1997; Skrutskie et al. 2006) and can be considered the central 3D structure of the rotating bar (Nataf et al. 2010; Wegg & Gerhard 2013). It is quite small with scale lengths $(h_x, h_y, h_z) = (0.70, 0.44, 0.18)$ kpc (Bland-Hawthorn & Gerhard 2016) and has a total stellar mass in the bulge region of $(1.4 - 1.7) \times 10^{10}$ M$_\odot$ (Portail et al. 2015).

Table 1.1. Summary of literature estimates on some of the structural parameters of the MW quoted throughout this work (see also review by Bland-Hawthorn & Gerhard 2016, B-H&G16).

Quantity		Literature estimate	Reference
Thin disk scale length	$R_{\mathrm{s,thin}}$	2.6 ± 0.5 kpc	B-H&G16, §5.1.2
Thick disk scale length	$R_{\mathrm{s,thick}}$	2.0 ± 0.2 kpc	B-H&G16, §5.1.2
Thin disk scale height	$z_{\mathrm{s,thin}}$	300^{+150}_{-80} pc	B-H&G16, §5.1.1
Thick disk scale height	$z_{\mathrm{s,thick}}$	900 pc $\pm 20\%$	Jurić et al. (2008)
Thick-to-thin disk surf. dens. ratio	$f_\Sigma(R_\odot)$	$12\% \pm 4\%$	B-H&G16, §5.1.3
Circular velocity at the Sun	$v_{\mathrm{circ}}(R_\odot)$	238 ± 15 km s^{-1}	B-H&G16, §6.4.2
Slope of the rotation curve	$\mathrm{d}(\ln v_{\mathrm{circ}})/\mathrm{d}(\ln R)\vert_{R_\odot}$	-0.06 ± 0.05	Bovy & Rix (2013)
Total surface mass density	$\Sigma_{\mathrm{tot,1.1kpc}}(R_\odot)$	70 ± 5 M$_\odot$pc^{-2}	B-H&G16, §5.4.2
Disk maximality	$\frac{v_{\mathrm{circ,disk}}}{v_{\mathrm{circ,tot}}}\Big\vert_{R=2.2R_{\mathrm{s,disk}}}$	$0.83 \pm 0.04^{(a)}$	Bovy & Rix (2013)
Total stellar mass	M_*	$(5 \pm 1) \times 10^{10}$ M$_\odot$	Bovy & Rix (2013), B-H&G16, §6.4.4
Stellar mass of the bulge	$M_{*,\mathrm{bulge}}$	$(1.4 - 1.7) \times 10^{10}$ M$_\odot$	Portail et al. (2015)
DM halo mass in r_{200}	$M_{\mathrm{DM},200}$	10^{12} M$_\odot$	B-H&G16, §6.3
DM density at the Sun	$\rho_{\mathrm{DM},\odot}$	0.0065 ± 0.0023 $M_\odot$ pc^{-3}	Zhang et al. (2013b)
Gas surface density at the Sun	$\Sigma_{\mathrm{gas}}(R_\odot)$	13.2 $M_\odot$ pc^{-2}	Flynn et al. (2006)
Bar pattern speed	Ω_b	43 ± 9 km s^{-1} kpc^{-1}	B-H&G16, §4.4

$^{(a)}$At 2.2 times a scale length of 2.15 kpc.

The bulge is composed of old stars (Ortolani et al. 1995) on box orbits in the triaxial potential of the bar. The stars have a range of metallicities (Zoccali et al. 2003): High metallicities indicate a fast enrichment history due to high gas densities early on in the center of the Galaxy. The origin of the low-metallicity stars is disputed (Bland-Hawthorn & Gerhard 2016). If the MW also contains a classical bulge component, formed during the initial collapse of the gas or through the accretion of satellites, is not clear. It is not expected that more than 25% of the bulge mass would belong to this classical bulge sub-component (Shen et al. 2010).
The innermost regions of the bulge contain a SMBH, Sagittarius A* (Balick & Brown 1974), which has a mass of $M_\bullet \sim 4.2 \times 10^6$ M$_\odot$ (Ghez et al. 2008; Gillessen et al. 2009b,a), and a nuclear star cluster with a half-light radius of only ~ 4.2 pc (Schödel et al. 2014).

Stellar Halo. The stellar halo is an extended low-mass spheroidal component in the MW, that reaches beyond 40 kpc. The stellar density falls off with (spherical) radius as a power-law and the motions are random, with a net-rotation of only ~ 40 km s^{-1}. The halo stars make up only 1% of the MW's stellar mass (e.g., Bland-Hawthorn & Gerhard 2016). The stars are expected to come from two different origins (Carney 2001): (i) Part of the halo stars formed in-situ early on together with the stellar bulge, but with a much lower metallicity (Eggen et al. 1962; Ortolani et al. 1995). (ii) A significant amount of stars was accreted from satellite galaxies (e.g., Searle & Zinn 1978; Helmi et al. 1999; Bell et al. 2008). Most of the halo stars are in substructures, in globular clusters, shells, and streams (e.g., Belokurov et al. 2006), which most likely used to be stellar clusters and dwarf galaxies that got stripped in the tidal field of the Galaxy. The most famous and massive stellar halo stream is the Sagittarius stream associated with the Sagittarius dwarf galaxy (Ibata et al. 1994; Newberg et al. 2002, 2003; Majewski et al. 2003).

1.2.2 The Disk

Historical definition of thin and thick stellar disk. The disk surface brightness profiles in many external galaxies appear to be best described by two exponential disk components of different scale heights (Yoachim & Dalcanton 2006). Gilmore & Reid (1983) found that, in the MW, the

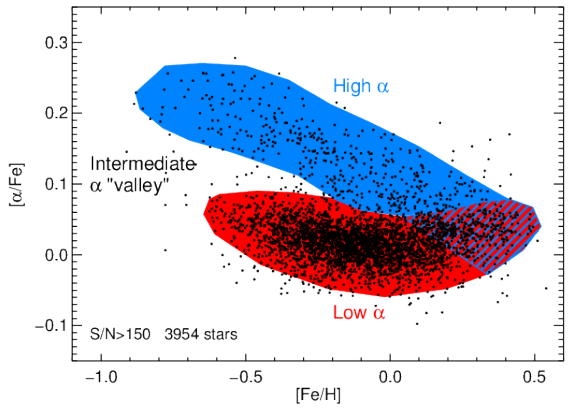


Figure 1.7: Distribution of APOGEE RC stars (Bovy et al. 2014) in the Solar neighborhood in the [α/Fe] vs. [Fe/H] plane. There are two chemically distinct populations, the high-α and the low-α sequence, which merge at the high-metallicity end. Historically, these two sequences correspond to the "thick" and the "thin" disk in the MW. **Figure credit:** Nidever et al. (2014).

stellar number counts toward the Galactic South pole also allowed the fit of two exponential profiles, with scale heights of ~ 300 pc and ~ 1 kpc. Given that the stars in the two components appeared to have also approximately different ages, chemistry (e.g., Fuhrmann 1998; Gratton et al. 2000), and kinematics, motivated the introduction of geometrically defined "thin" and "thick" disk components in the MW. The thick disk is on average older (~ 12 Gyr as determined from the main-sequence turn-off point by, e.g., Gilmore et al. 1995), more α-rich and metal-poor, and kinematically hotter than the thin disk.

A better definition: abundance and age. While the names "thin and thick disk" are still used, they have turned out to be partly misleading. The distribution of stars in the disk appears to have continuously varying properties with chemical abundance, with scale heights in the range $\sim 200 - 1000$ pc Bovy et al. (2012d,b) and vertical velocity dispersions $\lesssim 35$ km s^{-1} (Norris 1987) (see also Section 1.2.3). There is therefore no clear bimodality in the phase-space distribution.

Better names would be "low-α"/"high-α" or "α-young"/"α-old" disk, because the bimodality in chemistry and also age between the two disk components is more distinct and appears to be real (e.g., Haywood et al. 2013; Bensby et al. 2014; Nidever et al. 2014; Hayden et al. 2015; Masseron & Gilmore 2015). Also, age and chemistry are labels that are set by the birth of a star and cannot change drastically during Galaxy evolution[2] as compared to the spatial distribution. In the remainder of this work, we will use "thin/thick" interchangeably for "low-α/high-α" for historical reasons.

Figure 1.7 shows the distribution of Red Clump (RC) stars in the [α/Fe] vs. [Fe/H] plane (taken from the Sloan Digital Sky Survey III's Apache Point Observatory Galactic Evolution Experiment (APOGEE)), which demonstrates the clear bimodality in α. The two sub-populations merge at the high-metallicity end. If we compare this with the process of chemical enrichment described in Figure 1.5 and Section 1.1.4, this clearly suggests that the high-α disk stars must have formed earlier (probably around the peak of cosmic star formation at $z \sim 2$, i.e., $t_{\rm age} \sim 10 - 12$ Gyr), and had a faster enrichment history and therefore higher star formation efficiency (Wyse & Gilmore 1993; Nidever et al. 2014) than the low-α disk. The ages of the low-α disk stars span a range of < 1 Gyr to $10 - 12$ Gyr. The thin disk contains more stars than the thick disk; the thick-to-thin disk surface mass density ratio is $\sim 8 - 16\%$ (see Table 1.1 and Bland-Hawthorn & Gerhard 2016). The formation origin of the thick disk is still disputed; we give an overview in Section 1.2.5.

[2]Some chemical abundances do, however, change during the course of stellar evolution, e.g., carbon (C) and nitrogen (N). Matter, which has been processed by the CNO cycle and has therefore a different [C/N], is carried from the central regions to the surface during the first dredge-up on the giant branch (e.g., Martig et al. 2016).

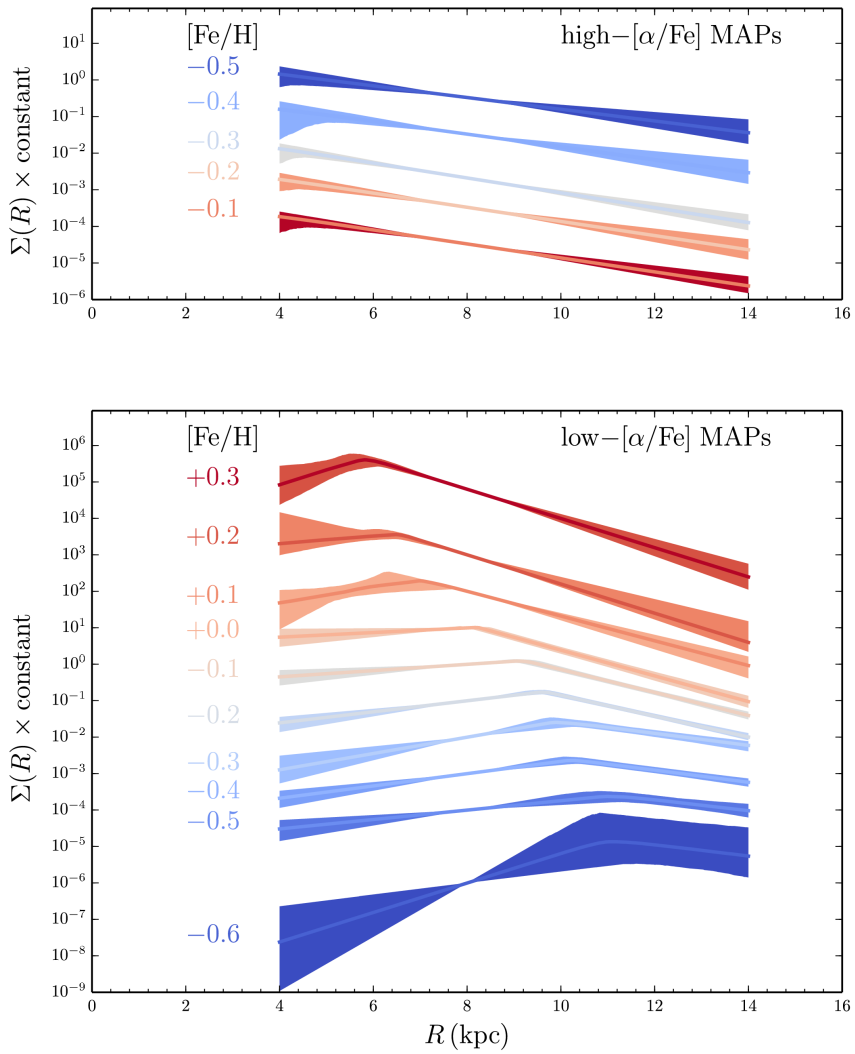


Figure 1.8: Stellar surface density profiles for the SF-corrected distribution of APOGEE RC giants in different abundance bins in. High-α MAPs follow single exponential profiles with scale lengths of $h_R \sim 2$ kpc, while low-α MAPs are best fitted by two exponentials, one rising, one falling, and a break radius. The break radius depends on the metallicity of the thin disk MAP, with more metal-rich stars being more centrally concentrated. **Figure credit:** Bovy et al. (2016b).

Gas disk. In addition to the bimodal stellar disk, the Galaxy also contains a thin layer of cold atomic ($\sim 80\%$) and molecular ($\sim 20\%$) hydrogen gas on circular orbits, with traces of higher elements and dust. This is the ISM; its surface density at the Sun is $\sim 13\ M_\odot\ \mathrm{pc}^{-2}$ as compared to $\sim 29\ M_\odot\ \mathrm{pc}^{-2}$ in stars (Flynn et al. 2006).

1.2.3 Stellar Populations in the Disk

Mono-Abundance Populations (MAPs). In the last few years, MAPs have provided insight into the structure and formation history of the MW disk (e.g., Bovy et al. 2012d,c,b, 2016b). A MAP is defined as the sub-population of disk stars that have the same [α/Fe] and [Fe/H] abundance. MAPs are not equivalent to mono-age populations, with the exception of the highest [α/Fe] MAPs (Minchev et al. 2017). Mono-age populations might be the more intuitive choice to track the formation history of the disk (Bird et al. 2013; Martig et al. 2014), but chemical abundances can be determined more reliably and precisely than stellar ages. The investigation of

MAPs is purely empirically motivated and does not assume that stars within a MAP were born together. It does, nevertheless, unveil the interesting chemical sub-structure of the Galactic disk. The study by Bovy et al. (2012d) revealed that the spatial number density distribution of each MAP of G-dwarf stars in the Sloan Extension for Galactic Understanding and Exploration (SEGUE) survey can be well-fitted by a single exponential in both Galactocentric R and z direction, with a smooth range of disk scale heights $h_z \in [200, 1000]$ pc and scale lengths $h_R > 1.5$ kpc (see Figure 4.16(b) in Chapter 4).
Bovy et al. (2012c) extended the discussion by also considering the vertical motions of the MAPs (see Figure 4.17(b)). They found vertical velocity dispersions for the MAPs in the range $15 - 30$ km s^{-1} for the thin and $32 - 52$ km s^{-1} for the thick disk (similar to Binney 2012b), which appears also to be independent of the height above the plane (see Figure 4.17(b)). This motivates the description of the MW disk as an "isothermal sheet" (Spitzer 1942; Mo et al. 2010, §11.1.1; see also Section 1.1.4) which assumes $\sigma_z(z) \sim$ const. The velocity dispersion decreases radially; it is, however, not fully clear how this decrease depends on the chemical abundances (see Figure 4.18(b)).
There is a clear smooth trend between [α/Fe], h_z, and σ_z: The more α-rich the MAP, the larger h_z and σ_z. In addition, there is also a continuous anti-correlation between h_R and h_z, i.e., MAPs are either centrally concentrated and kinematically hot (high h_z and σ_z), or concentrated in the plane of the disk with a long scale length. Most MAPs that traditionally belong to the thick/high-α/low-metallicity disk have very similar scale lengths around $h_R \sim 2$ kpc. The thin/low-α/high-metallicity MAPs span a wider range of scale lengths, between 2 kpc and > 5 kpc.
Bovy et al. (2016b) complemented this picture on the basis of APOGEE RC stars in the radial range $R \in [4, 14]$ kpc. They confirmed the previous findings for the thick/high-α disk (see Figure 1.8, upper panel). For the thin/low-α disk stars, they found, however, that the single exponential profile previously assumed does not fully describe the spatial distribution of MAPs. These populations rather resemble "donuts" around the GC that have a peak surface density at a certain break radius $R_{\rm break}$ (see Figure 1.8, lower panel). At smaller radii, the surface density rises, at larger radii it drops exponentially with R. There is a correlation between the break radius and the metallicity of the MAP, with the more metal-rich populations having the break radius at smaller R and are therefore more centrally concentrated. Mackereth et al. (2017) found, based on APOGEE stars and ages from Martig et al. (2016), that the width of the distribution around $R_{\rm break}$ becomes wider with age.
Stars in the thin disk are in general on more circular orbits than the thick disk stars. The radial velocity distribution of MAPs is, however, not as well studied as the spatial distribution and vertical motions. The radial velocity dispersion at the Sun is expected to be in the range $\sigma_R \sim (35 \pm 5)$ km s^{-1} and $\sigma_R \sim (50 \pm 5)$ km s^{-1} for the thin and thick disk, respectively (see compilation by Bland-Hawthorn & Gerhard 2016, §5.4).

Galaxy formation context. The observed structure of the MW's disk is overall consistent with the inside-out galaxy formation scenario that we sketched in Section 1.1.4, and can be explained in terms of chemical gas enrichment (e.g., Matteucci & Francois 1989; Chiappini et al. 2001; Minchev et al. 2017; Schönrich & McMillan 2017), also sketched in Section 1.1.4. The highest gas densities will result in the earliest and largest star-bursts; the α-old stars are therefore most centrally concentrated. The first stars enrich the gas from which subsequent stellar generations are born. With increasing radius, the gas surface density and therefore the initial star burst strength decreases. In addition, new primordial gas is mostly accreted at the edge of the disk. At each radius an equilibrium [Fe/H] is established between Fe enrichment and H gas accretion, with [Fe/H] decreasing outwards. Bovy et al. (2016b) suspect, that the break-radius in the low-α populations corresponds to this equilibrium radius. Subsequent radial re-distribution of stars

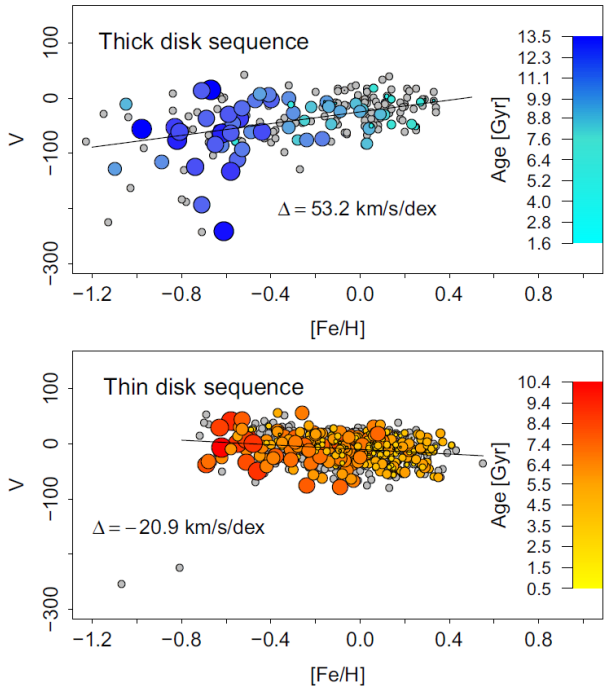


Figure 1.9: Heliocentric tangential velocity in the direction of Galactic rotation V vs. the metallicity [Fe/H] of the stars in the Solar neighborhood, color-coded by age. Haywood et al. (2013) has separated the stars according to thick and thin disk, demonstrating that metal-poor thick disk stars rotate slower in the solar neighborhood. The reason for the observed trend is explained in the text. **Figure credit:** Haywood et al. (2013).

lead to the observed donut-structure of these MAPs that gets wider with age (Mackereth et al. 2017).

The vertical structure is set by the star formation history of the gas disk and subsequent heating mechanisms. The overall disk exhibits vertical age (e.g., Martig et al. 2016) and metallicity gradients; in the MAP picture this is equivalent to h_z varying smoothly with [α/Fe] and [Fe/H] (see Figure 4.16(b) from Bovy et al. 2012d). If we consider [α/Fe] as being an approximate age indicator (e.g., Haywood et al. 2013), and [Fe/H] as a measure for the gas density and therefore burst strength of the birth location (see Figure 1.5), this suggests continuous rather than singular formation and heating events in the whole disk (see also the overview of thick disk formation scenarios in Section 1.2.5).

It has to be noted again that MAPs are no mono-age populations. In the solar vicinity there is a large scatter between age and metallicity (Edvardsson et al. 1993; Nordström et al. 2004). And within a given MAP, there can be negative radial age gradients, with stars at the same metallicity having formed either early at smaller radii, or later at larger radii (Minchev et al. 2017).

Simulations suggest that all mono-age populations flare, i.e., $h_z(R)$ rises with radius. This is suspected to be a consequence of secular evolution (i.e., radial migration) and environmental effects (i.e., interactions with satellites) (Minchev et al. 2015). Bovy et al. (2016b) found flaring only for the low-α MAPs. Minchev et al. (2017) attributed this to mixing of several mono-age populations with different scale lengths into a single MAP, which apparently removes the flare in the observations. This was only recently confirmed by Mackereth et al. (2017).

Excursion: Asymmetric drift in the Solar neighborhood. One well-studied feature of the MW disk is the dependence of the heliocentric velocity V (in the direction of Galactic rotation) on the metallicity of the stars in the Solar neighborhood (Kordopatis et al. 2017; Haywood et al. 2013; Allende Prieto et al. 2016; Wojno et al. 2016), which we show in Figure 1.9. Apparently, the thick disk is rotating slower than the thin disk.

The trends for the high- and low-α disks are easily explained in the MAP context with orbital blurring due to epicyclic motions. High-α/thick disk (low-α/thin disk) stars in the Solar neighborhood come mostly from inside (outside) of the Sun (see Figure 1.8) and are currently at

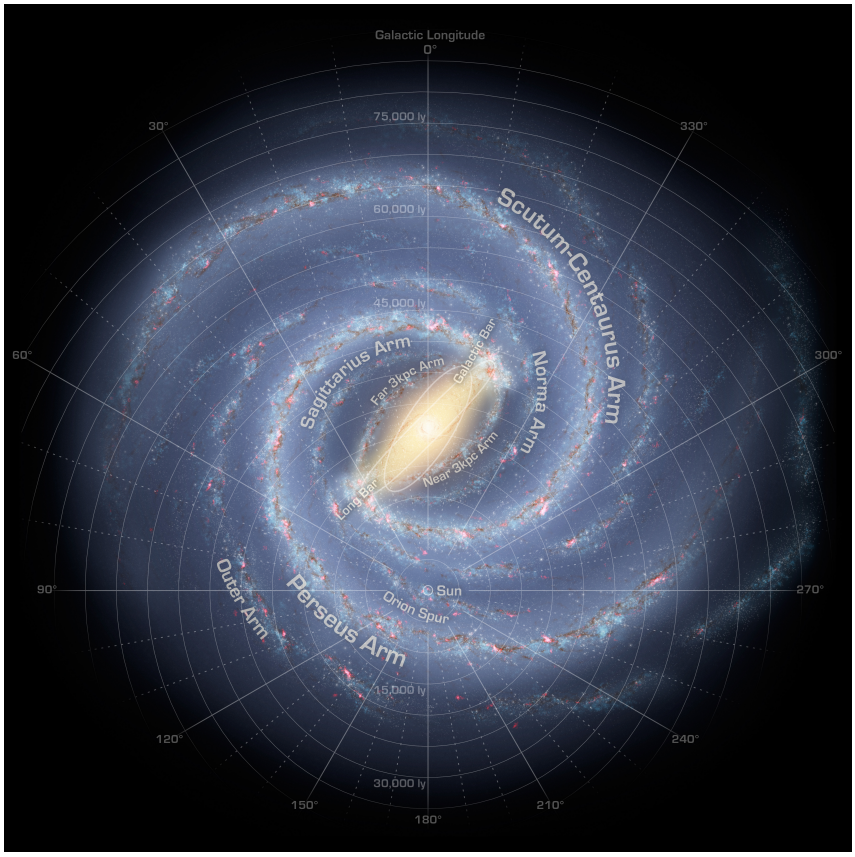


Figure 1.10: Artistic interpretation of the face-on MW as a grand-design barred spiral galaxy based on infrared stellar number counts by Churchwell et al. (2009). The Sun is located in the Local/Orion Arm. The exact number and strength of the spiral arms is not yet fully known (see text in Section 1.2.5). **Figure credit:** NASA/JPL-Caltech/Robert Hurt (`http://www.spitzer.caltech.edu/images/1925-ssc2008-10b-A-Roadmap-to-the-Milky-Way-Annotated`).

their orbital apocenters (pericenters). This reduces (increases) the observed V velocities with respect to the local $v_{\rm circ}(R_\odot)$, a process called *asymmetric drift*.
The stars with the highest [Fe/H] are on the most circular orbits and the observed heliocentric V corresponds to the disk's mean rotational velocity $v_{\rm circ}(R_\odot)$. Stellar populations with lower [Fe/H] have in general higher radial velocity dispersions (see, e.g., Figures 4.17(c) and 4.17(d)), which increases the effect of asymmetric drift. This explains the trend of *decreasing* V with decreasing [Fe/H] in the thick disk. In the thin disk, the decrease of [Fe/H] with Galactocentric radius might be enough to explain the opposite trend of V *increasing* with decreasing [Fe/H]. How much the radial velocity dispersion affects this trend in the thin disk, is not clear.

1.2.4 (Sub-)Structure in the Stellar Disk

Spiral Arms. We see our Galaxy only edge-on and cannot directly observe its spiral structure. From the observation of external galaxies we know that spiral arms are traced by star-forming regions. Star-forming regions emit characteristic light (see, e.g., Carroll & Ostlie 2007, §12.3). Young luminous stars are still shrouded in gas clouds and can act as natural MASER sources. Also, UV photons from the young blue stars ionize the hydrogen cloud and subsequent recombination in these HII regions creates Hα radiation. By measuring the distances to Hα clouds, MASER sources, and young OB-stars, e.g., via geometric parallaxes or stellar spectral models, the approximate spiral structure of the MW can be derived (e.g., Georgelin & Georgelin 1976; Reid et al. 2014). Figure 1.10 shows an illustration of the spiral structure based on star counts and geometric

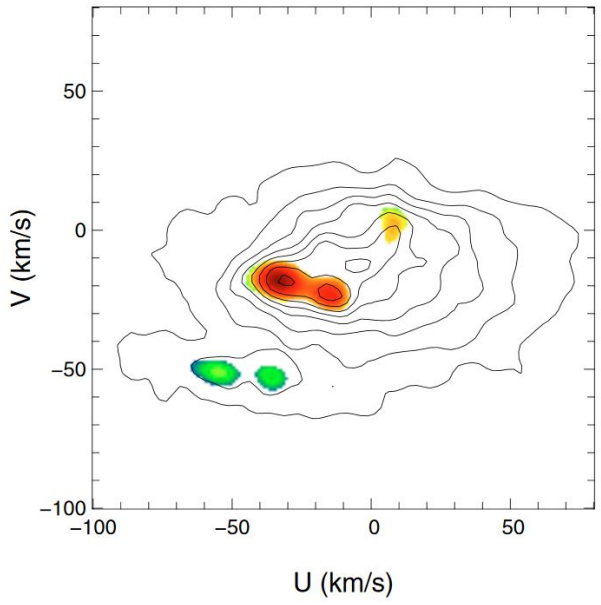


Figure 1.11: Velocity distribution of stars in the local sample of Hipparcos/*Tycho-2* stars by Famaey et al. (2005). U is the velocity in the direction towards the GC and V is in the direction of Galactic rotation, both with respect to the Sun which has $(U, V) = (0, 0)$. In an axisymmetric Galaxy, this distribution would be symmetric around the mean Galactic flow at $(U, V) \sim (-11, -12)$ km s^{-1} (Schönrich et al. 2010), i.e., the heliocentric velocities corrected by the Solar peculiar motion. The overdensities marked in color are called *moving groups* and correspond to streams or groups of stars in the Galactic disk that have unusual velocities. These moving groups are called *Hercules* (green), *Pleiades* (red left), *Hyades* (red right), and *Sirius* (yellow). See text in Section 1.2.4. **Figure credit:** Famaey et al. (2008).

assumptions by Churchwell et al. (2009). The Sun is located in the local Orion arm. Further inside towards the GC is the Sagittarius arm, further out the Perseus arm. Figure 1.10 is slightly misleading; we actually do not know much about the spiral arms of the MW (see Section 1.2.5).

Moving groups. The MW disk contains groups of stars that appear to move together on similar orbits and that stand out from the expected background distribution in density and/or kinematics (e.g., Klement et al. 2008; Bovy et al. 2009). Some examples in the Solar neighborhood's velocity distribution are shown in Figure 1.11. There are several possible origins: They are (i) open clusters, i.e., stars that are not gravitationally bound to each other but which were born together from the same molecular cloud, (ii) disrupted accreted satellites (analogous to streams in the halo; e.g., the Arcturus stream, Navarro et al. 2004), (iii) dynamic perturbations in the disk (e.g, Famaey et al. 2005; Sellwood 2010; Bovy & Hogg 2010).
Some of the moving groups in Figure 1.11 (Hercules, Hyades, Sirius, Arcturus) are heterogenous in chemical abundances and/or age (Famaey et al. 2008; Bensby et al. 2014) and are therefore not expected to have been born together. Instead, these moving groups could be created by resonances in the disk with the rotating Galactic bar (Dehnen 1998, 2000; Bovy et al. 2015; Fux 2001; Monari et al. 2017b; see also Section 1.4.5). Streaming motions measured in stellar velocities could also be caused by spiral arms (Quillen 2003; Quillen & Minchev 2005; Siebert et al. 2012).

Rings and warps. In addition to perturbations in the plane of the disk, like spiral arms, the disk plane itself appears to be subject to perturbations. Ring-like overdensities in the stellar number counts of main-sequence stars in the Sloan Digital Sky Survey (SDSS) have been discovered towards the Galactic anti-center (Newberg et al. 2002; Jurić et al. 2008). They appear to exhibit alternating offsets in north and south direction with increasing radius, which would be consistent with a radial wave of vertical oscillations in the disk (Xu et al. 2015) rather than having an accreted origin (Price-Whelan et al. 2015; Gómez et al. 2016). The most famous of this rings in the Monoceros ring (Rocha-Pinto et al. 2003).
There is also evidence that the Galactic disk is warped outside of the Solar circle (e.g., Poggio et al. 2017), i.e., the tilt angle of the disk varies with radius. This signature has been seen in gas (Burke 1957), as well as stars and dust (Drimmel & Spergel 2001).

1.2.5 Open Questions about the Milky Way's Structure

Overall structure of the disk. There exist photometric measurements of the disk scale length and height from star counts (Jurić et al. 2008), but it is important to compare this with dynamical mass measurements. While the mass-weighted disk scale length has already been measured thanks to the large radial coverage of surveys like SDSS (Bovy & Rix 2013), the disk scale height of the total mass is still not well constrained. Even though it is not expected from mass measurements of the MW disk (e.g., Kuijken & Gilmore 1989a; Holmberg & Flynn 2004), it would be an important test for ΛCDM to check if the disk has a dark component (e.g., Oort 1932; Fan et al. 2013; Randall & Reece 2014; Kramer & Randall 2016).
The relation between positions, velocities, chemical abundances, and ages of stars in the disk are subject of numerous studies, but there is still a lot to learn. Especially the radial kinematics in the disk are less well known than the vertical velocities.

Formation of the thick disk. The origin of the thick disk is a long standing question in galaxy formation theory. There are several proposed formation mechanisms.

(i) The thick disk was formed already as a thick disk...
  (a) ... during the dissipational collapse of the primordial gas cloud out of early turbulent gas (before $z \sim 1$; Gilmore et al. 1989; Bournaud et al. 2009).
  (b) ... from the debris of accreted satellites that deposited their stars high above the Galactic plane (Abadi et al. 2003; this was also suggested as formation mechanism for thick disks in external galaxies, see Yoachim & Dalcanton 2006, 2008).

(ii) The thick disk was formed as a thin disk after the dissipationless collapse of the gas cloud and experienced subsequent vertically heating through...
  (a) ... gradual secular evolution (Schönrich & Binney 2009b; Loebman et al. 2011; Minchev et al. 2015).
  (b) ... minor merger event(s) (Quinn et al. 1993; Villalobos & Helmi 2008).

Prochaska et al. (2000) argue that the abundance patterns of halo and bulge suggest that the thick disk was formed 10 Gyr ago from the same gas reservoir over the time scale of 1 Gyr. Bovy et al. (2016b) reason that the continuum in disk scale heights of MAPs across the disk suggest a gradual heating process. There is, however, no consensus yet on which mechanisms contribute and by how much.

The mass structure of spiral arms. We only see the MW edge-on. The spiral arms of the MW are therefore largely unknown. Neither do we have any reliable measurements for their pattern speed (which is suspected to be slower than the bar; e.g. Gerhard 2011), nor have dynamic features in the stars been attributed beyond doubt to be caused by spiral arms (Rix & Bovy 2013). This is related to the fact, that we do not know how strong the relative mass contrast of the spiral perturbations actually are in the disk.
The exact number of spiral arms in the MW is also still under debate. The distribution of star-forming regions traced, e.g., by H II regions (Georgelin & Georgelin 1976), maser sources (Reid et al. 2009, 2014), and young massive stars (Urquhart et al. 2014), suggest that the MW has four major spiral arms (see also Vallée 2008, 2014, and references therein). Observations in the infrared, e.g., of old RC giant stars in the Spitzer/GLIMPSE survey (Churchwell et al. 2009) or of stellar Near-Infrared (NIR) emission by the COBE satellite (Drimmel & Spergel 2001) indicate that the MW is a grand-design two-armed spiral. One hypothesis is that the four-armed spiral observed in young stars and gas is due to the response of the gas to the two-armed spiral in old stars (Drimmel 2000; Martos et al. 2004).

Star formation history. One of the key research questions in the field of galaxy formation, is to understand galactic star formation histories. In the MW, where we have chemical abundance measurements of many individual stars, this can be studied exceptionally well. If continually new stars were formed from the enriched ISM (see Section 1.1.4), the chemical abundances of stars observed today would mirror the chemical evolution path of the gas in Figure 1.5, with [α/Fe] being closely correlated with the age of the stars. [Fe/H] is in general also an age-indicator but is more dependent on the star-formation efficiency and history, which in turn depends strongly on the position within the galaxy. All of this can be put together to derive the MW star formation history analogous to the sketch in Section 1.2.3.
As star-formation is still on-going in the MW disk, the disk allows also to study the interplay of stellar formation and feedback, gas in- and outflows. Understanding the star-formation efficiency in the MW might help to understand why it is less efficient in other galaxies.

The role of radial migration. Radial migration is an important part of the secular evolution of galaxies. It is a process in which stars change the angular momentum of their orbit, and therefore their mean radius (called *churning*), but not necessarily their eccentricity (see, e.g., Sellwood & Binney 2002). It can be induced at resonances of non-axisymmetric, periodic perturbations like bars and/or spiral arms (e.g., Minchev et al. 2011)—we will look at this process in more detail in Section 1.4.5—, but also by interaction with satellites (Bird et al. 2012). It is thought that radial migration might at least conserve the vertical action (see Section 1.4.3) of the bulk of orbits (Minchev et al. 2012; Solway et al. 2012; Vera-Ciro & D'Onghia 2016). At larger radii, the surface density of the disk is smaller, the vertical restoring force as well, which leads—if the vertical action is conserved—to larger vertical excursions of the orbits and therefore flaring of the stellar population at larger radii (i.e., the vertical scale height $h_z \longrightarrow h_z(R)$ rises with radius; Loebman et al. 2011). It has been thought that stars would stay forever on their birth orbits, an assumption that has turned out to be too simplistic. In the recent years, it has been recognized that radial migration could have profound effects on the overall structure of the disk (Sellwood & Binney 2002; see also Section 1.1.5 on the origin of exponential disks) and might be fundamental in explaining some of the observed phase-space and abundance features of the disk (Roškar et al. 2008; Schönrich & Binney 2009a; see also Sections 1.2.3 and 1.2.4). Many aspects are still not fully understood, for example the effect of radial migration on the vertical heating of the disk (cf., e.g., Loebman et al. 2011; Minchev et al. 2012).

1.3 Galactic Archaeology in the Milky Way

The understanding of our universe and the formation and evolution of galaxies has seen an unprecedented advance in the last decades. This is due to the technological progress in the construction of telescopes and high-performance computing. We live in a golden era of Galactic surveys, that culminates with the *Gaia* satellite mission, and will soon see a surge of new discoveries about the MW galaxy. This section is dedicated to the strategies, goals, and challenges of Galactic archaeology, the discipline of understanding the present-day MW in the context of galaxy formation. For reviews see Turon et al. (2008) and Rix & Bovy (2013). The following section was inspired by a talk by Wyse (2017).

1.3.1 Introduction to Galactic Archaeology

Studying galaxy formation and evolution. To learn about the origin of galaxies, astronomers employ three complementary approaches:

- *External galaxies.* The evolution of galaxies takes place on time scales between the lifetime of a massive star ($\sim 10^7$ yrs) and the age of the universe ($\sim$ 13.7 Gyr), so we will always

observe just one single 2D snapshot in the life of each galaxy. Individual stars in external galaxies cannot be resolved, so it is only possible to measure integrated quantities. However, due to the vast amount of galaxies of different types, evolution stages, and inclinations, and the finite speed of light, we can study galaxy formation and evolution in a statistical sense from a large number of different galaxies at different redshift epochs. (See, e.g., reviews by Madau & Dickinson (2014) and Conselice (2014), and references therein.)

- *Numerical simulations.* N-body and hydrodynamical computer simulations allow to follow the formation and evolution of a few galaxies in time lapse. Simulations are limited by our current knowledge of physical processes and empirical relations. On the one hand, we will never know how much the simulated galaxy evolution mirrors the real physics. On the other hand, we can directly intervene in the formation process and vary the initial and regulatory conditions to investigate their effects on the galaxy structure. (See, e.g., reviews by Somerville & Davé (2015) and Naab & Ostriker (2016), and references therein.)

- *The Milky Way.* As we will lay out in detail in the following paragraphs, the MW complements these approaches: It is only one single galaxy, but it is real, and we can study its evolution in detail on a star-by-star basis (Turon et al. 2008; Binney 2011).

What is Galactic archaeology? *Galactic archaeology* uses the stars in a galaxy as the fossil records to reconstruct the formation history of the galaxy (e.g., Eggen et al. 1962; Freeman & Bland-Hawthorn 2002). The related field of *near-field cosmology* is especially interested in the earliest phases of star and galaxy formation to give constraints on cosmological initial conditions (Frebel & Norris 2015). We distinguish between "galactic archaeology" as the study of galaxy formation in external galaxies, and "Galactic archaeology", which attempts to dissect the MW's history. We focus on the latter.
Stars carry signatures of their formation:

- *Chemical abundances.* The chemical abundances in the atmosphere of stars are set by the ISM at the time of their formation through a combination of (i) Big Bang nucleosynthesis of abundances in the primordial gas and (ii) stellar evolution and chemical enrichment history within the Galaxy. They are expected to be life-long tags (except, see Dotter et al. 2017).

- *Age.* The age can be derived, for example, via stellar evolution models from stellar parameters like the effective surface temperature $T_{\rm eff}$, surface gravity $\log g$, and chemical abundances, as measured from spectroscopy.

- *Orbits.* The initial orbits of stars are set by the random motions in the gas from which they were formed and the overall gravitational potential. Subsequent evolution, e.g. through gravitational torques, radial migration (see Section 1.4.5), heating mechanisms, etc., might change the orbit. Some orbit characteristics, e.g. the vertical action (Minchev et al. 2012; Solway et al. 2012; Vera-Ciro & D'Onghia 2016), are expected to be approximately conserved.

Galactic archaeology requires measurements of all these stellar properties for a large population of stars to be able to deduce the star formation history, the origin of the abundances, initial conditions, and secular evolution that shaped the orbit distribution. Figure 1.12 shows, for example, a recent measurement of the distribution of stellar ages within the Galaxy—a valuable clue for inside-out formation (see Section 1.1.4).
One sub-field of Galactic archaeology is *chemical tagging* (Freeman & Bland-Hawthorn 2002; Hogg et al. 2016; Ting et al. 2016), which assumes that stars which have the same chemical abundances have been born together from the same molecular cloud and with similar orbits. If

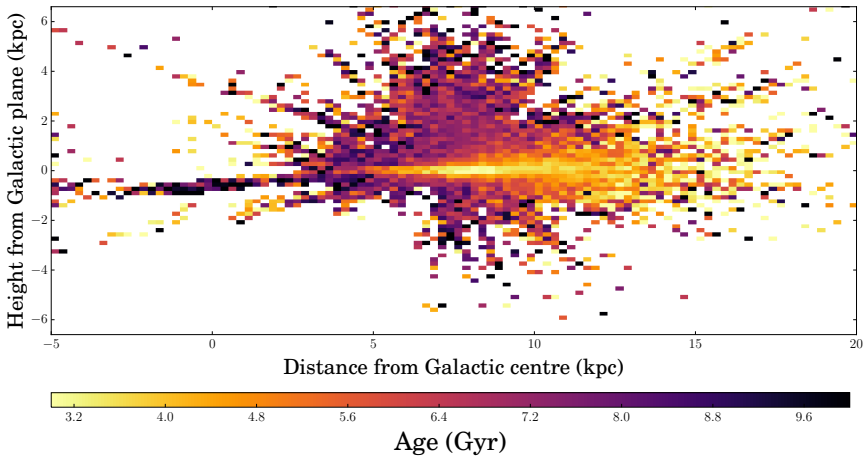


Figure 1.12: Distribution of APOGEE red giant stars in the Galactic meridional plane, color-coded according to their ages, derived with *The Cannon* (Ness et al. 2016). The youngest stars are in the plane of the disk and at large radii, while the central and high latitude regions are predominantly old. The radial distribution of ages is consistent with inside-out growth. Any theory of Galaxy assembly will have to be able to explain the overall age structure of the Galaxy. **Figure credit:** Melissa Ness (MPIA).

pairs and groups of such stars could be identified, and their orbits differ now, one might be able to learn about the orbit evolution that the stars experienced.

Specific goals. The first step in understanding the formation history of the MW is to have a good grip on its present structure, in particular the (i) gravitational potential and (ii) the stellar distribution.
The *Galactic gravitational potential* $\Phi(\boldsymbol{x})$ is fundamental for understanding the MW's DM and baryonic structure (Famaey 2012; Rix & Bovy 2013; Strigari 2013; Read 2014). It encompasses the total matter distribution and sets the force field in which stars and gas move. Or, in other words, it is the "invisible stage" of the Galaxy.
The *stellar Distribution Function (DF)*, $\mathrm{DF}(\boldsymbol{x}, \boldsymbol{v})$, is the "visible protagonist", as it captures what we observe. It is the basic constraint on the Galaxy's formation history (Binney 2013; Sanders & Binney 2015b). As mentioned in Section 1.1.3, galaxies should be considered as a superposition of stellar-population-depended DFs with different dynamical properties. Specifically, one is interested in the *orbit DF*, $\mathrm{DF}(\mathrm{orbit} \mid \Phi)$: the number of stars on a given orbit (see Section 1.4). This is physically most meaningful, because it directly sets (in combination with the potential) the overall appearance and evolutionary state of the Galaxy, while the observed positions and velocities of single stars are only one random realization of this orbit DF at a given point in time. If taken together with the ages t_{age} and chemical abundances $[X/\mathrm{H}]$ of stellar populations (with $X \in$ [Fe,Mg,O,C,...]), i.e., determining $\mathrm{DF}(\mathrm{orbit} \mid [\mathrm{X/H}], t_{\mathrm{age}})$, the DF is straight-forward to interpret in terms of "dynamical" evolution, enrichment history, etc. (Binney 2011).

1.3.2 The Milky Way as Ideal Site for Galactic Archaeology

Why study the MW? The MW is our home galaxy. The same physics that formed this special galaxy created also our Earth and us. Studying the MW's formation history means, therefore, studying our earliest origins.
Disk galaxies are an important regime in the diverse world of galaxies (see Section 1.1) and we believe the MW to be a typical disk galaxy, considering its size and properties (see, e.g., Mo et al. 2010; van der Kruit & Freeman 2011). It is close to the knee of the galaxy luminosity function, i.e., a L^* galaxy with stellar mass $M_* \sim 10^{11}$ $\mathrm{M}_\odot$, and close to the peak of the stellar-to-halo-mass

relation at a halo mass $M_{\rm vir,DM} \sim 10^{12}$ $M_\odot$ (see Section 1.1.5). The MW can, therefore, be considered a template for disk galaxy evolution. There is, for example, a surprising regularity in the MW disk considering relations between chemical abundances of stars and their positions, velocities and ages (see Section 1.2). This indicates that there are some mechanisms that shaped the MW disk that might be universal in all galaxies.
The MW is also special in the sense that it has avoided destructive mergers with neighboring galaxies so far. We should, thus, still be able to deduce many of the in-situ galaxy formation mechanisms.

The uniqueness of observations in the MW. The MW is the only galaxy that we can study up close and from within. It is the only galaxy in which we are able to resolve individual stars in order to measure their 3D positions, 3D motions and stellar parameters (chemical abundances, ages, masses, etc.) with high precision. The latter is of high importance for Galactic archaeology (see Section 1.3.1) and dynamical modeling (see Sections 1.4 and 1.5), where velocity errors of ~ 5 km s^{-1} in all three directions and abundance errors of ~ 0.1 dex and less are required (see also Chapters 2 and 4).
The open science questions on galaxy structure and evolution that we have posed in the previous sections 1.1.2, 1.1.5, and 1.2.5, can all be uniquely studied in the MW, because they profit from a star-by-star investigation.

1.3.3 Importance of Constraining the Galaxy's Potential and Disk DF

The holy grail in the fields of Galactic archaeology and Galactic dynamics is (i) to constrain the MW's gravitational potential Φ and today's stellar DF, and (ii) interpreting them in terms of galaxy formation and evolution. In the following we will give an overview of the different aspects and fields of research that benefit from having a reliable potential model and DF specifically for the MW.

Gravitational potential. The immediate result from measuring the MW's axisymmetric potential would be a better knowledge of its overall structural parameters, like the rotation curve, the disk scale lengths and heights, the shape of the halo etc. (Bland-Hawthorn & Gerhard 2016; see also Table 1.1), which are still less well known than in some external galaxies.
There is a lot to learn also beyond these model dependent numbers. The MW potential measurement could help to constrain galaxy formation and fundamental physics:

- A comparison of the measured potential with the observed distributions of stars and gas would reveal the DM—be it the halo or also a DM disk (see Section 1.2.5)—as the discrepancy. The overall spatial profile of the DM would help to understand the nature of DM (see Section 1.1.2).
- The relative amounts of gas, stars, and DM are also constraints for galaxy formation theory. For example, how the total stellar mass and star formation rate of a galaxy depends on the halo mass is an important question (Kauffmann et al. 1999; Moster et al. 2010; Behroozi et al. 2013; see also Section 1.1.5), and the MW could provide a reliable anchor point for this relation.
- A proper decomposition of the Galactic rotation curve into contributions of stellar disk and bulge, gas, and DM would help to resolve the question of the validity of the *maximum disk hypothesis* (see Section 1.1.5; Sackett 1997; Bovy & Rix 2013; Figure 1.13).
- A reliable astrophysical measurement for the DM density at the Sun is needed for DM detection experiments to estimate the expected interaction or annihilation rate (e.g., Su & Finkbeiner 2012; see also reviews by Feng 2010 and Strigari 2013).

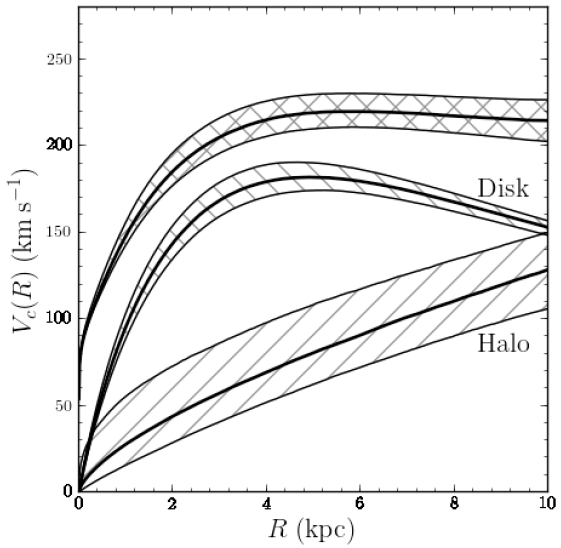


Figure 1.13: Decomposed rotation curve for the MW derived by Bovy & Rix (2013) from combining their radial profile for the vertical gravitational force derived with RoadMapping modeling from MAPs in the stellar disk, with terminal velocity measurements (Clemens 1985; McClure-Griffiths & Dickey 2007) and the stellar surface density from Zhang et al. (2013b). Understanding the relative contributions of stars and DM to the total gravitational potential of the MW is one of the principal goals of Galactic archaeology. **Figure credit:** Bovy & Rix (2013).

- A potential model could provide estimates for the mass contrast of the local spiral arm (see also Sections 3.5.4 and 5.1). This would help to understand the dynamical nature of spiral arms better, and to interpret local observations in stellar kinematics.

Having a reliable *axisymmetric potential model* for the MW is important in a wide range of studies:

- It is needed as the axisymmetric baseline for perturbation theory to better understand the dynamic effects of spiral arms and the Galactic bar (e.g., Binney 2013; Monari et al. 2016).
- It is needed to estimate reliable actions (see Section 1.4) and other orbital parameters for as many MW stars as possible (e.g., Bovy et al. 2012d; Anders et al. 2014; Myeong et al. 2017).
- It is needed in every study that requires orbit integration in the MW (e.g., de la Vega et al. 2015; Howes et al. 2015; Cantat-Gaudin et al. 2016; Hunt et al. 2016; O'Leary et al. 2016; Maji et al. 2017; Marchetti et al. 2017).
- It is needed for studies of accretion events and stellar streams in the Galactic halo (e.g., Fardal et al. 2015; Balbinot & Gieles 2017), as streams are created by the tidal disruption of globular clusters or dwarf galaxies which move in the MW potential.
- It could be needed for the comparison and/or even the setup of MW-like galaxy simulations, e.g., like those by D'Onghia et al. (2013) or Grand et al. (2017).

Stellar distribution function. The ultimate goal would be to have a set of DFs fully specifying the stellar content of the Galactic disk (and possibly also stellar halo and bulge) in terms of orbits, chemical abundances, masses, ages, relative numbers of stars in different sub-components, and possibly also including non-axisymmetric signatures of bar and spiral arms. So far, we are not even close to having a good handle on this.
Having *axisymmetric DFs* for some sub-populations, be it mono-abundance or mono-age populations (see Section 1.2.3) would be a useful first step:

- The DF could be used in chemical tagging (see Section 1.3.1) as smooth background model for the disk to find and characterize infall and substructures in orbit-abundance space.

Substructures that exist in orbit space independent of age and abundance were created by dynamical processes, like moving groups due to spiral arms and resonances. Substructures in orbit space that show up only in a small range of chemical abundances are expected to be connected via their formation within MW (see Section 1.2.4).

- The axisymmetric DF can be used as baseline for perturbation theory to investigate the dynamical effects of potential perturbations (e.g., Binney & Lacey 1988) like spiral arms and bars and to create orbit-based DFs for spiral arms (Monari et al. 2016) and the effects of resonances (Fouvry et al. 2015; Monari et al. 2017a).

- Galaxy simulations and chemical evolution studies that try to uncover the enrichment history of the MW (see, e.g., Minchev et al. 2017; Grand et al. 2017) could use an axisymmetric DF as reference for the disk configuration today.

- Together with a reliable potential model for the MW, the functional shape of the DF could be improved on the basis of new data (see, e.g., Section 1.5.2 for the DF used in this work). It could be, for example, extended to explicitly include chemical abundances (similar to Sanders & Binney 2015b).

- A DFs could be used to predict the information content of future Galactic surveys given a proposed survey Selection Function (SF).

Taken together, knowing the potential and DF of the MW could also help to learn more about the population of galaxies:

- The MW can be used to anchor extragalactic scaling relations. One example would be the Tully-Fisher relation, which is employed to estimate the distances to external disk galaxies (see Section 1.1.4; Klypin et al. 2002; Flynn et al. 2006; Hammer et al. 2007). This would require the total luminosity from the stellar DF and the circular velocity curve of the MW from the potential.

- The decomposition of bulge and disk components in external galaxies is often based on the fit of an exponential disk and (a Sersic) bulge to the surface brightness profile (e.g., de Jong 1996a; Graham 2001; MacArthur et al. 2003). It would be preferable if the decomposition would be performed on the basis of the dynamical properties of the components (e.g., Zhu et al. 2016). If we have found stellar DFs for the MW components, they will help to establish techniques also for the decomposition in external galaxies.

1.3.4 Galactic Surveys

To derive the MW's stellar DF and subsequently its gravitational potential, we first require measurements of 3D positions and velocities, and ideally also chemical abundances ([Fe/H], $[\alpha/\mathrm{Fe}]$, ...) and stellar labels ($\log g, T_{\mathrm{eff}}$) for a large ensemble of stars. The main reference for this section is the review by Rix & Bovy (2013).

Astrometric surveys. Imaging the plane of the sky allows to determine 2D positions, (R.A.,Dec.), for stars. If observations are repeated over a longer period of time, the change in the stars' positions allows the measurement of 2D proper motions, $\boldsymbol{\mu}$, and their distances via the trigonometric parallaxes, ϖ, from the Earth's motion around the Sun. The two largest astrometric surveys were performed by the High Precision Parallax Collecting Satellite (Hipparcos) (Perryman et al. 1997; ESA 1997) and *Gaia* (de Bruijne 2012), both space-based satellite missions, with the latter still on-going since 2013. Hipparcos recorded $\sim 10^5$ stars in a volume of ~ 200 pc around the Sun. The *Gaia* mission is expected to measure positions, proper motions, and parallaxes for

Table 1.2. Overview of the major spectroscopic MW surveys.

Survey name	Brightness limit	Number of stars	Velocity error	Resolution & signal-to-noise	Spectral range
Gaia-ESO (Gilmore et al. 2012)	$V < 19$ mag	$\sim 100,000$	~ 0.3 km s^{-1}	$R \sim 5,000 - 47,000$	NUV-NIR $3,000 - 11,000\AA$
Gaia's RVS[a] (Katz et al. 2004)	$G \lesssim 17$ mag	$(100 - 150) \times 10^6$	~ 2 km s^{-1}	$R = 11,500$ S/N > 10	NIR $8,450 - 8,720\AA$
GALAH (De Silva et al. 2015)	$V \approx 12 - 14$ mag	$< 1,000,000$ (9,860 in TGAS)	~ 0.2 km s^{-1}	$R \approx 28,000$ S/N > 100	visible-NIR $4,718 - 7,890\AA$
LAMOST/LEGUE (Cui et al. 2012)	$r < 19$ mag	$\sim 2,000,000$ ($\sim 100,000$ in TGAS)	~ 7 km s^{-1}[b]	$R \approx 1,800$ S/N > 10	visible-NIR $3,700 - 9,000\AA$
RAVE (Kunder et al. 2017)	$I \approx 9 - 12$ mag	457,588 (255,922 in TGAS)	~ 1.5 km s^{-1}	$R \approx 7,500$ S/N ~ 50	NIR $8,410 - 8,795$ Å
SDSS/APOGEE (Zasowski et al. 2013)	$H = 11 - 14$ mag	$\sim 100,000$ ($\sim 21,000$ in TGAS)	~ 0.1 km s^{-1}	$R \approx 22,500$ S/N > 100	infrared $15,100 - 17,000\AA$
SDSS/SEGUE (Yanny et al. 2009)	$g < 19$ mag	$\sim 350,000$	~ 4 km s^{-1}	$R \approx 2,000$ S/N ~ 25	visible-NIR $3,850 - 9,200\AA$

[a] First radial velocity measurements by *Gaia*'s Radial Velocity Spectrometer (RVS) will be available in 2018 (ESA 2017a).

[b] Schönrich & Aumer (2017).

$\sim 10^8$ stars out to ~ 10 kpc (Perryman 2005). In Section 4.2.1, we will have a closer look at the *Gaia* mission and the first Data Release (DR) from September 2016, in particular the Tycho-Gaia Astrometric Solution (TGAS).

Spectroscopic surveys. Splitting the light of an individual star into its full wavelength spectrum provides the third velocity component, the line-of-sight velocity, $v_{\rm los}$ of the star, and stellar labels and chemical abundances. In particular, the line-of-sight velocity manifests itself as the Doppler shift of the spectrum. Models for stellar atmospheres predict the strength of absorption lines in the spectrum depending on the stellar labels, and can be used to fit the observed spectrum. This allows the derivation of chemical abundances etc. to a precision of 0.05-0.2 dex (e.g., Asplund et al. 2009). Stellar ages and masses can be estimated from $\log g$ and $T_{\rm eff}$ together with chemical abundances and stellar evolution models (e.g., Soderblom 2010).
Table 1.2 summarizes several spectroscopic MW surveys. The different surveys supplement each other, as each has a different spectral range, resolution, and sample selection. In Section 4.2.2, we will present one of them in detail, the RAdial Velocity Experiment (RAVE) (Steinmetz et al. 2006).
Gaia will ultimately provide also radial velocities and metallicities. For other chemical abundances it is and will also in the future be necessary to combine *Gaia* measurements with other spectroscopic surveys.

Photometric surveys. Imaging of the sky allows the measurement of stellar fluxes in different wavelength passbands and therefore the determination of apparent brightness magnitudes and colors. Photometric colors can be used to estimate the metallicity of a star (e.g., Wallerstein 1962), but to lower precision (0.1-0.3 dex) than from spectra.
Some important photometric surveys are the Two Micron All Sky Survey (2MASS), an all-sky survey providing infrared magnitudes (Skrutskie et al. 2006), SDSS on the Northern sky, ranging from visible to NIR (Eisenstein et al. 2011), and Pan-STARRS, visible to NIR, which covers three quarters of the sky (Kaiser et al. 2002). They are all ground-based surveys. PanSTARRS and SDSS also provide some proper motions.

Distances. The distance to a star is the most difficult coordinate to measure. Trigonometric parallaxes have the advantage that they do not require spectroscopy and stellar models and work for any kind of bright star. The precision of trigonometric parallaxes, however, especially beyond a few hundred parsecs, are to date not yet precise enough for many purposes (see also Section 4.2.1). Alternatives are (spectro-)photometric distances. Some stellar types are standard candles in the sense that they have well-determined absolute luminosities, e.g., RC giants (see also Section 4.2.4), or variable stars for which the luminosity scales with the pulsation period. For non-standard candles, stellar evolution models can, on the basis of the spectroscopically determined stellar parameters, provide estimates for the expected absolute magnitude of a given star (e.g. Burnett & Binney 2010). The distance can then be calculated from the star's measured brightness and its known true brightness (*distance modulus*).

Survey Selection Function (SF). The principal problem of Galactic archaeology is that no survey can ever provide a full picture of the MW. This is mostly due to our specific location within the disk from which we can observe the Galaxy only edge-on. While we have an unobstructed view of the outer bulge and halo, the disk—which contains most of the stellar mass—is obscured by projection effects and interstellar dust: (i) crowding makes it difficult to disentangle the stellar content in regions of highest density; (ii) interstellar dust extinction absorbs star light which causes stars to appear fainter and more reddish. To correct for the latter is difficult (Bovy et al. 2016a) and requires at least a 2D map (Schlegel et al. 1998; Nidever et al. 2012)—or better 3D map (Green et al. 2015)—of dust in the Galaxy.
From a technical point of view, surveys are always brightness-limited: the fainter a star, the lower its photon signal-to-noise ratio at a given exposure time, which makes it more difficult to detect faint stars above the noise level of a CCD chip, or to distinguish physical absorption and extinction lines from noise in stellar spectra. In addition, surveys have often complex target selection strategies that can depend on position, color, etc. (see for example Section 4.2.2 for a summary of the target selection of RAVE). Any stellar catalog will therefore be incomplete. The volume that a survey covers within the Galaxy (the *effective survey volume*) is always spatially restricted to a few kiloparsecs or much less and depends strongly on the stellar population in question.
Characterizing the survey SF and its completeness to be able to correct for it, is a highly complex problem (e.g., Bovy et al. 2014; Wojno et al. 2017; Bovy 2017; see also Section 4.3 on the SF of *Gaia* and *RAVE*). It is, however, essential and necessary in Galactic archaeology to dissect signatures of the underlying physical stellar DF and selection effects (e.g., Rix & Bovy 2013).

1.4 Galactic Dynamics

The movements of stars in galaxies are rich in structure and complexity. In this section, we will give an overview of the principles of Galactic dynamics (Section 1.4.1), with a strong focus on orbits in the Galactic disk and an easy to understand introduction to orbital actions (Sections 1.4.2-1.4.4). We conclude with a short excursion on resonant orbits in Section 1.4.5, which is not subject of this thesis, but very important for the dynamics of the Galactic disk. The principal reference for Galactic dynamics is Binney & Tremaine (2008).

1.4.1 Galactic Dynamics in an Axisymmetric, Collisionless Framework

Galaxies are collisionless systems. Galaxies can be considered as collisionless stellar systems, as compared to, for example, star clusters. The following simple argument shows why. The change in a star's velocity v due to a gravitational encounter with another star of mass m at distance b is $\Delta v \sim 2Gm/(bv)$ (Binney & Tremaine 2008, §1.2.1). We assume for simplicity that

the cumulative effect of orbit-changing stellar encounters is dominated by a single significant encounter. The velocity change is significant enough to modify the orbit if $\Delta v \sim v$. The cross-section of a significant encounter is therefore $\sigma \sim \pi \left(2Gm/v^2\right)^2$, the corresponding mean free path $\lambda = (n\sigma)^{-1}$ with the stellar number density n, and the average time between significant encounters $t_{\rm enc} = \lambda/v$. For $m = 1$ M$_\odot$, $v \sim 20$ km s^{-1} (the Sun's approximate random motion with respect to the Galactic flow, see also Figure 4.2), and $n < 1$ pc^{-3} in the Galactic disk, $t_{\rm enc}$ is much longer than the age of the universe. Encounters between stars are therefore rare and the galaxy effectively collisionless. It is therefore the joint gravitational potential of all matter constituents in the galaxy that determines the motions and orbits of stars.

Axisymmetry & steady state. The principal assumptions of Galactic dynamics (in addition to the Galaxy being collisionless) are

(i) *The Galaxy is in a steady state.* Any function describing the overall state of the system—for example the gravitational potential Φ or the stellar DF (see next paragraph)—does not explicitly depend on and evolve with time, i.e. $\partial/\partial t = 0$.

(ii) *The Galaxy is axisymmetric.* Let us consider the MW in terms of a cylindrical coordinate system (R, ϕ, z), with the Galactic center at its origin, and the $(z = 0)$-plane being aligned with the Galactic disk. The Galactic potential can then, to first order, be approximated as being symmetric around the z-axis and the $(z = 0)$-plane. We will see in Section 1.4.2 that this allows three integrals of motions. The assumption of axisymmetry is related to the assumption that the Galactic disk is well-mixed in phase: It is assumed that all stars had sufficient time to orbit away from their birth positions after the formation of the Galaxy. Due to variations in their initial velocities, they are now uniformly distributed across all azimuthal angles (or phases) ϕ in the Galactocentric cylindrical coordinate frame. No function decribing the state of the system depends therefore explicitly on ϕ, i.e., $\partial/\partial\phi = 0$.

It is obvious from Section 1.2 that these two assumptions are clearly not fully satisfied in the MW. We will dedicate the whole Chapter 3 on the topic of assuming axisymmetry if the system in truth is not.

Distribution functions and Jeans theorem. A stellar system is an ensemble of stars, with the ith star being at time t located at position $\boldsymbol{x}_i$ and having velocity $\boldsymbol{v}_i$. As the MW disk has $\sim 10^{11}$ stars, it is in practice more feasible to describe the ensemble in terms of a probability function $\mathrm{DF}(\boldsymbol{x}, \boldsymbol{v}, t)$. The number of expected stars with $(\boldsymbol{x}_i, \boldsymbol{v}_i)$ is then $\mathrm{DF}(\boldsymbol{x}_i, \boldsymbol{v}_i, t)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{v}$. In a collisionless system, in which all stars move according to the equations of motion in a smooth external potential, the probability density has to evolve continuously in phase-space. This is equivalent (see §4.1 in Binney & Tremaine 2008) to the condition

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{DF}(\boldsymbol{x}, \boldsymbol{v}, t) \tag{1.4}$$

$$= \frac{\partial \mathrm{DF}}{\partial t} + \boldsymbol{v} \cdot \frac{\partial \mathrm{DF}}{\partial \boldsymbol{x}} - \frac{\partial \Phi}{\partial \boldsymbol{x}} \cdot \frac{\partial \mathrm{DF}}{\partial \boldsymbol{v}}, \tag{1.5}$$

which is called the *collisionless Boltzmann equation* (where we used the laws of motions $\dot{\boldsymbol{x}} = \boldsymbol{v}$ and $\dot{\boldsymbol{v}} = -\partial\Phi/\partial\boldsymbol{x}$ in the last step).
An integral of motion is any function $I(\boldsymbol{x}, \boldsymbol{v})$ that is conserved along a stellar orbit, i.e.,

$$0 = \frac{\mathrm{d}}{\mathrm{d}t} I = \boldsymbol{v} \cdot \frac{\partial I}{\partial \boldsymbol{x}} - \frac{\partial \Phi}{\partial \boldsymbol{x}} \cdot \frac{\partial I}{\partial \boldsymbol{v}}. \tag{1.6}$$

From comparison with the collisionless Boltzmann equation it becomes immediately evident that in a steady-state system, i.e., $\partial/\partial t = 0$, any function that is itself a function of integrals of

motion only,

$$\mathrm{DF}\left(\boldsymbol{I}\left[\boldsymbol{x}\left(t\right),\boldsymbol{v}\left(t\right)\right]\right), \tag{1.7}$$

solves the collisionless Boltzmann equation. This is known as the *Jeans theorem* (Jeans 1915).

Integrals of motion. It depends on the symmetries of the potential, how many integrals of motion it allows. In any static, time-independent potential, the Hamiltonian function, i.e., the total energy, is conserved. Spherically symmetric potentials are rotationally symmetric around all three axes, therefore the three components of the angular momentum are conserved. In spheroidal systems (oblate or prolate), there is symmetry around the z-axis and the $(z = 0)$-plane, which allows two integrals of motion in addition to the energy, one of them being the z-component of the angular momentum. In general triaxial potentials, rotational symmetry is not granted anymore and therefore the angular momentum is not conserved. Because of the mirror symmetry at two planes, triaxial potentials still allow in total (and maximal) three integrals of motion (see also the footnote on page 36). In the following sections we introduce a convenient coordinate system, the *action-angle variables*. The three actions are integrals of motion that have intuitive physical meaning and are therefore a most convenient choice for the arguments of a DF. In other words, any function of the actions will automatically describe a steady-state, collisionless system. So the only challenge is to pick a functional form for an action-based DF that best fits the system of consideration, in our case the MW disk (see also Section 1.5.2).

1.4.2 Heuristic Introduction to Actions

Near-circular orbits. The paths that stars follow in the smooth gravitational potential of the Galaxy are called *orbits*. Orbits can look quite complex, as demonstrated in Panels (a)-(c) of Figure 1.14. This figure displays four typical example orbits in the Galactic disk, integrated using the `galpy.orbit` `python` module (Bovy 2015) in the axisymmetric, time-independent MW-like potential `MWPotential2014` by Bovy (2015). We show a circular orbit (pink), a rosette orbit in the plane (green), a shell orbit (orange), and an orbit that will eventually fill an approximately cylindrical annulus in the Galactic disk (blue). Describing the time-evolution of a star along its orbit requires the full 6D $(\boldsymbol{x}(t), \boldsymbol{v}(t))$ path, which is very complicated and unpractical. Panel (c) in Figure 1.14 indicates, however, that all four example orbits are in fact quite similar: They all follow the circular orbit at the so-called *guiding-center radius* $R_g = 8$ kpc and $z = 0$ with velocity $v_{\mathrm{circ}}(R_g) = 220$ km s^{-1} around the GC, with only small oscillations around the mean position. The overall mean rotation motion of the four stars is therefore very easily described by (i) the azimuthal position ϕ within the Galaxy, which changes periodically from 0 to 2π with the *circular frequency*

$$\Omega \equiv \frac{v_{\mathrm{circ}}(R_g)}{R_g} = \sqrt{\frac{1}{R}\left(\frac{\partial\Phi}{\partial R}\right)}\bigg|_{R=R_g,z=0} \tag{1.8}$$

as the star moves along the orbit, and (ii) the constant guiding-center radius R_g. This is the first indication that there is at least one quantity that is conserved during the orbit—a so-called *integral of motion*—and that it is R_g, or, equivalently, the *z-component of the angular momentum*

$$L_z \equiv \frac{1}{2\pi}\int_0^{2\pi} v_T \cdot R \ \mathrm{d}\phi \equiv R_g \cdot v_{\mathrm{circ}}(R_g), \tag{1.9}$$

where R and v_T are the star's instantaneous radial position and tangential velocity. The integral over ϕ is basically taking the average of all $v_T \cdot R$ along the orbit, which is the definition of the guiding-center radius, $L_z = R_g \cdot v_{\mathrm{circ}}(R_g)$, in a given potential Φ. That the z-component of the angular momentum is conserved in an axisymmetric potential is easily understood, because there are no forces that act on the star in tangential direction.

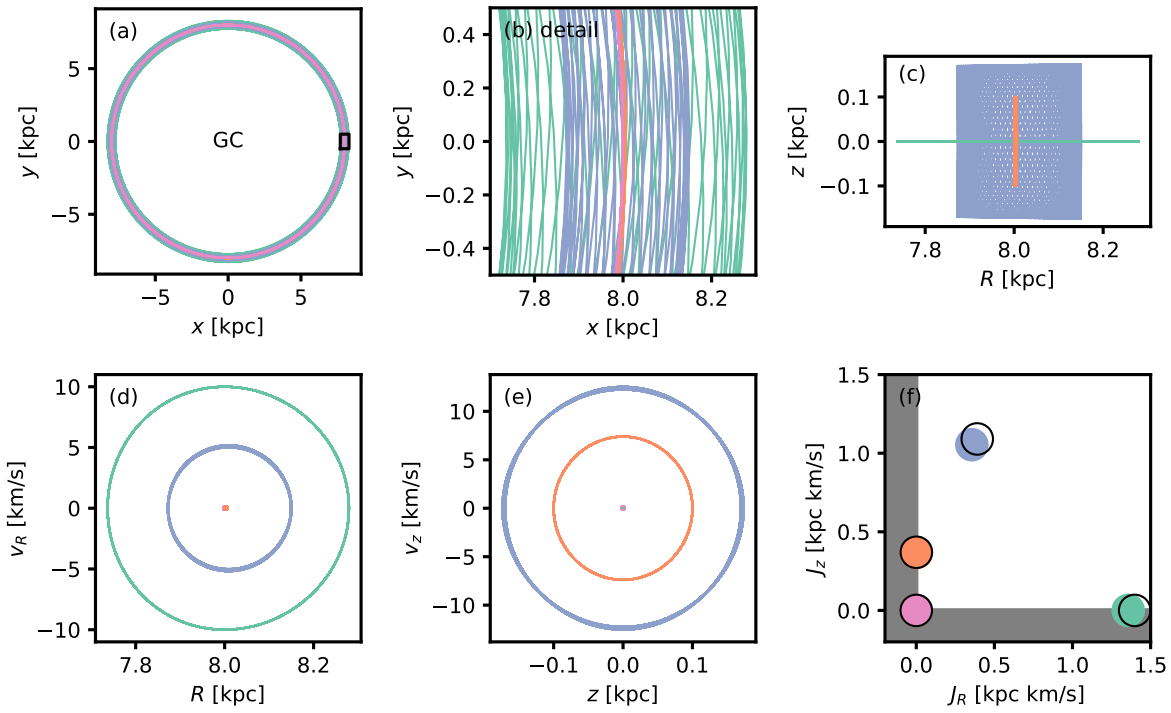


Figure 1.14: Four typical thin disk orbits in the MW and their actions. The orbits were integrated using `galpy` in the `MilkyWayPotential2014` by Bovy (2015). We show a circular orbit (pink), a shell orbit (orange), a planar rosette orbit (green), and a general orbit with excursions both in the plane and perpendicular to the plane at $z = 0$ (blue). Panel (a) shows the four orbits in Galactocentric Cartesian (x, y) coordinates around the GC at $(x, y) = (0, 0)$. Panel (b) shows a detail of the orbits around $(x, y) = (8, 0)$ kpc. Panel (c) shows the projection into Galactocentric cylindrical (R, z) coordinates (the meridional plane). Panel (d) shows the orbits in the radial velocity vs. position plane (R, v_R), and Panel (e) in the vertical velocity vs. position plane (z, v_z). The pink circular and the orange shell orbit are single dots at $(R, v_R) = (8 \text{ kpc}, 0)$ in Panel (d); the pink circular and the green planar orbit in Panel (e) are dots at $(z, v_z) = (0, 0)$. The last Panel (f), shows the radial and vertical action J_R and J_z (which are defined in the range $[0, \infty]$) for all four orbits calculated using the *Stäckel fudge* (colored dots; see Section (1.4.4) and estimated as the area inside the orbits in Panels (d) and (e) (black circles; see text for details).

Integrals of motion. Let's now have a closer look at the small oscillations that the stars perform around their guiding-center radius. In Panel (d) of Figure 1.14, we plot the instantaneous radial velocity, v_R, vs. the radial position, R, of the star, and analogous the vertical velocity, v_z, vs. the height, z, in Panel (e). Interestingly, the projection of the orbits into these two planes become closed curves. Instead of describing the motion in a two-dimensional plane, one would now only have to describe the motion along a one-dimensional line. That both the raduak and vertical component of the motion allow this reduction from 2D to 1D, suggest that an axisymmetric potential allows two more conserved quantities along an orbit in addition to L_z. A very popular choice for an integral of motion is of course the total energy

$$E = \frac{1}{2}\left(v_R^2 + v_T^2 + v_z^2\right) + \Phi(\boldsymbol{x}) \tag{1.10}$$

along the orbit. Choosing E and L_z as two of the three allowed integrals of motions in an axisymmetric potential makes the third integral of motion, I_3, rather cumbersome. There exists no analytic expression for I_3, and it also does not have a simple physical interpretation. Given E and L_z, I_3 quantifies how the fraction of the total energy that is not in the rotation is distributed between motions in the R- and z-direction.

Actions. By looking at Panels (d) and (e), we can come up with a more convenient choice for the additional two integrals of motions that are allowed by axisymmetric potentials. The closed curves look so much like circles or ellipses, that one is tempted to simplify the two-dimensional motion in the velocity vs. position plane to a one-dimensional motion described in polar coordinates with (i) a constant radius and (ii) a single angle-coordinate (θ_R and θ_z in Panel (d) and (e), respectively) that changes between 0 and 2π, analogously to the azimuthal motion around the GC. Defining a radius coordinate in a plane where x- and y-axis have different units is, of course, not sensible. So instead of a radius, we use the area inside the ellipse as the conserved quantity. The relation between the amplitude in position and the corresponding amplitude in the velocity is unambiguously set by the potential in which the star moves and the guiding-center radius. So the area inside the curves in Panels (d) and (e) are equivalent to a radius coordinate. We now introduce the *radial action* as the (scaled) area inside the orbit,

$$\begin{aligned} J_R &\equiv \frac{1}{2\pi}\int_{\text{area inside orbit}} \mathrm{d}R\ \mathrm{d}v_R && (1.11)\\ &= \frac{1}{2\pi}\oint_{\text{orbit}} v_R\ \mathrm{d}R && (1.12)\\ &= \frac{1}{\pi}\int_{R_{\min}}^{R_{\max}} v_R\ \mathrm{d}R && (1.13) \end{aligned}$$

where we used Stokes' theorem in the first step. The prefactor of $(2\pi)^{-1}$ is in analogy to Equation (1.9). $R_{\min}$ and $R_{\max}$ are the peri- and apocenter of the orbit. It is obvious that the radial action is an integral of motion with intuitive physical meaning: If it is zero, like for in the pink circular orbit or the orange shell orbit, the orbit does not oscillate at all in radial direction. The larger J_R, the larger the excursions of the orbit, both in radial position and velocity, from the circular orbit at the guiding-center radius. Analogously, we can introduce the *vertical action*

$$\begin{aligned} J_z &\equiv \frac{1}{2\pi}\int_{\text{area inside orbit}} \mathrm{d}z\ \mathrm{d}v_z && (1.14)\\ &= \frac{1}{2\pi}\oint_{\text{orbit}} v_z\ \mathrm{d}z && (1.15)\\ &= \frac{2}{\pi}\int_{0}^{z_{\max}} v_z\ \mathrm{d}z. && (1.16) \end{aligned}$$

Epicycle approximation. For the small oscillations in our example orbits, the motion in R and z can be considered as decoupled from the motion in ϕ and from each other. The *Hamiltonian function* (which is equal to the total energy in Equation (1.10) and determines the motion of a star in the potential Φ; see also Section 1.4.3) can be re-written as

$$\begin{aligned} \mathscr{H}(R,z,v_R,v_z) &= \frac{v_R^2}{2}+\frac{v_z^2}{2}+\Phi_{\text{eff}}(R,z) && (1.17)\\ \text{with } \Phi_{\text{eff}}(R,z) &\equiv \frac{L_z^2}{2R^2}+\Phi(R,z) && (1.18) \end{aligned}$$

with the effective potential Φ_{eff}, which is a function of the axisymmetric gravitational potential $\Phi(R,z)$ and the centrifugal potential $L_z/(2R^2)$, and uses $L_z = v_T \cdot R = \text{const}$ for a given orbit. A star with L_z that moves in this effective potential will effectively feel no forces if it is exactly at R_g and $z = 0$. If it is at $R > R_g$ or $z \neq 0$, it will feel the gravitational force pulling it towards the GC or back to the midplane. If it is at $R < R_g$, it will be pulled outwards by the centrifugal force. In other words, as long as the amplitude is small, it will oscillate like a simple harmonic

oscillator in both R- and z-direction,

$$\ddot{R} = -\kappa^2(R - R_g), \qquad \ddot{z} = -\nu^2 z \tag{1.19}$$

with the *radial* or *epicyclic frequency* κ and the *vertical frequency* ν,

$$\kappa^2(R_g) = \left.\frac{\partial^2 \Phi_{\text{eff}}}{\partial R^2}\right|_{R=R_g, z=0} \tag{1.20}$$

$$\nu^2(R_g) = \left.\frac{\partial^2 \Phi_{\text{eff}}}{\partial z^2}\right|_{R=R_g, z=0}. \tag{1.21}$$

Using the definition of Φ_{eff} in Equation (1.18) and the circular frequency Ω in Equation (1.8), this becomes

$$\kappa(R_g) = \left.\sqrt{R\frac{\partial \Omega^2}{\partial R} + 4\Omega^2}\right|_{R=R_g, z=0} \tag{1.22}$$

$$\nu(R_g) = \left.\frac{\partial^2 \Phi}{\partial z^2}\right|_{R=R_g, z=0}. \tag{1.23}$$

These two frequencies are the frequencies by which the angles θ_R and θ_z change along the closed curves in Panels (d) and (e). In case of the most general of our four example orbit classes, the blue one in Figure 1.14, κ and ν are not necessarily multiples of Ω, so when the star is again at the same azimuthal position ϕ after time $2\pi/\Omega$, it will not be again at the same (R, z) position, but at a slightly different one. This causes the orbit to look like a rosette that fills a cylindrical annulus given enough time. There are some special R_g for which the frequencies are multiples of each other, and the orbit will become a closed curve (see also Section 1.4.5).

Summary. To summarize, an axisymmetric potential allows three integrals of motion. In the case of near-circular orbits considered in this section (which is true for most orbits in the Galactic disk), a convenient choice are the actions $(J_R, J_\phi \equiv L_z, J_z)$. Given the potential, these three numbers fully specify a given orbit and their values tell us about the amount of oscillation in the three coordinate directions (R, ϕ, z) that the star performs along its orbit. There are also three angles and frequencies associated with the actions, the azimuth $\theta_\phi \equiv \phi$ and the circular frequency Ω, and the angles θ_R and θ_z in the v_R-vs.-R and v_z-vs.-z planes, in which the star oscillates like a harmonic oscillator with the frequencies κ and ν. It is obvious that these quantities make it much simpler to describe orbits, than using $(\boldsymbol{x}, \boldsymbol{v})$ or even the classical integrals (E, L_z) together with the notorious I_3 integral.

Need for more general definition. The last panel in Figure 1.14, Panel (f), shows the radial and vertical actions for the four orbits. The colored dots were calculated using the action estimation method *Stäckel fudge*, which we will introduce in Section 1.4.4. The black circles are the values for the actions which were calculated very simply by estimating the area of the ellipses in Panels (d) and (e) as $J_R \approx \pi \cdot \max[R(t) - R_g] \cdot \max[|v_R(t)|]/(2\pi)$ and $J_z \approx \pi \cdot \max[|z(t)|] \cdot \max[|v_z(t)|]/(2\pi)$. They agree very well, illustrating that the motivation of actions in this section is valid. There are, however, for the orbits with slightly larger excursions from the circular orbit some small discrepancies. This indicates that the above definition of actions is not yet general enough.

1.4.3 Formal Introduction to Actions

Any axisymmetric, time-independent potential allows three integrals of motion and the description of orbits in terms of action-angle coordinates. An orbit that possesses action-angle coordinates is called a *regular orbit* and it is quasi-periodic, i.e., its Fourier decomposition consists only of linear combinations of three fundamental frequencies.[3] The actions are integrals of motions, and their canonical conjugate generalized positions are three angles which are periodic in the fundamental frequencies. Together with their intuitive physical meaning, which we have motivated in the previous section, this makes actions excellent orbit labels. We will now formally introduce action-angle coordinates, broadly following Binney & Tremaine 2008, §3.5.

A convenient coordinate system for orbits. In the previous section we have only considered special orbits, for which the *epicycle approximation* was valid. In this approximation, the motions in all coordinate directions can be considered as de-coupled from each other. This is true if the amplitude in R and z is small enough that $\kappa(R) \approx \kappa(R_g)$ and $\nu(R) \approx \nu(R_g)$ along the orbit. If the motions start to be coupled to each other, the lines in Panels (d) and (e) in Figure 1.14 start to smear out. For the blue orbit this effect can be already seen.
It is, however, still true that axisymmetric potentials allow three integrals of motion. So it is also generally true, that there has to exist a coordinate system for regular orbits consisting of three generalized "positions" θ_i and generalized "momenta" J_i with the condition that

$$\text{Condition (1):} \qquad \dot{J}_i \stackrel{!}{=} 0. \tag{1.24}$$

In addition, as we are interested in bound orbits that revolve around the Galaxy periodically, we would like the θ_i to be periodic in $[0, 2\pi]$ by construction and evolve linearly in time

$$\text{Condition (2):} \qquad \theta_i(t) \stackrel{!}{=} \theta_{i,0} + \Omega_i \cdot t \tag{1.25}$$

according to some constant frequency Ω_i. We are looking therefore for a coordinate system that satisfies these conditions for any orbit and reduces to the very simple to understand action-angle coordinates that we have considered in the previous section in the limiting case of near-circular orbits.

Hamilton equations and canonical coordinates. The motion of a particle in a potential is described in *Hamiltonian mechanics* by the *Hamiltonian function*, $\mathscr{H}$, that is a function of positions $\boldsymbol{q}$ and their canonical conjugate momenta $\boldsymbol{p}$ and which is equal to the sum of kinetic energy K and potential energy V,

$$\mathscr{H}(\boldsymbol{q}, \boldsymbol{p}) = K + V. \tag{1.26}$$

The relation between canonical positions, velocities $\dot{\boldsymbol{q}}$, and their conjugate momenta is defined by the *Lagrangian function*, $\mathscr{L}$, via

$$\boldsymbol{p} = \frac{\partial \mathscr{L}}{\partial \dot{\boldsymbol{q}}} \quad \text{with} \quad \mathscr{L} = \sum_i \dot{q}_i p_i - \mathscr{H}. \tag{1.27}$$

[3] All orbits in Stäckel potentials (which we will introduce in Section 1.4.4) and most orbits in typical axisymmetric and triaxial galaxy potentials are regular orbits that have three integrals of motions. The more a potential deviates from a Stäckel potential, however, the larger the region in phase-space within which orbits are irregular and stochastic. Stochastic orbits have a chaotic appearance and less than three integrals of motions—sometimes only the energy is conserved. In this work we focus on regular orbits only.

The time evolution of $\boldsymbol{q}$ and $\boldsymbol{p}$ is given by the *Hamilton equations*

$$(1) \quad \dot{\boldsymbol{p}} = -\frac{\partial \mathscr{H}}{\partial \boldsymbol{q}}, \qquad (2) \quad \dot{\boldsymbol{q}} = +\frac{\partial \mathscr{H}}{\partial \boldsymbol{p}}. \tag{1.28}$$

Classically, $\boldsymbol{q} = \boldsymbol{x}$ and $\boldsymbol{p} = \boldsymbol{v}$ in the usual position-velocity phase-space. The Hamiltonian is then $\mathscr{H} = \boldsymbol{v}^2/2 + \Phi(\boldsymbol{x})$, and the Hamilton equations are simply $\dot{\boldsymbol{x}} = \boldsymbol{v}$ and $\dot{\boldsymbol{v}} = -\partial\Phi/\partial\boldsymbol{x} \equiv \boldsymbol{F}(\boldsymbol{x})$, the (gravitational) force acting on the particle with mass $m = 1$.

It would be convenient if our desired unknown coordinates $(\boldsymbol{\theta}, \boldsymbol{J})$ were also canonical conjugate coordinates, i.e.

$$\text{Condition (3):} \qquad \mathrm{d}\boldsymbol{q}\,\mathrm{d}\boldsymbol{p} = \mathrm{d}\boldsymbol{\theta}\,\mathrm{d}\boldsymbol{J}, \tag{1.29}$$

or, phrased differently, the Jacobian determinant $|\partial(\boldsymbol{J}, \boldsymbol{\theta})/\partial(\boldsymbol{q}, \boldsymbol{p})| = 1$. The big advantage would be that the phase-space density would be conserved under the transformation $(\boldsymbol{x}, \boldsymbol{v}) \longrightarrow (\boldsymbol{\theta}, \boldsymbol{J})$, i.e., the number of stars in a little 6D box $\mathrm{d}^3x\,\mathrm{d}^3v$ around $(\boldsymbol{x}, \boldsymbol{v})$ would be the same as in the box $\mathrm{d}^3\theta\,\mathrm{d}^3J$ around the corresponding $(\boldsymbol{\theta}, \boldsymbol{J})$. If Condition (3) is satisfied, then $(\boldsymbol{\theta}, \boldsymbol{J})$ can also be used as arguments for the Hamilton equations, and Conditions (1) and (2) become

$$(1) \quad \dot{\boldsymbol{J}} = -\frac{\partial \mathscr{H}}{\partial \boldsymbol{\theta}} = 0 \tag{1.30}$$

$$(2) \quad \dot{\boldsymbol{\theta}} = +\frac{\partial \mathscr{H}}{\partial \boldsymbol{J}} = \boldsymbol{\Omega}(\boldsymbol{J}). \tag{1.31}$$

Poincaré invariant. We will now show that the so-called *Poincaré invariant* is the integral of motion that satisfies the conditions (1) and (2). For that, we introduce a general closed curve γ_i in the 6D phase space $(\boldsymbol{q}, \boldsymbol{p})$. This curve can be parametrized by an angle θ_i which is supposed to evolve in time according to our requirement in Equation (1.25). The point $(\boldsymbol{q}(\theta_i), \boldsymbol{p}(\theta_i))$ on γ_i at $\theta_i = 0$ is the same as the point at $\theta_i = 2\pi$. Next, we introduce the quantity

$$A_i \equiv \oint_{\gamma_i} \boldsymbol{p}\,\mathrm{d}\boldsymbol{q} = \int_0^{2\pi} \boldsymbol{p}\frac{\partial \boldsymbol{q}}{\partial \theta_i}\,\mathrm{d}\theta_i, \tag{1.32}$$

where we have used the parametrization via θ_i of the line integral in the second step. How does this quantity change with time? The total time-derivative (with $\theta \equiv \theta_i$ for readability), following Fitzpatrick (2014) §2.7, is

$$\frac{\mathrm{d}A_i}{\mathrm{d}t} = \int_0^{2\pi} \frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{p}\frac{\partial \boldsymbol{q}}{\partial \theta}\right)\,\mathrm{d}\theta \tag{1.33}$$

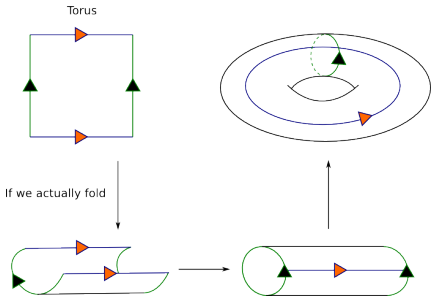
$$= \int_0^{2\pi} \left(\frac{\partial \boldsymbol{p}}{\partial t} + \frac{\partial \boldsymbol{p}}{\partial \theta}\dot{\theta}\right)\frac{\partial \boldsymbol{q}}{\partial \theta} + \boldsymbol{p}\left(\frac{\partial^2 \boldsymbol{q}}{\partial t \partial \theta} + \frac{\partial^2 \boldsymbol{q}}{\partial \theta^2}\dot{\theta}\right)\,\mathrm{d}\theta \tag{1.34}$$

$$\overset{\text{Eq. (1.25)}}{=} \int_0^{2\pi} \left(\frac{\partial \boldsymbol{p}}{\partial t}\frac{\partial \boldsymbol{q}}{\partial \theta} + \frac{\partial \boldsymbol{q}}{\partial \theta}\frac{\partial \boldsymbol{p}}{\partial \theta}\Omega\right)\mathrm{d}\theta + \int_0^{2\pi} \boldsymbol{p}\frac{\partial}{\partial \theta}\left(\frac{\partial \boldsymbol{q}}{\partial t} + \frac{\partial \boldsymbol{q}}{\partial \theta}\Omega\right)\,\mathrm{d}\theta \tag{1.35}$$

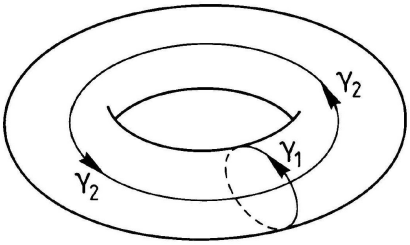
$$\overset{\text{int. by parts}}{=} \int_0^{2\pi} \left(\frac{\partial \boldsymbol{p}}{\partial t}\frac{\partial \boldsymbol{q}}{\partial \theta} + \frac{\partial \boldsymbol{q}}{\partial \theta}\frac{\partial \boldsymbol{p}}{\partial \theta}\Omega\right)\mathrm{d}\theta + \underbrace{\left[\boldsymbol{p}\left(\frac{\partial \boldsymbol{q}}{\partial t} + \frac{\partial \boldsymbol{q}}{\partial \theta}\Omega\right)\right]_0^{2\pi}}_{=0} - \int_0^{2\pi} \frac{\partial \boldsymbol{p}}{\partial \theta}\left(\frac{\partial \boldsymbol{q}}{\partial t} + \frac{\partial \boldsymbol{q}}{\partial \theta}\Omega\right)\,\mathrm{d}\theta \tag{1.36}$$

$$= \int_0^{2\pi} \frac{\partial \boldsymbol{p}}{\partial t}\frac{\partial \boldsymbol{q}}{\partial \theta} + \frac{\partial \boldsymbol{q}}{\partial \theta}\frac{\partial \boldsymbol{p}}{\partial \theta}\Omega - \frac{\partial \boldsymbol{p}}{\partial \theta}\frac{\partial \boldsymbol{q}}{\partial t} - \frac{\partial \boldsymbol{p}}{\partial \theta}\frac{\partial \boldsymbol{q}}{\partial \theta}\Omega\,\mathrm{d}\theta \tag{1.37}$$

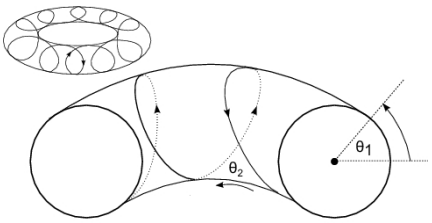
$$= \int_0^{2\pi} \frac{\partial \boldsymbol{p}}{\partial t}\frac{\partial \boldsymbol{q}}{\partial \theta} - \frac{\partial \boldsymbol{q}}{\partial t}\frac{\partial \boldsymbol{p}}{\partial \theta}\,\mathrm{d}\theta \tag{1.38}$$



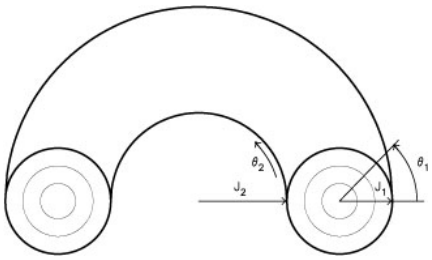
(a) Folding a 2-surface to form a periodic 2-torus.



(b) Two closed independent curves on the surface of a 2-torus.



(c) A star orbits periodically on the torus surface in the angles θ_i.



(d) The actions J_i determine the size of the torus.

Figure 1.15: Illustration of the torus description for stellar orbits in axisymmetric galaxy potentials (in 2 instead of 3 dimensions for presentation purposes). The three integrals of motion reduce the 6D orbit motion in $(\boldsymbol{x}, \boldsymbol{v})$ to a motion on a 3D-surface. The motion is quasi-periodic on this plane, so this plane is topologically the surface of a 3-torus (Panel 1.15(a), here a 2-torus). It is possible to define on this surface three closed curves that cannot be transformed into each other (Panel 1.15(b), here two curves). Each of this curves allows the introduction of a periodic angle coordiante θ_i. The stars oscillate on this surface. The motion can therefore be decomposed into the periodic motions in these angles θ_i (Panel 1.15(c), here two angles). The size of the orbital torus is set by the actions J_i, which are related to the torus radii via $r_i = \sqrt{J_i}$ (Panel 1.15(d)). **Figure credit:** Panel(a): Math Explorer's Club (2009); Panel (b): Binney & Tremaine (2008); Panel (c): Brink et al. (2015); Panel (d): Merritt (1999).

Using the Hamilton equations in Equation (1.28), this becomes

$$\begin{aligned}
\frac{\mathrm{d}A_i}{\mathrm{d}t} &\overset{\text{Eq. (1.28)}}{=} \int_0^{2\pi} -\frac{\partial \mathscr{H}}{\partial \boldsymbol{q}}\frac{\partial \boldsymbol{q}}{\partial \theta} - \frac{\partial \mathscr{H}}{\partial \boldsymbol{p}}\frac{\partial \boldsymbol{p}}{\partial \theta}\ \mathrm{d}\theta && (1.39)\\
&= -\int_0^{2\pi} \frac{\mathrm{d}\mathscr{H}}{\mathrm{d}\theta}\ \mathrm{d}\theta && (1.40)\\
&= 0. && (1.41)
\end{aligned}$$

In the last two steps we have used that the Hamiltonian $\mathscr{H}(\boldsymbol{q}, \boldsymbol{p})$ takes the same value at $\theta_i = 0$ and $\theta_i = 2\pi$ and does not explicitly depend on $\boldsymbol{\theta}$ (Equation 1.30). A_i is therefore a conserved quantity under time evolution. It is called the *Poincaré invariant*. Using Stokes' theorem and the fact that $(\boldsymbol{\theta}, \boldsymbol{J})$ are canonical conjugate coordinates (Equation 1.29), we get

$$\begin{aligned}
A_i &= \oint_{\gamma_i} \boldsymbol{p}\ \mathrm{d}\boldsymbol{q} = \int_{\text{area inside } \gamma_i} \mathrm{d}\boldsymbol{q}\,\mathrm{d}\boldsymbol{p} = \int_{\text{area inside } \gamma_i} \mathrm{d}\boldsymbol{\theta}\,\mathrm{d}\boldsymbol{J} && (1.42)\\
&= \int_{\text{area inside } \gamma_i} \mathrm{d}\theta_i\,\mathrm{d}J_i = \oint_{\gamma_i} J_i\ \mathrm{d}\theta_i = J_i \int_0^{2\pi} \mathrm{d}\theta_i = 2\pi J_i && (1.43)\\
\Rightarrow J_i &= \frac{1}{2\pi}\oint_{\gamma_i} \boldsymbol{p}\ \mathrm{d}\boldsymbol{q}, && (1.44)
\end{aligned}$$

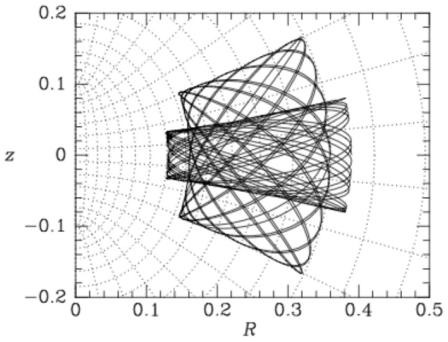


Figure 1.16: *Dotted lines:* Lines of constant u (ellipses) and v (hyperbola) of the spheroidal coordinate system defined in Equation (1.46) shown in the meridional plane. *Solid lines:* Two numerically integrated orbits in a realistic (logarithmic) galaxy potential. The orbits are bounded by lines of constant u and v. **Figure credit:** Binney & Tremaine (2008), their figure 3.27.

as general definition for the action J_i, analogous to Equations (1.12) and (1.15).

Torus description of an orbit. Let us now consider the orbit of a star in 6D $(\boldsymbol{q}, \boldsymbol{p})$. The three integrals of motion permitted by the axisymmetric potential reduce the orbit to a closed 3D surface in this 6D space. This surface has the same topology as the surface of a 3-torus.
To better illustrate this, we show in Figure 1.15(a) a 2-torus, which looks like a donut. While moving along its orbit, a star would spiral on the surface around the torus (see Figure 1.15(c)). It is possible to define two closed curves γ_i on this surface that define 2D-areas that are perpendicular to each other (see Figure 1.15(b)). A 3-torus would correspondingly allow three γ_i curves. Each γ_i allows the definition of a (θ_i, J_i) pair according to the previous section. The spiralling motion of the star on the torus surface can therefore be decomposed into three decoupled linear motions in θ_1, θ_2, and θ_3. The actions define the size of the torus (see Figure 1.15(d)).

1.4.4 Action Estimation

Need for approximations. The art of calculating actions J_i is to find the three 2D planes in phase-space in which the periodic motions decouple and on which γ_i of the orbit is a closed curve. In an axisymmetric potential, the actions can also be defined to be $\boldsymbol{J} \equiv (J_R, J_\phi = L_z, J_z)$ and are calculated as

$$J_i \equiv \frac{1}{2\pi} \oint_{\text{orbit}} p_i \, \mathrm{d}q_i \quad \text{with } i \in [R, \phi, z]. \tag{1.45}$$

In this case, because the motions do not fully decouple in R and z (see Figure 1.14, blue curve), γ_i does not close after one period, and the line integral needs to be evaluated over the full orbit from $q_i(t = 0)$ to $q_i(t = \infty)$. To integrate the full orbit up to infinity is computationally very expensive and, in practice, not a feasible approach. This demonstrates why actions have originally been used only in Solar system dynamics (Delaunay 1860) and in Galactic dynamics only recently: In spherically symmetric potentials, like the Kepler potential, four integrals of motions are allowed, $(E, \boldsymbol{L})$ and the motions in the coordinate directions (R, θ, ϕ) fully decouple, making the action calculation $(J_R, J_z \sim J_\theta, J_\phi = L_z)$ straight-forward. For general axisymmetric galaxy potentials and especially if actions for a large number of stellar orbits are required, the accurate calculations of actions can only be performed using supercomputers.

Prolate confocal coordinates. Figure 1.16 shows two example orbits in an axisymmetric galaxy potential (solid lines), which are more general than the near-circular orbits in Figure 1.14. We introduce now a coordinate system, the *prolate confocal coordinates* (u, v, ϕ), which are related to the normal Galactocentric cylindrical coordinates (R, ϕ, z) by

$$R = \Delta \cdot \sinh u \cdot \sin v, \qquad z = \Delta \cdot \cosh u \cdot \cos v, \qquad \phi = \phi. \tag{1.46}$$

Δ is called the focal distance and defines this coordinate system. In Figure 1.16, lines of constant u and v are overplotted (dotted lines), which take the form of ellipses or hyperbolas, respectively. It is now evident that these more general orbits are not expected to decouple in (R, z), but might in the (u, v) coordinates, because the orbits are bounded by lines of constant u and v. For near-circular orbits in the plane, u is close to a radial coordinate, and v a measure for the height above the plane. The canonical conjugate momenta of u and v, $p_u(u, v, \dot{u})$ and $p_v(u, v, \dot{v})$, can be calculated using Equation (1.27) and the Lagrangian $\mathscr{L}\left(R(u, v), z(u, v)\right)$.

Stäckel potentials. For a special kind of axisymmetric potential, the expressions for p_u and p_v can be simplified and expressed in terms of integrals of motions as $p_u(u, E, L_z, I_3)$ and $p_v(v, E, L_z, I_3)$. Then the motions in u and v are decoupled and the calculation of actions becomes straight-forward. In order to do this, we use the Hamiltonian which is, for an axisymmetric potential in cylindrical coordinates (R, z),

$$\mathscr{H}(R, z, p_R, p_z) = \frac{1}{2}\left(p_R^2 + p_z^2 + \frac{L_z^2}{R^2}\right) + \Phi(R, z), \tag{1.47}$$

and in (u, v) coordinates

$$\mathscr{H}(u, v, p_u, p_v) = \frac{p_u^2 + p_v^2}{2\Delta^2(\sinh^2 u + \sin^2 v)} + \frac{L_z^2}{2\Delta^2 \sinh^2 u \sin^2 v} + \Phi(u, v). \tag{1.48}$$

Setting $\mathscr{H} = E$, we get the Hamilton-Jacobi equation

$$2\Delta^2 \cdot E \cdot (\sinh^2 u + \sin^2 v) = p_u^2 + p_v^2 + \frac{L_z^2}{\sin^2 v} + \frac{L_z^2}{\sinh^2 u} + 2\Delta^2 \cdot (\sinh^2 u + \sin^2 v) \cdot \Phi(u, v).$$

This equation can only be solved with the *separation of variables* ansatz to decouple the motions in u and v, if the potential is *separable*, i.e., of the form

$$(\sinh^2 u + \sin^2 v) \cdot \Phi(u, v) \equiv U(u) - V(v) \tag{1.49}$$

$$\Rightarrow \qquad \Phi(u, v) = \frac{U(u) - V(v)}{(\sinh^2 u + \sin^2 v)}. \tag{1.50}$$

Axisymmetric potentials of this form are called *Stäckel potentials* and they resemble real galaxies (Batsleer & Dejonghe 1994; Famaey & Dejonghe 2003; Binney & Tremaine 2008, §3.5.3). We separate the variables,

$$2\Delta^2\left(E\sinh^2 u - U(u)\right) - p_u^2 - \frac{L_z^2}{\sinh^2 u} = \frac{L_z^2}{\sin^2 v} + p_v^2 - 2\Delta^2\left(E\sin^2 v + V(v)\right) \equiv 2\Delta^2 I_3$$

and set both sides equal to a constant $2\Delta^2 I_3$ that is independent of u and v. I_3 is an integral of motion. Solving for p_u and p_v and using Equation (1.45), the actions are

$$J_R \approx J_u = \frac{1}{\pi}\int_{u_{\min}}^{u_{\max}} p_u(u)\ \mathrm{d}u \tag{1.51}$$

$$\text{with} \quad p_u^2(u) = 2\Delta^2\left(E\sinh^2 u - I_3 - U(u)\right) - \frac{L_z}{\sinh^2 u} \tag{1.52}$$

$$J_z \approx J_v = \frac{1}{\pi}\int_{v_{\min}}^{v_{\max}} p_v(v)\ \mathrm{d}v \tag{1.53}$$

$$\text{with} \quad p_v^2(v) = 2\Delta^2\left(E\sin^2 v + I_3 + V(v)\right) - \frac{L_z}{\sin^2 v}, \tag{1.54}$$

with $u_{\mathrm{min/max}}$ and $v_{\mathrm{min/max}}$ being the roots of $p_u(u)$ and $p_v(v)$, i.e., the boundary curves in Figure 1.16. (E, L_z, I_3) can be calculated from a single given point $(R, z, \boldsymbol{v})$ on the orbit in the known potential. To calculate the actions in Stäckel potentials, there is therefore no need to integrate the orbit in time up to infinity, a single integration of a simple function suffices.

Stäckel fudge. The *Stäckel fudge* is an action estimation method for general axisymmetric potentials $\Phi(u, v)$ proposed by Binney (2012a). Here, we summarize the algorithm. At its core, the method simply pretends the potential was a Stäckel potential and calculates the actions following Equations (1.51)-(1.54). This requires two preparation steps: (i) setup of two functions $\delta U(u)$ and $\delta V(v)$, and (ii) pre-calculating the integrals of motion $[I_3 + U(u_0)]$ and $\left[I_3 + V\left(\frac{\pi}{2}\right)\right]$. Step (i) uses Equation (1.49) to setup

$$\delta U(u) \equiv U(u) - U(u_0), \qquad \delta V(v) \equiv V(v) - V\left(\frac{\pi}{2}\right) \tag{1.55}$$

using the potential $\Phi(u, v)$. If the potential is close to a Stäckel potential, $\delta U(u)$ will only weakly depend on v, and analogously for $\delta V(v)$. Step (ii) calculates from the currently observed phase-space position (u', v', p'_u, p'_v) the following quantities: E and L_z, $U(u') = \delta U(u') + U(u_0)$ and $V(v') = \delta V(v') + V(\pi/2)$. All of this is then inserted into Equations (1.52) and (1.54) and the resulting equations $p_u^2(u') = p'_u$ and $p_v^2(v') = p'_v$ are solved for $[I_3 + U(u_0)]$ and $\left[I_3 + V\left(\frac{\pi}{2}\right)\right]$. Step (iii) replaces in Equations (1.52) and (1.54)

$$I_3 + U(u) \to [I_3 + U(u_0)] + \delta U(u), \qquad I_3 + V(v) \to \left[I_3 + V\left(\frac{\pi}{2}\right)\right] + \delta V(v) \tag{1.56}$$

and the functions $p_u(u)$ and $p_v(v)$ can then be integrated to calculate the actions. The *Stäckel fudge* and the choice of Δ and u_0 for a given potential is also discussed in Binney (2012a), Bovy & Rix (2013), and Sanders & Binney (2016). In general, for a MW-like potential in the Solar neighborhood, $\Delta = 0.45 R_\odot$ is a good choice. A `python` implementation of the *Stäckel fudge* can be found in the `galpy` package by Bovy (2015).

Overview of action estimation in different potentials. Action calculation from a star's phase-space coordinates, $(\boldsymbol{x}, \boldsymbol{v}) \xrightarrow{\Phi} \boldsymbol{J}$, is typically very computationally expensive. Actions can, for general potentials, only be estimated. The triaxial Stäckel potentials (de Zeeuw 1985) are the most general potentials that allow exact action calculations using a single quadrature; earlier in this section we have introduced the special case of axisymmetric Stäckel potentials. The spherical isochrone potential (Henon 1959; Binney & Tremaine 2008, §3.5.3) is the most general potential for which the action calculation is analytic without any integration. The Kepler potential is a limiting case of the isochrone potential. In all other potentials, actions have to be numerically estimated; see Sanders & Binney (2016) for a recent review of action estimation methods, and also Bovy (2014) for a general method to compute $(\boldsymbol{x}, \boldsymbol{v}) \longrightarrow (\boldsymbol{J}, \boldsymbol{\theta})$.
Sanders & Binney (2016) sort action estimation methods for general galaxy potentials into convergent and non-convergent approaches. The accuracy of the latter is set by how close the potential in question is to a separable potential; the accuracy can therefore not be improved by investing more computation time. The *Stäckel fudge* belongs to these non-convergent methods. There are two other (non-convergent) ways to employ the convenient properties of the Stäckel potential to estimate actions:

- *Stäckel expansion.* If a simple functional form for $U(u)$ and $V(v)$ in Equation (1.50) is assumed—for example the Kuzmin-Kutuzov potential—, this can be used to expand the true potential as a sum of many of these Stäckel potentials with different parameters. Batsleer & Dejonghe (1994) used, for example, two components to reproduce the MW

potential (see also Section 2.2.3), and Famaey & Dejonghe (2003) used three components. Any sum of Stäckel potentials is itself a Stäckel potential as long as all components have the same focal length Δ. While the action estimation is computationally fast, this has two drawbacks: (i) We are restricted to potentials that have a fixed Δ, which makes it hard to reproduce very flattened disk-like and round halo-like components simultaneously (see also Section 2.3.6). (ii) There might be Stäckel potentials that fit locally—i.e., in the region of the stellar orbit—better than an overall best fit expansion, and would therefore allow much better accuracy. Because of the latter, the *Stäckel fudge* is in general preferred over this *Stäckel expansion* approach.

- *Stäckel fitting.* Computationally much slower but more accurate is the *Stäckel fitting* procedure (Dejonghe & de Zeeuw 1988; Sanders 2012; Sanders & Binney 2016), that—for each single star—first integrates part of the orbit, finds an optimal Δ, and then numerically the locally best $U(u)$ and $V(v)$. According to Sanders & Binney (2016), *Stäckel fitting* takes ~ 0.005 sec for one action calculation as compared to only $\sim 10^{-5}$ sec with the *Stäckel fudge*.

Among the convergent action estimation methods presented by Sanders & Binney (2016) are the *iterative torus construction* (McMillan & Binney 2008; Sanders & Binney 2015a) and the *generating function from orbit integration* (Sanders & Binney 2014) method. Both methods work by numerically finding the coordinate transformation—i.e., the generating function—between the action-angles in a toy isochrone potential (for which the action calculation is analytic) and the true potential. These methods can become highly accurate, but are very slow. They can take between 10^{-3} sec and 0.2 sec depending on orbit and method (Sanders & Binney 2016). The best compromise between accuracy and speed for orbits in the thin disk is therefore the *Stäckel fudge*, which we will use throughout this work.

1.4.5 Excursion: Resonances in the Disk

The previous sections considered Galactic dynamics in the axisymmetric limit. In reality, the MW is not axisymmetric. Especially radial migration triggered at the resonances of the bar and spiral arms can have profound implications for stellar orbits. We do not investigate the effects of bar resonances in this work at all, but recognize that it could be an important caveat. Thus, we give here an overview. The main references are Minchev (2017) and Binney & Tremaine (2008). We consider a bar that is rotating with angular frequency Ω_b. Stars in the galaxy are moving with circular frequency $\Omega_*(R) = v_{\rm circ}(R)/R$ (Equation 1.8).

Co-rotation. Stars at R where $\Omega_*(R) = \Omega_b$ stand still with respect to the triaxial bar; they are co-moving. If the star is located slightly behind (in front of) the bar, it will be constantly accelerated (decelerated) by the non-axisymmetric gravitational force of the bar and moves outwards (inwards), i.e. L_z increases (decreases).

Lindblad resonances. In the galactic rest-frame, a star has the orbital frequencies Ω_* and κ_* (Equation 1.22); in a frame co-moving with the bar it has the circular frequency $\Omega'_* = \Omega_b - \Omega_*$. At radii, where $\kappa_*(R)$ is not an integer multiple of $\Omega'_*(R)$, the orbits are rosettes in the co-moving frame and the acceleration and deceleration due to the bar from different pass-bys cancel each other out. In the case that $\kappa_*(R) = \pm 2\Omega'_*(R)$, the orbit is a closed curve in the frame moving with the bar. The star encounters the bar always in exactly the same configuration and epicycle phase after each period. Hence, the accelerating effect of the bar builds up over time, can significantly change the orbit, and affects both epicycle and circular motion. These two special radii are called the Lindblad resonances. The outer Lindblad resonance in the MW is expected

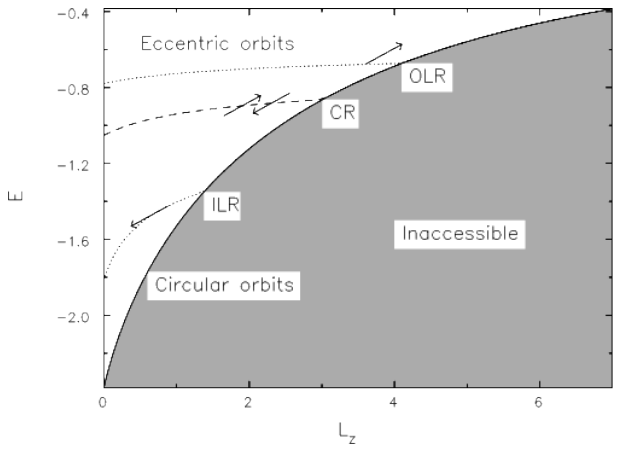


Figure 1.17: Classical Lindblad diagram, showing the energy E vs. the angular momentum L_z of stellar orbits. Circular orbits are located on the dividing line between the accessible and inaccessible integral space. The co-rotation radius (CR), inner and outer Lindblad resonances (ILR and OLR) caused by a periodic potential perturbation with pattern speed Ω_b are shown, where resonances can trigger radial migration along lines of constant Jacobi Energy $E_J = R - L_z \cdot \Omega_b$. **Figure credit:** Sellwood & Carlberg (2014).

to be located in the Solar neighborhood and to be responsible for the observed moving groups (e.g., Dehnen 2000; see also Section 1.2.4).

Lindblad diagram. The above behavior of radial migration around the resonances can be illustrated using the classical Lindblad diagram in Figure 1.17, which shows energy E vs. L_z of the stellar orbits. Above the line of circular orbits, $E_c(L_z)$, lie the eccentric orbits, with $E - E_c(L_z)$ being a measure for the eccentricity. As discussed in the previous sections, axisymmetric time-independent potentials allow three integrals of motions. An axisymmetric potential with a periodic perturbation only allows one, the *Jacobi energy*

$$E_J = E - L_z \cdot \Omega_b = \text{const.} \Rightarrow \delta E = \Omega_b \cdot \delta L_z, \tag{1.57}$$

from which follows that if a star changes its orbit, $\delta E \neq 0$ and $\delta L_z \neq 0$, it will always migrate along lines with slope Ω_b in the Lindblad diagram. At co-rotation resonance (CR), the line of circular orbits $E_c(L_z)$ also has the slope Ω_b. Around the CR, radial migration changes L_z, but there will be no radial heating, as $E - E_c(L_z)$ will stay the same. Circular orbits at the Lindblad resonance experience radial heating in addition to churning of their L_z (Lynden-Bell & Kalnajs 1972). On average, stars that get scattered at the ILR move inwards, while at the OLR they move outwards.

1.5 Dynamical Modeling for the Milky Way Disk

The *Gaia* mission (Perryman et al. 2001) and other Galactic surveys provide data of unprecedented quantity and precision for the positions $\boldsymbol{x}$ and instantaneous velocities $\dot{\boldsymbol{x}} = \boldsymbol{v}$ of stars in the MW. The acceleration $\ddot{\boldsymbol{x}}$ of the stellar movements would constitute a direct measurement of the gravitational forces and therefore mass distribution within our Galaxy. Our current technology, however, is not yet able to detect $\ddot{\boldsymbol{x}}$. This is the principal problem of Galactic dynamics.

The goal of Galactic dynamical modeling is therefore to turn position-velocity data of individual stars into constraints both on the gravitational potential and on the DF of stellar orbits. Yet, rigorous and practical modeling tools of high enough sophistication to fully exploit the information content of the data are scarce (Rix & Bovy 2013).

In this section, we mention some recent modeling attempts in the literature (Sections 1.5.1 and 1.5.4) and introduce RoadMapping, our action-based dynamical modeling approach (Sections 1.5.2-1.5.3) that will be the topic of this PhD thesis.

1.5.1 Overview of Dynamical Modeling Methods

There are a variety of practical approaches to the dynamical modeling of discrete collisionless tracers, such as the stars in the MW. Most of them—explicitly or implicitly—describe the stellar distribution through a DF. Not all of them avoid binning of the data.

- *Jeans modeling* (Jeans 1915). The Jeans equations relate the velocity ellipsoid $\langle v_i v_j \rangle$ with $i, j \in [R, \phi, z]$ of the stars at a given position within the Galaxy to the gravitational forces and the spatial DF of stars. The Jeans equations can be derived from the first velocity moments of the collisionless Boltzmann equation in Equation (1.5). This set of equations is not closed and cannot be uniquely solved. Jeans modeling therefore always requires assumptions on the stellar system in question. If a solution is found, it is not guaranteed that the corresponding DF is physical, i.e., non-negative everywhere. Jeans modeling has been used in the MW to estimate the total surface mass density in the disk (Kuijken & Gilmore 1989b,a, 1991; Zhang et al. 2013b), the DM density at the Sun (Bovy & Tremaine 2012; Garbari et al. 2012; Zhang et al. 2013b), and the velocity ellipsoid tilt (Büdenbender et al. 2015).

- *Schwarzschild's orbit superposition technique* (Schwarzschild 1979). In a given galaxy potential, orbits are integrated. An orbit library is constructed that samples the full integral-of-motion space. By assigning weights to the different orbits (or, in other words, a number of stars or amount of emitted light), a mock observation of the observed luminosity distribution and kinematics can be constructed. These weights constitute effectively the stellar orbit DF of the system. By comparison with observations, the gravitational potential and the DF of the system can be recovered. This modeling is computationally expensive and the resulting numerical DF difficult to handle. Schwarzschild modeling is mostly used in external galaxies (e.g., Rix et al. 1997; van de Ven et al. 2008; Zhu et al. 2016), but recently also for the MW's nuclear star cluster and SMBH (Feldmeier-Krause et al. 2017).

- *Made-to-measure (M2M) modeling* (Syer & Tremaine 1996; de Lorenzi et al. 2007). M2M modeling is based on N-body galaxy simulations. The particle orbits in the simulation are integrated in the composite gravitational potential. The masses or weights of the particles are adapted on-the-fly to improve the fit between the observed data and the N-body model. M2M provides a self-consistent potential and DF pair in the form of particle weights. The advantage is that the method does not require restrictive assumptions, like, e.g., axisymmetry, so it can reproduce bars and spiral arms. It is, however, computationally expensive and the initial conditions are unknown. Bissantz et al. (2004) and Long et al. (2013) used M2M modeling to construct models for the Galactic bulge and bar. M2M modeling has recently been adapted to account for discrete *Gaia*-like data (Hunt & Kawata 2013, 2014; Bovy et al. 2017).

- *Action-based modeling.* Recently, Binney (2012b) and Bovy & Rix (2013) proposed constraining the MW's gravitational potential by combining parameterized axisymmetric potential models with DFs that are simple analytic functions of the three orbital actions to model discrete data. This action-based modeling is formally very similar to Schwarzschild modeling in the sense that it considers galaxies as superposition of orbits: Instead of numerically integrated orbits it uses orbital actions and tori, and instead of discrete orbit weights it employs physically motivated and analytic action-based DFs. Several different approaches have been suggested for action-based modeling for the MW: Bovy & Rix (2013), Piffl et al. (2014), Sanders & Binney (2015b), and Das & Binney (2016) have used parametric DFs and fit for the data by calculating $(\boldsymbol{x}, \boldsymbol{v}) \xrightarrow{\Phi} \boldsymbol{J}$. Magorrian (2014) proposes marginalization over a non-parametric DF. Torus modeling projects orbital tori into phase-space, $\boldsymbol{J} \xrightarrow{\Phi} (\boldsymbol{x}, \boldsymbol{v})$,

which is computationally expensive but allows the data-model comparison directly in the space of observables (McMillan & Binney 2008, 2012, 2013; Binney & McMillan 2016).

1.5.2 The Quasi-isothermal Distribution Function

Action-based DFs. A stellar Distribution Function $\mathrm{DF}(\boldsymbol{x}, \boldsymbol{v})$ can be considered as the probability of a star to be found at $(\boldsymbol{x}, \boldsymbol{v})$. Using orbit DFs in terms of $(\boldsymbol{J}, \boldsymbol{\theta})$ instead, has the advantage that the distribution of stars in $\boldsymbol{\theta}$ is uniform: Due to the linear relation between $\boldsymbol{\theta}$ and time (Equation 1.25), and our assumption that the Galactic disk is axisymmetric and well-mixed in phase (Section 1.4.1), any angle is equally likely to be occupied by a star. Hence, the DF reduces effectively to a function of the actions $\boldsymbol{J}$ only. As $|\partial(\boldsymbol{J}, \boldsymbol{\theta})/\partial(\boldsymbol{x}, \boldsymbol{v})| = 1$, the function $\mathrm{DF}(\boldsymbol{J})$ can still be thought of as a probability in $(\boldsymbol{x}, \boldsymbol{v})$.

The quasi-isothermal Distribution Function (qDF). The action-based quasi-isothermal Distribution Function (qDF) by Binney (2010) and Binney & McMillan (2011) is a simple DF which we will employ as a specific example throughout this work to describe individual stellar sub-populations. This is motivated by the findings of Bovy et al. (2012b,c,d) and Ting et al. (2013) on the simple phase-space structure of stellar MAPs (see Section 1.2.3) and Bovy & Rix (2013)'s successful application (see Section 1.5.3). The qDF has the form

$$\mathrm{qDF}(\boldsymbol{J} \mid p_{\mathrm{DF}}) = f_{\sigma_R}\left(J_R, L_z \mid p_{\mathrm{DF}}\right) \times f_{\sigma_z}\left(J_z, L_z \mid p_{\mathrm{DF}}\right) \tag{1.58}$$

with some free parameters, p_{DF}, and

$$\begin{aligned} f_{\sigma_R}\left(J_R, L_z \mid p_{\mathrm{DF}}\right) &= n(R_g) \times \frac{\Omega}{\pi \sigma_R^2(R_g)\kappa} \exp\left(-\frac{\kappa J_R}{\sigma_R^2(R_g)}\right) \\ &\quad \times \left[1 + \tanh\left(L_z/L_0\right)\right] \qquad (1.59) \\ f_{\sigma_z}\left(J_z, L_z \mid p_{\mathrm{DF}}\right) &= \frac{\nu}{2\pi\sigma_z^2(R_g)} \exp\left(-\frac{\nu J_z}{\sigma_z^2(R_g)}\right) \qquad (1.60) \end{aligned}$$

(Binney & McMillan 2011). Here, R_g, Ω, κ, and ν are functions of L_z and denote, respectively, the guiding-center radius (Equation 1.8), circular frequency (Equation 1.8), radial/epicycle frequency (Equation 1.22), and vertical frequency (Equation 1.23) of the near-circular orbit with angular momentum L_z in a given potential. In practice, we will use the radial action $J_R = J_u$ from Equation (1.51) and the vertical action $J_z = J_v$ from Equation (1.53).

Interpretation and informativeness on the potential. The qDF consists of four terms that are responsible for describing different aspects of the Galactic disk. To better explain them, we first have a look at the Hamiltonian in action-angle coordinates, which is, according to Equations (1.30)-(1.31),

$$\begin{aligned} \mathscr{H}(\boldsymbol{\theta}, \boldsymbol{J}) = \sum_i^3 \Omega_i \cdot J_i &= \Omega_R \cdot J_R + \Omega_\phi \cdot L_z + \Omega_z \cdot J_z \qquad (1.61) \\ &\approx \kappa \cdot J_R + \Omega \cdot L_z + \nu \cdot J_z, \qquad (1.62) \end{aligned}$$

where we have replaced the fundamental frequencies Ω_i with the near-circular orbit frequencies. In the limit of near-circular orbits, these frequencies are equivalent.

- *The vertical profile.* The form of Equation (1.60) was chosen in analogy to $\exp\left(-E_z/\sigma_z^2\right)$, a so-called isothermal sheet (Spitzer 1942; Binney 2010), which is a classical choice for a DF describing the vertical profile of a stellar disk. From the Hamiltonian in Equation

(1.62) and the harmonic oscillator follows, that the average energy stored in the vertical motion is $E_z = \langle v_z^2 \rangle \approx \nu \cdot J_z$ (Binney 2010; Binney & McMillan 2011). This DF generates an approximately constant velocity dispersion σ_z at any height ($\rightarrow$ "isothermal" in z; see Figure 1.18(c)). The qDF allows $\sigma_z(R_g)$ to vary with radius ($\rightarrow$ "quasi-isothermal"). The relation between the maximum height above the plane and the maximum velocity a star can reach, given a fixed value for the action J_z, is set by the potential (see Figure 1.14, panel (e)). The vertical tracer density profile implied by a population with given σ_z is therefore directly informative about the potential. (This relation is also captured by the vertical Jeans equation; see Binney & Tremaine (2008), §4.8.2; see also Section 4.7.1, where we will apply the vertical Jeans equation).

- *The underlying radial tracer density profile.* $n(R_g)$ in Equation (1.59) sets the approximate radial surface density profile of the tracers. If all stars were currently at their guiding-center radius (i.e., orbital mean radius), or all on circular orbits (i.e., $\sigma_R = \sigma_z = 0$), $n(R_g)$ would exactly describe the radial tracer density profile.

- *The radial profile.* Orbits oscillate around their R_g, with the stars staying longer at their apo-center and shorter at their peri-center distance. The true tracer surface density at R is therefore determined by the number of circular orbits with $R_g = R$, plus the number of stars with apo-center at R (i.e., $n(R_g)$ at $R_g < R$), plus the number of stars with peri-center at R (i.e., $n(R_g)$ at $R_g > R$). We expect, however, more stars from the inside than from the outside to currently reside at R ("asymmetric drift"). As we saw in Figure 1.14, panel (d), the relation between the maximum radial velocity and maximum excursion at given J_R is set by the potential. If we had a realistic model for $n(R_g)$, and a good model describing the distribution of J_R at given R_g, the true observed number of stars at R would be directly informative about the potential.

  A model for the J_R distribution is given in Equation (1.59), which is motivated by Shu (1969), who proposed $\exp\left(-(E_{\rm inplane} - E_c)/\sigma_R^2\right)$ for the distribution of orbits in a planar disk with radial velocity dispersion σ_R. E_c is the energy of a circular orbit at R_g, and $E_{\rm inplane}$ the energy of the in-plane motion. The Hamiltonian (Equation 1.62) indicates that the energy that is stored in the radial oscillation only is $E_R = E_{\rm inplane} - E_c \approx \kappa \cdot J_R$ (Binney 2010). Properly normalized, this leads to the $\exp\left(-\frac{\kappa J_R}{\sigma_R^2(R_g)}\right)$ term in Equation (1.59).

- *Suppressing counter-rotation.* The term $[1 + \tanh(L_z/L_0)]$ in Equation (1.59) suppresses counter-rotation for orbits in the disk with $L_z < L_0$ (with $L_0 \sim 10$ km s^{-1} kpc).

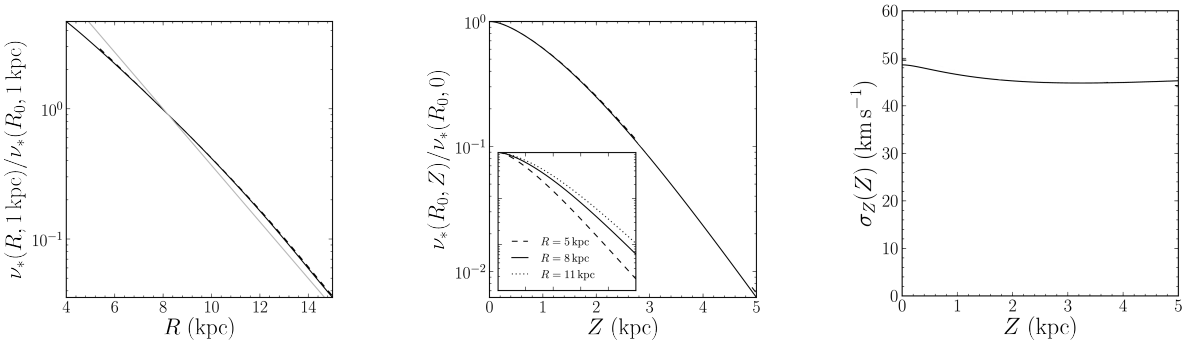
Scaling profiles. We are left with the decision for a good model for the stellar number density $n(R_g)$, and the radial and vertical velocity dispersion profiles $\sigma_R(R_g)$ and $\sigma_z(R_g)$. Following Bovy & Rix (2013), we choose the functional forms

$$n(R_g \mid p_{\rm DF}) \propto \exp\left(-\frac{R_g}{h_R}\right) \tag{1.63}$$

$$\sigma_R(R_g \mid p_{\rm DF}) = \sigma_{R,0} \times \exp\left(-\frac{R_g - R_\odot}{h_{\sigma,R}}\right) \tag{1.64}$$

$$\sigma_z(R_g \mid p_{\rm DF}) = \sigma_{z,0} \times \exp\left(-\frac{R_g - R_\odot}{h_{\sigma,z}}\right). \tag{1.65}$$

The qDF has a set of five free parameters $p_{\rm DF}$: 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



(a) Radial tracer density distribution.

(b) Vertical tracer density distribution.

(c) Vertical velocity dispersion as function of height.

Figure 1.18: Projection of the qDF (for some fiducial model parameters and potential) into the space of observables (using the Stäckel fudge for the action estimation). Panel 1.18(a) shows that the radial tracer density profile implied by the qDF (black) is close to a pure exponential (grey). Panel 1.18(b) demonstrates that the vertical tracer density resembles a flattened exponential which flares slightly with radius. Panel 1.18(c) shows the vertical velocity dispersion which is almost constant with height above the plane—the characteristic that gives the qDF its name. **Figure credit:** Bovy & Rix (2013).

lengths $h_{\sigma,R}$ and $h_{\sigma,z}$, that describe the radial decrease of the velocity dispersion. Figure 1.18 shows a projection of the qDF into observable space, demonstrating the exponential profile and the quasi-isothermal velocity dispersion.

1.5.3 Introducing RoadMapping Modeling

Idea and Motivation: Pulling everything together. In Sections 1.1 and 1.2, we have set the scene for the MW in the context of galaxy formation and evolution and have pointed to several open questions. Based on this, we have motivated in Section 1.3 why learning about the gravitational potential and stellar DF of the MW is crucial. We have learned that the motions of the stars are the fossil records of the MW's evolution (Section 1.3.1) and that the proper physical way to consider stellar motions is in terms of orbits and orbit-based DFs (Section 1.4). When dealing with orbits, it is most convenient to use actions as labels for the orbits (Section 1.4.3). An orbit (or an action calculation) needs a single 6D phase-space observation for a star as starting point, and the gravitational potential under whose forces the star subsequently moves. Fitting an orbit DF to observed $(\boldsymbol{x}, \boldsymbol{v})$ of a population of stars inherently contains the requirement to either know the true potential, or to directly fit a potential model simultaneously with the DF. This is the core principle of the orbit-based modeling techniques presented in Section 1.5.1.

In theory, one would have to test all (!) possible combinations of gravitational potentials and DFs to find the pair that makes the observed data most likely. This is, in practice, not feasible. We have to resort to parametric models for the potential and DF, whose free model parameters offer enough flexibility to be able to capture the main features of the true potential and DF.

The largest amount of 6D phase-space measurements ever taken for a stellar population is by far that of the MW disk (see Section 1.3.4). However, the chemo-kinematical structure of the MW disk is highly complex (see Section 1.2). A lot of effort has been put into constructing DF models that can describe the whole disk at once (e.g., Sanders & Binney 2015b). However, the more complex a model is, the more difficult it is to interpret the fitting results.

In Section 1.2.3, we have learned that the complexity of the MW disk can be broken down into simple functional forms when considering MAPs in [Fe/H] and [α/Fe] separately (Bovy et al. 2012b,d,c, 2016b). Ting et al. (2013) and Bovy & Rix (2013) proposed to make use of this in

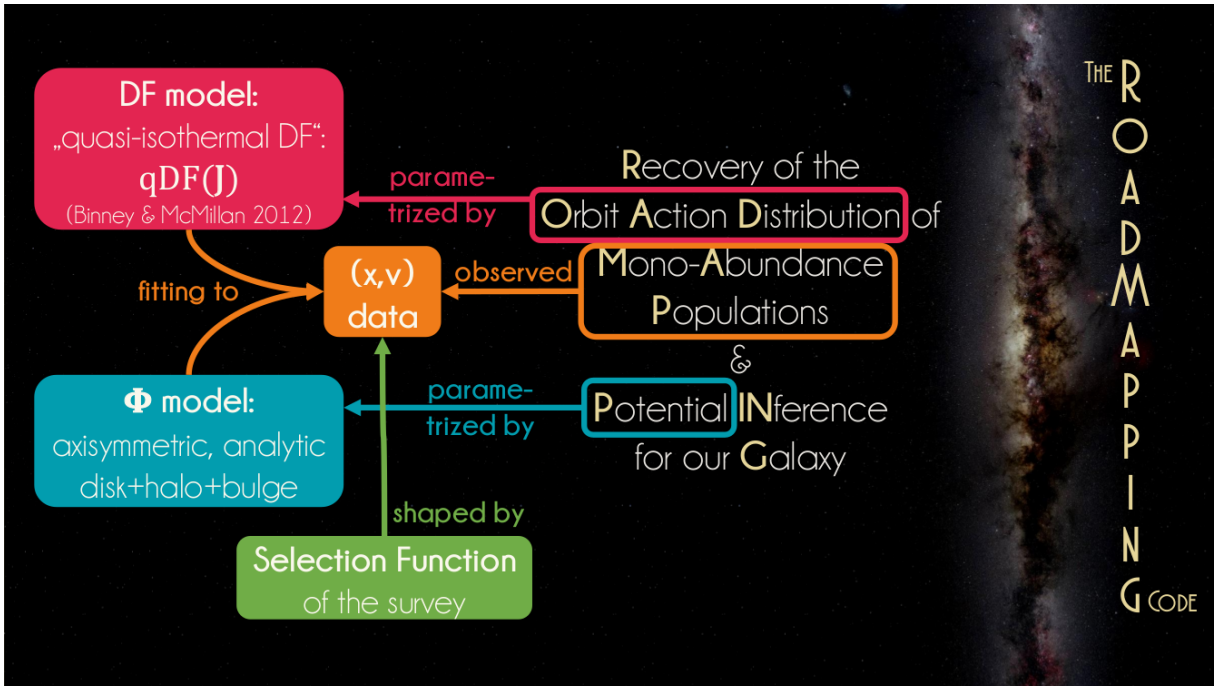


Figure 1.19: Schematic idea behind the RoadMapping modeling approach.

dynamical modeling of the MW. In particular, Ting et al. (2013) showed that the very simple qDF (Binney & McMillan 2011), which we have introduced in Section 1.5.2, is flexible enough to capture the phase-space structure of MAPs. They also showed, using mock data, that the approach of fitting a qDF to a single MAP could indeed work to recover the MW's potential. Thus, the idea of RoadMapping was born:

RoadMapping:
Recovery of the **O**rbit **A**ction **D**istribution
of **M**ono-**A**bundance **P**opulations
and **P**otential **IN**ference
for our **G**alaxy

Modeling characteristics. Using MAPs has several implications.

(i) All stars of different MAPs move in the same gravitational potential. By modeling several MAPs separately, we get independent constraints on the same potential. This offers good cross-checking possibilities.

(ii) We do not use any stellar abundance information explicitly in the dynamical modeling. By using MAPs, the chemistry is, however, implicitly taken into account.

(iii) Using MAPs is purely empirically motivated. It does not assume that the stars in a given MAP were born together and/or have the same age. Still, it is reminiscent of chemical tagging, as it investigates the orbits of stars with similar abundances (see Section 1.3.1).

(iv) RoadMapping is no self-consistent modeling in which the tracer DF also implies (part of) the total matter distribution. Potential model and qDF model for a single sub-population are considered as independent in RoadMapping. Because a single MAP has negligible mass as compared to the whole disk, this assumption is acceptable.

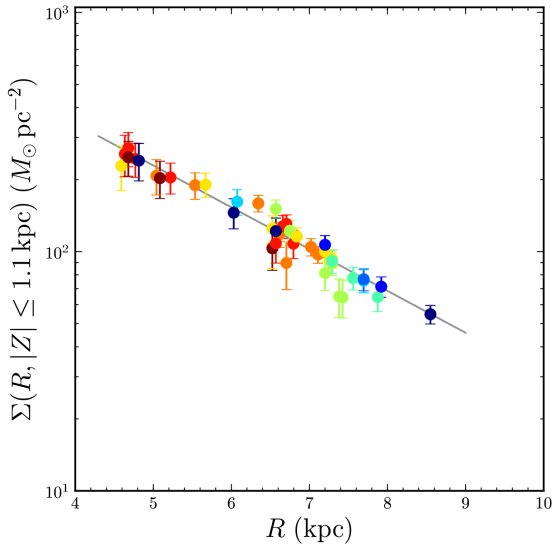


Figure 1.20: Total surface mass density profile for the MW disk within $|z| \leq 1.1$ kpc measured by Bovy & Rix (2013) from MAPs of SEGUE G-dwarfs using RoadMapping modeling. For each MAP, the surface density was determined at one best radius (see Sections 4.6.2 and 4.7.2). Taken together, they implied an exponential surface density profile with scale length $R_s = 2.5 \pm 0.1$ kpc. Each dot is color-coded by the $[\alpha/\mathrm{Fe}]$ of the respective MAP, ranging from $[\alpha/\mathrm{Fe}] = 0.025$ (blue) to $[\alpha/\mathrm{Fe}] = 0.475$ (brown). **Figure credit:** Bovy & Rix (2013).

RoadMapping constructs purely axisymmetric models for the MW. In non-axisymmetric potentials, the action estimation becomes very difficult. The orbit DF assumes that the disk is fully phase-mixed (i.e., independent of the angles and a function of the actions only) and therefore by construction axisymmetric.
One crucial modeling ingredient of RoadMapping is the SF of the survey that provides the observed stellar positions and velocities (Rix & Bovy 2013, see also Section 1.3.4). Brightness limits and incompleteness in the SF can shape the observed distribution strongly and even remove all signatures of the underlying physical DF if we do not correct for it.
We will describe the actual fitting process in detail in Section 2.2. Here, we only mention the philosophy behind it: "Binning is sinning!" All the information should be taken into account. For each star, we ask independently what is the probability of this star coming from the physical DF in the potential, given its position and velocity with measurement uncertainties, and given the survey SF. We do this in a full Bayesian framework, that also allows us to include previous knowledge in the form of priors.

The first RoadMapping application. Bovy & Rix (2013) were the first to put the RoadMapping modeling approach described above into practice. They used 43 MAPs of SDSS/SEGUE G-dwarf stars (Yanny et al. 2009) in the Galactic disk in the abundance range $[\mathrm{Fe/H}] \in [-1.3, 0.3]$ and $[\alpha/\mathrm{Fe}] \in [0.0, 0.5]$. They accounted for the complex but known SF of the kinematic tracers (Bovy et al. 2012d). To avoid spiral arm effects, they did not use in-plane motions and marginalized the likelihood over the radial and tangential velocity components. qDF parameters governing the radial velocities were not fitted. For each MAP the modeling resulted in an independent estimate of the same gravitational potential. Their potential model had only two free parameters (disk scale length and relative halo-to-disk contribution to the radial force at the solar radius). To account for missing model flexibility, they constrained the surface mass density for each MAP only at one best radius. Taken as an ensemble, they constrained the disk surface density within 1.1 kpc (the mode height of the data) over a wide range of radii ($\sim 4 - 9$ kpc; see Figure 1.20). The profile they derived in this fashion had a scale length of $R_s = 2.5$ kpc and was in the regime $R > 6.6$ kpc later confirmed by Piffl et al. (2014) using a different action-based procedure. Combining their measurements of the vertical force (which is equivalent to the surface density at $z = 1.1$ kpc) with previous constraints, they managed to decompose the rotation curve into halo and disk component (see Figure 1.13).

This successful application was the starting point for the in-depth analysis of the characteristics of RoadMapping modeling in this thesis and for a wide range of improvements, which we will describe and contrast with Bovy & Rix (2013) in the following chapters.

1.5.4 Other Modeling Approaches using Action-based DFs

In some sense, the RoadMapping approach focuses on constraining the potential, treating the DF parameters as nuisance parameters.
Magorrian (2014) introduced a framework that avoids specific parameterizations of action-based DFs and marginalizes over all possible DFs to constrain the potential. While this is the proper way to treat a nuisance DF, it appears to be computationally very challenging.
Since the first application of RoadMapping by Bovy & Rix (2013), there have been two similar efforts to constrain the Galactic potential and/or orbit DF for the disk.
Piffl et al. (2014) fitted both a potential and a DF($\boldsymbol{J}$) to giant stars from the RAVE survey (Steinmetz et al. 2006) and the vertical stellar number density profiles in the disk by Jurić et al. (2008). They did not include any chemical abundances in the modeling. Instead, they used a superposition of action-based DFs to describe the overall stellar distribution at once: a superposition of qDFs for cohorts in the thin disk, a single qDF for the thick disk stars, and an additional DF for the halo stars. Taking proper care of the SF requires a full likelihood analysis, which is computationally expensive. Piffl et al. (2014) choose to circumvent this difficulty by directly fitting (i) histograms of the three velocity components in eight spatial bins to the velocity distribution predicted by the DF and (ii) the vertical density profile predicted by the DF to the profiles by Jurić et al. (2008). The vertical force profile of their best fit mass model nicely agrees with the results from Bovy & Rix (2013) for $R > 6.6$ kpc. The disadvantage of their approach is that by binning the stars spatially, a lot of information is not used.
Sanders & Binney (2015b) focused on understanding the abundance-dependence of the DF, relying on a fiducial potential. They developed extended distribution functions (eDF), i.e., functions of both actions and metallicity for a superposition of thin and thick disks, each consisting of several cohorts described by qDFs, a DF for the halo, a functional form of the metallicity of the interstellar medium at the time of birth of the stars, and a simple prescription for radial migration. They applied a full likelihood analysis accounting for selection effects and found a best fit for the eDF in the fixed fiducial potential by Dehnen & Binney (1998) to the stellar phase-space data of the Geneva-Copenhagen Survey (Nordström et al. 2004; Holmberg et al. 2009), metallicity determinations by Casagrande et al. (2011), and the stellar density curves by Gilmore & Reid (1983). Their best fit predicted the velocity distribution of SEGUE G-dwarfs (Ahn et al. 2014) quite well, but had biases in the metallicity distribution, which they accounted to being a problem with the SEGUE metallicities.
Das & Binney (2016) proceeded recently in a similar fashion to constrain an eDF for halo stars.

CHAPTER 2

RoadMapping Modeling & the Breakdown of its Assumptions

2.1 Preface

Introduction. We have motivated the idea behind RoadMapping dynamical modeling for the MW disk in Section 1.5.3. Bovy & Rix (2013) first suggested this approach to constrain the gravitational potential and action-based DF and performed the first application of RoadMapping to SEGUE G-dwarf stars. They made, however, a number of quite severe and idealizing assumptions about the potential, the DF, and the knowledge of observational effects. These idealizations could plausibly translate into systematic errors on the inferred potential, well above the formal error bars of upcoming surveys with their wealth and quality of data.
In this chapter, we present an improved, refined, flexible, robust, and well-tested version of the original dynamical modeling machinery by Bovy & Rix (2013). In addition, our goal is to explore which of the assumptions that Bovy & Rix (2013) made and which other aspects of data, model, and machinery limit RoadMapping's recovery of the true gravitational potential. We investigate this through differential test cases with idealized mock data.

Scope and structure of this chapter. We examine the following aspects of the RoadMapping machinery that become especially important for a large number of stars:

(i) Numerical inaccuracies must not be an important source of systematics (Section 2.2.8).

(ii) As parameter estimates become much more precise, we need more flexibility in the potential and DF model and efficient strategies to find the best fit parameters. The improvements made in RoadMapping as compared to the machinery used in Bovy & Rix (2013) are presented in Section 2.2.10.

(iii) RoadMapping should be an unbiased estimator (Section 2.3.1).

We also explore how different aspects of the observational experiment design impact the parameter recovery:

(i) We consider the importance of the survey volume geometry, size, shape, and position within the MW to constrain the potential (Section 2.3.2).

(ii) We ask what happens if our knowledge of the sample SF is imperfect, and potentially biased (Section 2.3.3).

(iii) We investigate how to best account for individual, and possibly misjudged, measurement uncertainties (Section 2.3.4).

(iv) We determine which of several stellar sub-populations is best for constraining the potential (Section 2.3.7).

One of the strongest assumptions is restricting the dynamical modeling to a certain family of parameterized functions for the gravitational potential and the DF. We investigate how well we can hope to recover the true potential, when our models do not encompass the true

(i) DF (Section 2.3.5)

(ii) and potential (Section 2.3.6).

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 outlined in Sections 2.2.1-2.2.4 following the procedure in Sections 2.2.5-2.2.6 and analyzed according to the description of the RoadMapping machinery in Sections 2.2.7-2.2.10. Section 2.3 compiles our results on the investigated modeling aspects. In particular, our key results about the systematics introduced by using wrong DF or potential models are presented in the Sections 2.3.5 and 2.3.6. Section 2.4 finally summarizes and discusses our findings.

Attribution. This chapter is based on the published and peer-reviewed paper by Trick et al. (2016a). All the ideas in this work were developed in close consultation with **Hans-Walter Rix (MPIA, Heidelberg)** and **Jo Bovy (Uni Toronto)**.
My accomplishments in this paper are:

- Developing a well-tested dynamical modeling code implementing the RoadMapping concept.
- Coding and performing a vast suite of mock data tests, investigating many different aspects of RoadMapping modeling, and proving its robustness, even if some model assumptions break down.
- Developing from scratch a range of different geometrical SF models that can be used (i) together with potential and DF models from the `galpy python` library (Bovy 2015) to create mock data sets and (ii) for the dynamical modeling in the RoadMapping machinery.
- Implementing the `KuzminKutuzovStaeckelPotential`, now part of the `galpy` package.
- Writing all the text, interpretations, and conclusions, as well as creating all the plots in the original paper, Trick et al. (2016a), which found their way in large parts into this chapter.
- Coming up with the acronym RoadMapping, *Recovery of the Orbit Action Distribution of Mono-Abundance Populations and Potential INference for our Galaxy.*

My collaborators specifically contributed the following:

- **Jo Bovy (Uni Toronto)** provided a simple version of the original RoadMapping code by Bovy & Rix (2013) as starting point for this work. It was, however, later completely rewritten and restructured by me. My code still uses various components from Jo Bovy's `galpy` package, including implementations of potential and DF models, and action estimation methods.

- **Hans-Walter Rix (MPIA, Heidelberg)** suggested the main aspects that should be tested with RoadMapping in this chapter. He also helped with the phrasing of a few sentences in this paper.

In addition, we thank

- **Glenn van de Ven (MPIA, Heidelberg)** for suggesting the use of Kuzmin-Kutuzov Stäckel potentials in this case study.
- the anonymous referee for her/his very helpful comments and suggestions.
- **James J. Binney** and **Payel Das (University of Oxford)** for valuable discussions.
- **Rene Andrae (MPIA)** for an estimate of *Gaia* distance uncertainties.

2.2 RoadMapping Dynamical Modeling of Mock Data

In this section we summarize the basic elements of RoadMapping, the dynamical modeling machinery presented in this work, which in many respects follows Bovy & Rix (2013) and makes extensive use of the `galpy` Python package for galactic dynamics[4] (Bovy 2015).

2.2.1 Coordinate System

Our modeling takes place in the Galactocentric rest-frame with cylindrical coordinates $\boldsymbol{x} \equiv (R, \phi, z)$ and corresponding velocity components $\boldsymbol{v} \equiv (v_R, v_T, v_z)$. If the stellar phase-space data is given in observed heliocentric coordinates, position $\tilde{\boldsymbol{x}} \equiv (\text{R.A.}, \text{decl.}, m - M)$ in right ascension R.A., declination decl. and distance modulus $(m - M)$, and velocity $\tilde{\boldsymbol{v}} \equiv (\mu_{\text{R.A.}} \cdot \cos(\text{decl.}), \mu_{\text{decl.}}, v_{\text{los}})$ as proper motions and line-of-sight velocity, the data $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{v}})$ has to be converted into the Galactocentric rest-frame coordinates $(\boldsymbol{x}, \boldsymbol{v})$ using the Sun's position and velocity. We assume for the Sun

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

2.2.2 Action Estimation

According to Sanders & Binney (2016) the best compromise of speed and accuracy for the Galactic disk is the *Stäckel fudge* by Binney (2012a) for axisymmetric potentials (see Section 1.4.4). In addition, we use action interpolation grids (Binney 2012a; Bovy 2015) to speed up the calculation. The latter is one of the improvements employed by RoadMapping, which was not used in Bovy & Rix (2013).

2.2.3 Potential Models

For the gravitational potential in our modeling we assume a family of parameterized models. We use: a MW-like potential with disk, halo, and bulge (`DHB-Pot`); the spherical isochrone potential (`Iso-Pot`); and the 2-component Kuzmin-Kutuzov Stäckel potential (Batsleer & Dejonghe 1994; `KKS-Pot`), which also displays a disk and halo structure. Table 2.1 summarizes all reference potentials used in this chapter together with their free parameters, p_Φ. The true circular velocity at the Sun was chosen to be $v_{\text{circ}}(R_\odot) = 230$ km s^{-1} for all potential models. The `Iso-Pot`

[4] `galpy` is an open-source code that is being developed on `http://github.com/jobovy/galpy`. The latest documentation can be found at `http://galpy.readthedocs.org/en/latest/`.

Table 2.1. Axisymmetric gravitational potential models used throughout this chapter.

Name	Potential Model	Parameters p_Φ		Action Calculation
`Iso-Pot`	isochrone potential[(a)] (Henon 1959)	b	0.9 kpc	*analytic and exact* (Binney & Tremaine 2008, §3.5.2)
`KKS-Pot`	2-component Kuzmin-Kutuzov- Stäckel potential![(b)] disk and halo (Batsleer & Dejonghe 1994)	Δ $\left(\frac{a}{c}\right)_{\text{Disk}}$ $\left(\frac{a}{c}\right)_{\text{Halo}}$ k	0.3 20 1.07 0.28	*exact* using interpolation on action grid (Binney 2012a; Bovy 2015)
`DHB-Pot`	Disk+Halo+Bulge potential![(c)] Miyamoto-Nagai disk, NFW halo, Hernquist bulge (same as `MW14-Pot`, except of bulge)	a_{disk} b_{disk} f_{halo} a_{halo} f_{bulge} a_{bulge}	3 kpc 0.28 kpc (fixed) 0.35/0.95 16 kpc (fixed) 0.05/1.0 (fixed) 0.6 kpc (fixed)	*approximate* using *Stäckel fudge* (Binney 2012a) and interpolation on action grid
`MW14-Pot`	MW-like potential![(d)] Miyamoto-Nagai disk, NFW halo, cut-off power-law bulge (Bovy 2015)			*approximate* using *Stäckel fudge*

Note. — The potential parameters are fixed for the mock data creation at the values given in this table, which we susequently aim to recover with RoadMapping. The parameters of `DHB-Pot` and `KKS-Pot` were chosen to resemble the `MW14-Pot` (see Figure 2.1). We use $v_{\text{circ}}(R_\odot) = 230$ km s^{-1} as the circular velocity at the Sun for all potentials in this chapter.

[(a)] The free parameter of the spherical `Iso-Pot` is the isochrone scale length b.

[(b)] The coordinate system of each of the two Stäckel-potential components of the `KKS-Pot` is $R^2/(\tau_{i,p}+\alpha_p) + z^2/(\tau_{i,p}+\gamma_p) = 1$ with $p \in \{\text{Disk}, \text{Halo}\}$ and $\tau_{i,p} \in \{\lambda_p, \nu_p\}$. Both components have the same focal distance $\Delta \equiv \sqrt{\gamma_p - \alpha_p}$, to ensure that the superposition itself is a Stäckel potential. The axis ratio of the coordinate surfaces, $(a/c)_p := \sqrt{\alpha_p/\gamma_p}$, describes the flatness of each component. k is the relative contribution of the disk mass to the total mass.

[(c)] The parameters of the `DHB-Pot` are the Miyamoto-Nagai disk scale length a_{disk} and height b_{disk}, the NFW halo scale length a_{halo} and its relative contribution to $v^2_{\text{circ}}(R_\odot)$ with respect to the total disk+halo contribution, f_{halo}, and the Hernquist bulge scale length, a_{bulge}, and its contribution to the total $v^2_{\text{circ}}(R_\odot)$, f_{bulge}. We keep all except $v_{\text{circ}}(R_\odot), a_{\text{disk}}$ and f_{halo} fixed to their true values in the analysis.

[(d)] The `MWPotential2014` by Bovy (2015) (see their Table 1) has $v_{\text{circ}}(R_\odot) = 220$ km s^{-1}. We use, however, $v_{\text{circ}}(R_\odot) = 230$ km s^{-1}.

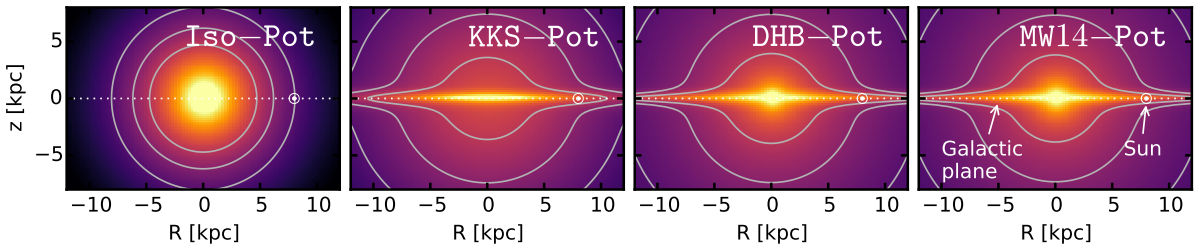


Figure 2.1: Density distribution of the four reference galaxy potentials in Table 2.1. These potentials are used throughout this chapter to create and model mock data with RoadMapping.

Table 2.2. Reference Parameters for the qDF in Equations (1.58)-(1.65).

Name	qDF Parameters $p_{\rm DF}$				
	h_R [kpc]	$\sigma_{R,0}$ [km s^{-1}]	$\sigma_{z,0}$ [km s^{-1}]	$h_{\sigma,R}$ [kpc]	$h_{\sigma,z}$ [kpc]
`hot`	2	55	66	8	7
`cool`	3.5	42	32	8	7
`cooler`	3	27.5	33	8	7
`colder`	$2+X\%$	$55-X\%$	$66-X\%$	8	7
`warmer`	$3.5-X\%$	$42+X\%$	$32+X\%$	8	7

Note. — These parameters are used to create 6D phase-space mock data sets for stellar populations of different kinematic temperature. The parameter X is explained in Section 2.2.4.

allows both accurate and particularly fast action calculations (see Section 1.4.4); we therefore use it for tests requiring a large number of analyses. The `KKS-Pot` and `DHB-Pot` were chosen for their more realistic galaxy shape and because their closed-form expression for $\Phi(R,z)$ makes the computation of forces and densities fast and easy. The `KKS-Pot` also allows for exact action calculations, while the `DHB-Pot` has physically more intuitive potential parameters but requires the *Stäckel fudge* to estimate actions. The parameter values of `KKS-Pot` and `DHB-Pot` in Table 2.1 were chosen to resemble the MW potential from Bovy (2015) (`MW14-Pot`). The density distribution of all these potentials is illustrated in Figure 2.1.

2.2.4 Stellar DF Models and Tracer Density

To describe the phase-space distribution of stars in this study, we use the stellar DF introduced and motivated in Section 1.5.2: the qDF. According to Equations (1.58)-(1.65), it has five free parameters $p_{\rm DF}$: 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. RoadMapping allows to fit any number of DF parameters simultaneously, while Bovy & Rix (2013) kept $\{\sigma_{R,0}, h_{\sigma,R}\}$ fixed.

Throughout, this chapter we make use of a few example stellar populations whose qDF parameters are given in Table 2.2. Most tests use the `hot` and `cool` qDFs, which correspond to kinematically hot and cool populations, respectively. The `warmer` (`cooler` and `colder`) qDFs in Table 2.2 were chosen to have the same anisotropy $\sigma_{R,0}/\sigma_{z,0}$ as the `cool` (`hot`) qDF, with X being a free parameter describing the temperature difference. Hotter populations have shorter tracer scale lengths (Bovy et al. 2012d) and the velocity dispersion scale lengths were fixed according to Bovy et al. (2012c).

One indispensable step in our dynamical modeling technique (Sections 2.2.7-2.2.8), as well as in creating mock data (Section 2.2.6), is to calculate the (axisymmetric) spatial tracer density $\rho_{\rm DF}(\boldsymbol{x} \mid p_\Phi, p_{\rm DF})$ for a given DF and potential. Analogously to Bovy & Rix (2013),

$$\begin{aligned}
&\rho_{\rm DF}(R, |z| \mid p_\Phi, p_{\rm DF}) \\
&= \int_{-\infty}^{\infty} {\rm DF}(\boldsymbol{J}[R,z,\boldsymbol{v} \mid p_\Phi] \mid p_{\rm DF})\, {\rm d}^3 v \\
&\approx \int_{-n_\sigma\sigma_R(R|p_{\rm DF})}^{n_\sigma\sigma_R(R|p_{\rm DF})} \int_{-n_\sigma\sigma_z(R|p_{\rm DF})}^{n_\sigma\sigma_z(R|p_{\rm DF})} \int_0^{1.5 v_{\rm circ}(R_\odot)} \\
&\qquad {\rm DF}(\boldsymbol{J}[R,z,\boldsymbol{v} \mid p_\Phi] \mid p_{\rm DF})\, {\rm d}v_T\, {\rm d}v_z\, {\rm d}v_R,
\end{aligned} \tag{2.1}$$

where $\sigma_R(R \mid p_{\rm DF})$ and $\sigma_z(R \mid p_{\rm DF})$ are given by Equations (1.64) and (1.65).[5] Each integral is evaluated using a N_vth order Gauss-Legendre quadrature. For a given p_Φ and $p_{\rm DF}$ we explicitly calculate the density on $N_x \times N_x$ regular grid points in the (R, z) plane and interpolate $\ln \rho_{\rm DF}$ in between using bivariate spline interpolation. The grid is chosen to cover the extent of the observations (for $|z| \geq 0$, because the model is symmetric in z by construction). The total number of actions to be calculated to set up the density interpolation grid is $N_x^2 \times N_v^3$, which is one of the factors limiting the computation speed. To complement the work by Bovy & Rix (2013), we will specifically work out in Section 2.2.8 and Figure 2.4 how large N_x, N_v and n_σ have to be chosen to get the density with a sufficiently high numerical accuracy.

2.2.5 Selection Functions

Any survey's SF can be understood as defining an effective sample sub-volume in the space of observables, e.g., position on the sky (limited by the pointing of the survey), distance from the Sun (limited by brightness and detector sensitivity), colors, and metallicity of the stars (limited by survey mode and targeting). The SF can therefore be thought of as having both a spatial small scale structure (due to pencil beam pointing, dust obscuration, etc.) and some overall spatial characteristics (e.g., mean height above the plane and mean Galactocentric radius of the stars). The treatment of realistic and complex SFs was already demonstrated in Bovy & Rix (2013) (who used the pencil beam SF of the SEGUE survey; Bovy et al. 2012d) and Bovy et al. (2016a) (who investigated the effect of dust extinction). In this chapter, we aim to make a generic and basic exploration of search volume shapes and, as shown by Bovy et al. (2016a), this should be possible without explicitly considering the spatial SF substructure. Inspired by the contiguous nature of the *Gaia* SF, which is basically only limited by a magnitude cut, and the fact that this magnitude cut would—in the absence of small scale structure—translate to a sharp distance cut for standard candle tracer populations like red clump stars, we therefore use in our modeling a simple spatial SF of spherical shape with radius $r_{\rm max}$ around the Sun,

$$\mathrm{SF}(\boldsymbol{x}) \equiv \begin{cases} \text{completeness}(\boldsymbol{x}) & \text{if } |\boldsymbol{x} - \boldsymbol{x}_\odot| \leq r_{\rm max}, \\ 0 & \text{otherwise}, \end{cases} \tag{2.2}$$

We set $0 \leq \text{completeness}(\boldsymbol{x}) \leq 1$ everywhere inside the observed volume, so it can be understood as a position-dependent detection probability for a star at $\boldsymbol{x}$. Unless explicitly stated otherwise, we simplify to $\text{completeness}(\boldsymbol{x}) = 1$. Additionally, we use in Figure 2.8 (Test 4 in Table 2.3) and Figures 2.2-2.3 for illustrative purposes some rather unrealistic survey volumes which are angular segments of a cylindrical annulus (wedge), i.e., the volume with $R \in [R_{\rm min}, R_{\rm max}], \phi \in [\phi_{\rm min}, \phi_{\rm max}], z \in [z_{\rm min}, z_{\rm max}]$ within the model Galaxy.

2.2.6 Creating Mock Data

The mock data in this chapter is generated according to the following procedure.
We assume that the positions and velocities of our stellar mock sample are indeed drawn from our assumed family of potentials (Section 2.2.3) and DFs (Section 2.2.4), with given parameters p_Φ and $p_{\rm DF}$. The DF is in terms of actions, while the transformation $(\boldsymbol{x}_i, \boldsymbol{v}_i) \xrightarrow{\Phi} \boldsymbol{J}_i$ is computationally much less expensive than its inversion. We therefore employ the following efficient two-step method for creating mock data, which also accounts for a spatial survey SF$(\boldsymbol{x})$ in Section 2.2.5. In the first step we draw stellar positions $\boldsymbol{x}_i$. We start by setting up the interpolation grid for

[5] The integration ranges over the velocities are motivated by Figure 2.3 and n_σ should be chosen as $n_\sigma \sim 5$ (see Figure 2.4). The integration range $[0, 1.5 v_{\rm circ}(R_\odot)]$ over v_T is, in general, sufficient, only for observation volumes with larger mean stellar v_T does this upper limit need to be increased.

the tracer density $\rho(R, |z| \mid p_\Phi, p_{\rm DF})$ generated according to Section 2.2.4.[6] Next, we sample random positions (R_i, z_i, ϕ_i) uniformly within the observable volume. Using an Monte Carlo (MC) rejection method we then shape the samples distribution to follow $\rho(R, |z| \mid p_\Phi, p_{\rm DF})$. To apply a non-uniform completeness function, we use the rejection method a second time. The resulting set of positions, $\boldsymbol{x}_i$, follows the distribution $p(\boldsymbol{x}) \propto \rho_{\rm DF}(R, |z| \mid p_\Phi, p_{\rm DF}) \times {\rm SF}(\boldsymbol{x})$.
In the second step we draw velocities, $\boldsymbol{v}_i$. For each of the positions (R_i, z_i) we first sample velocities from a Gaussian envelope function in velocity space which is then shaped toward ${\rm DF}(\boldsymbol{J}[R_i, z_i, \boldsymbol{v} \mid p_\Phi] \mid p_{\rm DF})$ using a rejection method. We now have a mock data set satisfying $(\boldsymbol{x}_i, \boldsymbol{v}_i) \longrightarrow p(\boldsymbol{x}, \boldsymbol{v}) \propto {\rm DF}(\boldsymbol{J}[\boldsymbol{x}, \boldsymbol{v} \mid p_\Phi] \mid p_{\rm DF}) \times {\rm SF}(\boldsymbol{x})$.
Measurement uncertainties can be added to the mock data by applying the following modifications to the above procedure. We assume Gaussian uncertainties in the heliocentric phase-space coordinates $\tilde{\boldsymbol{x}} = ({\rm R.A.}, {\rm decl.}, (m - M)), \tilde{\boldsymbol{v}} = (\mu_{\rm R.A.} \cdot \cos {\rm decl.}, \mu_{\rm decl.}, v_{\rm los})$ (see Section 2.2.1). In the case of distance and position uncertainties, stars virtually scatter in and out of the observed volume. To account for this, we draw the *true* $\boldsymbol{x}_i$ from a volume that is larger than the actual observation volume, perturb the $\boldsymbol{x}_i$ according to the position uncertainties and then reject all stars that now lie outside of the observed volume. This mirrors the random scatter around the detection threshold for stars whose distances are determined from the apparent brightness and the distance modulus. We then sample *true* $\boldsymbol{v}_i$ (given the *true* $\boldsymbol{x}_i$) as described above and perturb them according to the velocity uncertainties.
We show examples of mock data sets (without measurement uncertainties) in action space (Figure 2.2) and configuration space $(\boldsymbol{x}, \boldsymbol{v})$ (Figure 2.3). The mock data generated from the qDF follow the expected distributions in configuration space. 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. Circular orbits with $J_R = 0$ and $J_z = 0$ can only be observed in the Galactic mid-plane. The different ranges of angular momentum, L_z, in the two example observation volumes reflect $L_z \sim R \times v_{\rm circ}$ and the volumes' different radial extent. The volume at larger z contains stars with higher J_z. An orbit with $L_z \ll$ or $\gg L_z(R_\odot)$ can only reach into a volume at $\sim R_\odot$, if it is more eccentric and has therefore larger J_R. This, together with the effect of asymmetric drift, explains the asymmetric distribution of J_R vs. L_z in Figure 2.2.

2.2.7 Data Likelihood

As data, D, we consider here the positions and velocities of a sub-population of mock data stars within a given survey SF, ${\rm SF}(\boldsymbol{x})$,

$$\begin{aligned} D \equiv \{\boldsymbol{x}_i, \boldsymbol{v}_i \mid \quad & (\text{star } i \text{ in given sub-population}) \\ \wedge \quad & ({\rm SF}(\boldsymbol{x}_i) > 0)\}. \end{aligned}$$

For simplicity, we assume in most tests of this study contiguous, spherical SFs centered on the Sun, which are functions of $\boldsymbol{x}$ only and which we motivate in Section 2.2.5. The maximum radius of this spherical observed volume is denoted by $r_{\rm max}$.
We fit a model potential and DF (here, the qDF) which are specified by a number of fixed and free model parameters,

$$p_M \equiv \{p_{\rm DF}, p_\Phi\}.$$

The orbit of the ith star in a potential with p_Φ is labeled by the actions $\boldsymbol{J}_i \equiv \boldsymbol{J}[\boldsymbol{x}_i, \boldsymbol{v}_i \mid p_\Phi]$ and the DF evaluated for the ith star is then ${\rm DF}(\boldsymbol{J}_i \mid p_M) \equiv {\rm DF}(\boldsymbol{J}[\boldsymbol{x}_i, \boldsymbol{v}_i \mid p_\Phi] \mid p_{\rm DF})$.

[6] For the creation of the mock data we use $N_x = 20$, $N_v = 40$ and $n_\sigma = 5$ in Equation (2.1).

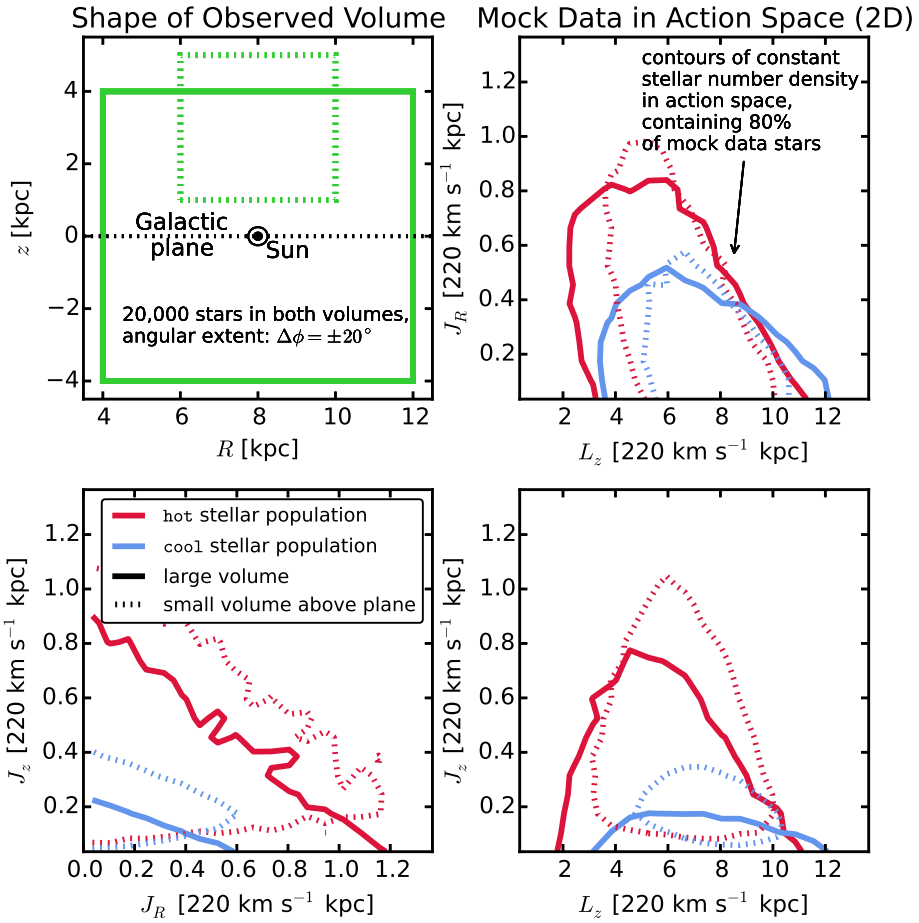


Figure 2.2: Distribution of mock data in action space (2D iso-density contours, enclosing 80% of the stars) depending on shape and position of a wedge-like survey observation volume (upper left panel, see also Section 2.2.5), and temperature of the stellar population (indicated in the legend). The four mock data sets are generated in the `KKS-Pot` from Table 2.1 from either the `hot` or `cool` DF in Table 2.2. The distribution in action space visualizes how orbits with different actions reach into different regions within the Galaxy. The corresponding mock data in configuration space is shown in Figure 2.3.

The likelihood of the data given the model is, following Bovy & Rix (2013) and McMillan & Binney (2013),

$$\begin{aligned}
&\mathscr{L}(D \mid p_M) \\
&\equiv \prod_i^{N_*} p(\boldsymbol{x}_i, \boldsymbol{v}_i \mid p_M) \\
&= \prod_i^{N_*} \frac{\mathrm{DF}(\boldsymbol{J}_i \mid p_M) \cdot \mathrm{SF}(\boldsymbol{x}_i)}{\int \mathrm{DF}(\boldsymbol{J} \mid p_M) \cdot \mathrm{SF}(\boldsymbol{x})\, \mathrm{d}^3x\, \mathrm{d}^3v} \\
&\propto \prod_i^{N_*} \frac{\mathrm{DF}(\boldsymbol{J}_i \mid p_M)}{\int \rho_{\mathrm{DF}}(R, |z| \mid p_M) \cdot \mathrm{SF}(\boldsymbol{x})\, \mathrm{d}^3x},
\end{aligned} \tag{2.3}$$

where N_* is the number of stars in D, and in the last step we used Equation (2.1).[7] $\prod_i^{N_*} \mathrm{SF}(\boldsymbol{x}_i)$

[7] Because $|\partial(\boldsymbol{J}, \boldsymbol{\theta})/\partial(\boldsymbol{x}, \boldsymbol{v})| = 1$, the integration over phase-space in the normalization term can be performed either over $(\boldsymbol{J}, \boldsymbol{\theta})$ or Cartesian $(\boldsymbol{x}, \boldsymbol{v})$.

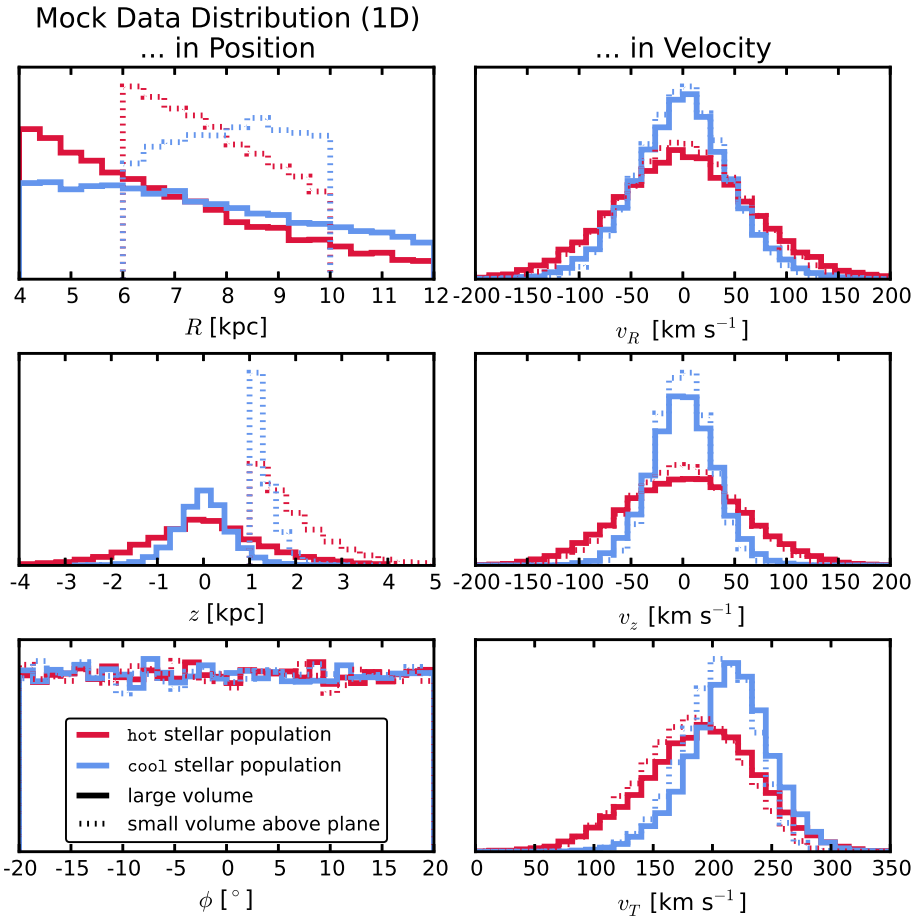


Figure 2.3: Distribution of the mock data from Figure 2.2 in configuration space. The corresponding observation volumes (as indicated in the legend) are shown in Figure 2.2, upper left panel. The 1D histograms illustrate that qDFs generate realistic stellar distributions in Galactocentric coordinates $(R, z, \phi, v_R, v_z, vT)$. 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 of the order of $\sim \sigma_{R,0}$ and $\sim \sigma_{z,0}$ (see Table 2.2). The distribution of tangential velocities, v_T, is skewed because of asymmetric drift.

is independent of p_M, so we treat it as unimportant proportionality factor. We find the best fitting p_M by maximizing the posterior Probability Distribution Function (PDF), $\mathrm{PDF}(p_M \mid D)$, which is, according to Bayes' theorem

$$\mathrm{PDF}(p_M \mid D) \propto \mathscr{L}(D \mid p_M) \cdot p(p_M),$$

where $p(p_M)$ is some prior probability distribution on the model parameters. We assume flat priors in both p_Φ and

$$p_{\mathrm{DF}} := \{\ln h_R, \ln \sigma_{R,0}, \ln \sigma_{z,0}, \ln h_{\sigma,R}, \ln h_{\sigma,z}\} \tag{2.4}$$

(see Section 2.2.4) throughout this chapter. Then PDF and likelihood are proportional to each other and differ only in units.

In this case, where we use uninformative priors, a maximum likelihood estimation procedure (e.g., via the expectation-maximization algorithm and parameter uncertainty estimates from the Fisher information matrix) would lead to the same result as the Bayesian inferential procedure

described in this chapter (see Section 2.2.10). We expect, however, that in due course increasingly informative priors will become available (like, e.g., rotation curve measurements from maser sources by Reid et al. 2009) and Bayesian inference is therefore the preferred framework.

2.2.8 Likelihood Normalization

The normalization in Equation (2.3) is a measure for the total number of tracers inside the survey volume,

$$M_{\mathrm{tot}} \equiv \int \rho_{\mathrm{DF}}(R, |z| \mid p_M) \cdot \mathrm{SF}(\boldsymbol{x})\, \mathrm{d}^3x. \tag{2.5}$$

In the case of an axisymmetric Galaxy model and $\mathrm{SF}(\boldsymbol{x}) = 1$ within the observation volume (as in most tests in this chapter), the normalization is essentially a two-dimensional integral in the (R, z) plane over ρ_{DF} with finite integration limits. We evaluate the integrals using Gauss-Legendre quadratures of order 40. The integral over the azimuthal direction can be solved analytically.
It turns out that a sufficiently accurate evaluation of the likelihood is computationally expensive, even for only one set of model parameters. This expense is dominated by the number of action calculations required, which in turn depends on N_* and the numerical accuracy of the tracer density interpolation grid with $N_x^2 \times N_v^3$ grid points in Equation (2.1) needed for the likelihood normalization in Equation (2.5). The accuracy of the normalization has to be chosen high enough, such that the resulting numerical error

$$\delta M_{\mathrm{tot}} \equiv \frac{M_{\mathrm{tot,approx}}(N_x, N_v, n_\sigma) - M_{\mathrm{tot}}}{M_{\mathrm{tot}}} \tag{2.6}$$

does not dominate the numerically calculated log-likelihood, i.e.,

$$\begin{aligned} &\ln \mathscr{L}_{\mathrm{approx}}(D \mid p_M) \\ &= \sum_i^{N_*} \ln \mathrm{DF}(\boldsymbol{J_i} \mid p_M) - N_* \ln(M_{\mathrm{tot}}) \\ &-N_* \ln(1 + \delta M_{\mathrm{tot}}), \end{aligned} \tag{2.7}$$

with

$$\ln(1 + \delta M_{\mathrm{tot}}) \leq \frac{1}{N_*}, \tag{2.8}$$

and therefore $\delta M_{\mathrm{tot}} \lesssim 1/N_*$. Otherwise, numerical inaccuracies could lead to systematic biases in the potential and DF recovery. For data sets as large as $N_* = 20,000$ stars, which in the age of *Gaia* could very well be the case, one needs a numerical accuracy of 0.005% in the normalization. We made sure that this is satisfied for all analyses in this chapter. Figure 2.4 demonstrates how the numerical accuracy for analyses with the `DHB-Pot` depends on the spatial and velocity resolution of the grid and that the accuracy we use, $N_x = 16$, $N_v = 24$ and $n_\sigma = 5$, is sufficient.[8] It has to be noted, however, that the optimal values for N_x, N_v, and n_σ depend not only on N_*, but also on the kinematic temperature of the population (and to a certain degree even on the choice of potential[9]) and it has to be checked on a case-by-case basis what the optimal accuracy is.

[8] The accuracy used in this chapter's analyses is slightly higher than in Bovy & Rix (2013), where N_* was only a few ~ 100.

[9] In Figure 2.19 we will show a comparison for the qDF parameters of two very similar mock data distributions in two different potentials, the `MW14-Pot` and a best fit potential of the parametric form of the `KKS-Pot`. As some of the qDF parameters in both potentials are very different, and even more different from the actual physical scale

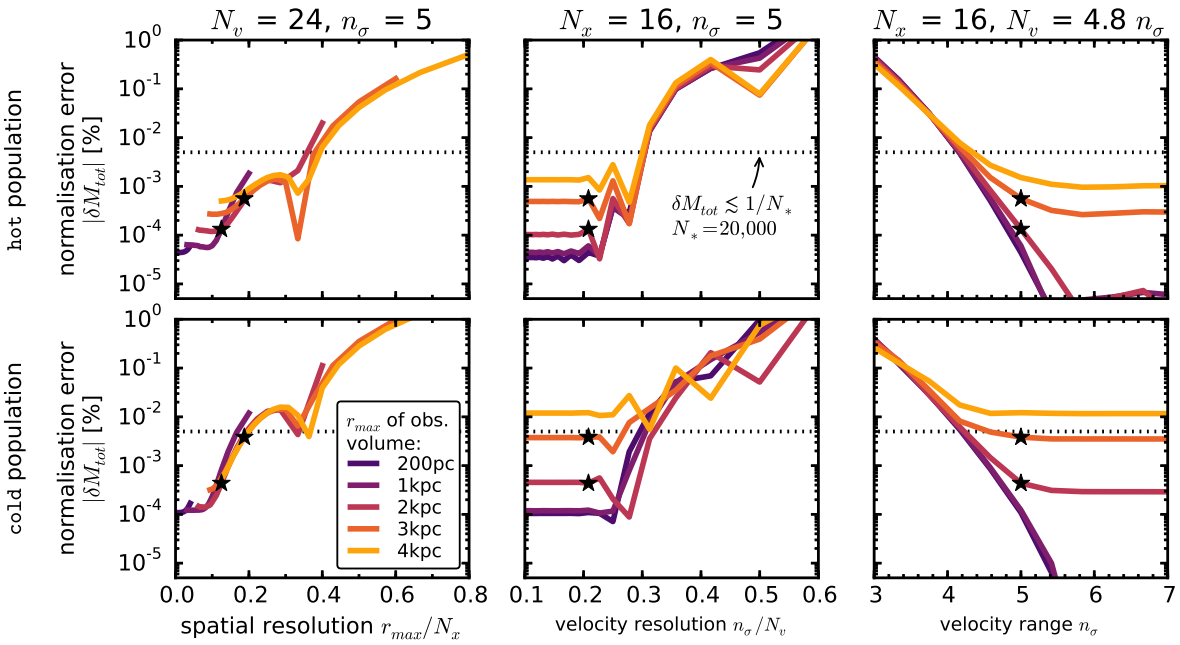


Figure 2.4: Relative error of the likelihood normalization, $\delta M_{\rm tot}$, in Equation (2.6), depending on the accuracy of the grid-based density calculation in Equation (2.1) (and surrounding text) in five spherical observation volumes with different radius, $r_{\rm max}$, and depending on the kinematic temperature of the population in the `DHB-Pot`. (Test 1 in Table 2.3 summarizes the model parameters.) The tracer density in Equation (2.1) is calculated on $N_x \times N_x$ spatial grid points in $R \in [R_\odot \pm r_{\rm max}]$ and $|z| \in [0, r_{\rm max}]$. The integration over the velocities is performed with Gauss-Legendre quadratures of order N_v within an integration range of $\pm n_\sigma$ times the dispersion $\sigma_R(R)$ and $\sigma_z(R)$ (and $[0, 1.5 v_{\rm circ}]$ in v_T). (We vary N_x, N_v and n_σ separately, and keep the other two fixed at the values indicated above each panel.) We calculate the "true" normalization $M_{\rm tot}$ in Equation (2.6) with high accuracy as $M_{\rm tot} \equiv M_{\rm tot,approx}(N_x = 32, N_v = 68, n_\sigma = 7)$. The black stars indicate the accuracy used in analyses with the `DHB-Pot`, Tests 5 and 7. It is better than 0.005% (dotted line), which is required for $N_* = 20,000$ stars. We find that the spatial resolution of the grid is important and depends on the kinematic temperature of the population, as cooler populations have a steeper density gradient in the z-direction, which has to be sampled sufficiently.

McMillan & Binney (2013), who use a similar modeling approach and likelihood normalization, argued that the required accuracy for the normalization scales as $\log_{10}(1 + \delta M_{\rm tot}) \leq 1/N_* \Rightarrow \delta M_{\rm tot} \lesssim 2.3/N_*$, which is satisfied for our tests as well. They evaluate the integrals in the normalization via Monte-Carlo integration with $\sim 10^9$ sample points in action space. Our approach uses a tracer density interpolation grid for which the resolution needs to be optimized by hand, but it has the advantage that it then only requires the calculation of $N_x^2 \times N_v^3 \sim 4{\cdot}10^6 - 10^7$ actions per normalization.

2.2.9 Measurement Uncertainties

Measurement uncertainties of the data have to be incorporated in the likelihood. We assume Gaussian uncertainties in the observable space $\boldsymbol{y} \equiv (\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{v}}) = (\text{R.A.}, \text{decl.}, (m - M), \mu_{\text{R.A.}} \cdot \cos(\text{decl.}), \mu_{\text{decl.}}, v_{\rm los})$, i.e., the ith star's observed $\boldsymbol{y}_i$ is drawn from the normal distribution $N[\boldsymbol{y}_i', \delta\boldsymbol{y}_i] \equiv \prod_i^6 N[y_{i,k}', \delta y_{i,k}] = \prod_i^6 \exp\{-(y_k - y_{i,k}')^2/(2\delta y_{i,k}^2)\}/\sqrt{2\pi\delta y_{i,k}^2}$, with $\boldsymbol{y}_i'$ being the star's true phase-space position, $\delta\boldsymbol{y}_i$ its uncertainty, and y_k the kth coordinate component of

lengths and velocity dispersions, an optimal n_σ has to be estimated first for a given potential model before running the RoadMapping analysis.

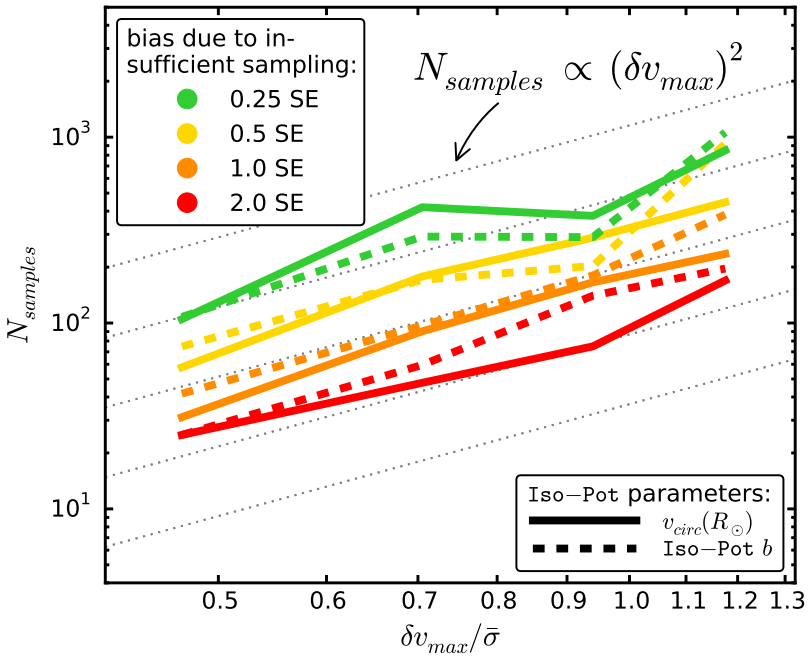


Figure 2.5: Number of MC samples, $N_{\rm samples}$, needed for the numerical convolution of the model probability with the measurement uncertainties in Equation (2.10), given the maximum velocity uncertainty, $\delta v_{\rm max}$, within the stellar sample with respect to the sample's kinematic temperature, $\bar{\sigma}$. Insufficient sampling introduces systematic biases in the parameter recovery; the size of the bias (in units of the standard error (SE) on the parameter estimate) is indicated in the legend. The relation found here, $N_{\rm samples} \propto \delta v_{\rm max}^2$, was distilled from analyses (with different $N_{\rm samples}$) of mock data sets with different proper motion uncertainties $\delta\mu \in [2, 5]$ mas yr^{-1} (see Test 2 in Table 2.3). As the reference for the converged convolution integral, we used $N_{\rm samples} = 800$ and 1200 for $\delta\mu \leq 3$ mas yr^{-1} and $\delta\mu > 3$ mas yr^{-1}, respectively (see also left panels in Figure 2.11). We plot $\delta v_{\rm max}$ in units of the sample temperature, which we quantify by $\bar{\sigma} \equiv (\sigma_{R,0} + \sigma_{z,0})/2$ (see Table 2.2 for the `hot` qDF).

$\boldsymbol{y}$. Stars follow the DF($\boldsymbol{J}[\boldsymbol{y}' \mid p_\Phi] \mid p_{\rm DF}$) ($\equiv$ DF($\boldsymbol{y}'$) for short) convolved with the measurement uncertainties $N[0, \delta\boldsymbol{y}_i]$. The SF($\boldsymbol{y}$) acts on the space of (uncertainty affected) observables. Then the probability of one star becomes

$$\begin{aligned} &\tilde{p}(\boldsymbol{y}_i \mid p_\Phi, p_{\rm DF}, \delta\boldsymbol{y}_i) \\ \equiv\ &\frac{{\rm SF}(\boldsymbol{y}_i) \cdot \int {\rm DF}(\boldsymbol{y}') \cdot N[\boldsymbol{y}_i, \delta\boldsymbol{y}_i]\, {\rm d}^6 y'}{\int ({\rm DF}(\boldsymbol{y}') \cdot \int {\rm SF}(\boldsymbol{y}) \cdot N[\boldsymbol{y}', \delta\boldsymbol{y}_i]\, {\rm d}^6 y)\, {\rm d}^6 y'}. \end{aligned} \tag{2.9}$$

In the case of uncertainties in distance and/or (R.A., decl.) the evaluation of this is very expensive computationally, especially if the stars have heteroscedastic $\delta\boldsymbol{y}_i$, which is the case for realistic data sets, and the normalization needs to be calculated for each star separately. In practice, we compute the convolution using MC integration with $N_{\rm samples}$ samples,

$$\begin{aligned} &\tilde{p}_{\rm approx}(\boldsymbol{y}_i \mid p_\Phi, p_{\rm DF}, \delta\boldsymbol{y}_i) \\ \approx\ &\frac{{\rm SF}(\tilde{\boldsymbol{x}}_i)}{M_{\rm tot}} \cdot \frac{1}{N_{\rm samples}} \sum_n^{N_{\rm samples}} {\rm DF}(\tilde{\boldsymbol{x}}_i, \boldsymbol{v}[\boldsymbol{y}'_{i,n}]) \end{aligned} \tag{2.10}$$

with

$$\boldsymbol{y}'_{i,n} \sim N[\boldsymbol{y}_i, \delta\boldsymbol{y}_i].$$

In addition, this approximation assumes that the star's position, $\tilde{\boldsymbol{x}}_i$, is perfectly measured. As the SF is also velocity independent, this simplifies the normalization drastically to $M_{\rm tot}$ in Equation (2.5). Measurement uncertainties in R.A. and decl. are often negligible anyway. The uncertainties in the Galactocentric velocities $\boldsymbol{v}_i = (v_{R,i}, v_{T,i}, v_{z,i})$ depend not only on $\delta\boldsymbol{\mu}$ and $\delta v_{\rm los}$ but also on the distance and its uncertainty, which we do *not* neglect when drawing MC samples, $\boldsymbol{y}'_{i,n}$, from the full uncertainty distribution, $N[\boldsymbol{y}_i, \delta\boldsymbol{y}_i]$.
An analogous but one-dimensional treatment of measurement uncertainties in only v_z was already applied by Bovy & Rix (2013). Similar approaches ignoring measurement uncertainties in the likelihood normalization and using MC sampling of the error ellipses were also used by McMillan & Binney (2013) and Das & Binney (2016). In Section 2.3.4, Figure 2.11 (Test 6.2 in Table 2.3), we will investigate the breakdown of our approximation for non-negligible distance uncertainties. Figure 2.5 demonstrates that in the absence of position uncertainties, the $N_{\rm samples}$ needed for the convolution integral to converge depends as

$$N_{\rm samples} \propto (\delta v)^2$$

on the uncertainties in the (1D) velocities. Figure 2.5 is based on analyses of mock data sets with different proper motion uncertainties, $\delta\mu$ (see Test 2 in Table 2.3 for all parameters). The proper motion uncertainty, $\delta\mu$, translates to heteroscedastic velocity uncertainties according to

$$\delta v[\mathrm{km\ s^{-1}}] \equiv 4.74047 \cdot r[\mathrm{kpc}] \cdot \delta\mu[\mathrm{mas\ yr^{-1}}],$$

with r being the distance of the star to the Sun. Stars with larger δv require more $N_{\rm samples}$ for the integral over their measurement uncertainties to converge; Figure 2.5 therefore shows how the $N_{\rm samples}$—needed for the PDF of the *whole* data set to be converged—depends on the *largest* velocity error, $\delta v_{\rm max} \equiv \delta v(r_{\rm max})$, within the data set.
These mock data sets contained each $N_* = 10,000$ stars. We found that for $N_* = 5,000$ the required $N_{\rm samples}$ to reach a given accuracy becomes smaller for $v_{\rm circ}(R_\odot)$, but remains similar for b. The former is consistent with our expectation that we need higher accuracy and therefore more $N_{\rm samples}$ for larger data sets. The latter seems to be a special property of the `Iso-Pot` (see also the discussion in Section 2.3.3).

2.2.10 Fitting Procedure

To search the $(p_\Phi, p_{\rm DF})$ parameter space for the maximum of the PDF in Equation (2.3), we go beyond the single fixed grid search by Bovy & Rix (2013) and employ an efficient two-step procedure: Nested-grid search and Monte Carlo Markov Chain (MCMC).
The first step employs a nested-grid search to find the approximate peak and width of the PDF in the high-dimensional p_M space with a low number of likelihood evaluations:

- *Initialization.* For N_p free model parameters, p_M, we start with a sufficiently large grid with 3^{N_p} regular points.
- *Evaluation.* We evaluate the PDF at each grid-point similar to Bovy & Rix (2013) (their Figure 9). An outer loop iterates over the potential parameters, p_Φ, and pre-calculates all $N_* \times N_{\rm samples} + N_x^2 \times N_v^3$ actions required for the likelihood calculation (see Equations (2.1), (2.3) and (2.10)). Then an inner loop evaluates Equation (2.3) (or (2.10)) for all DF parameters, $p_{\rm DF}$, in the given potential.
- *Iteration.* For each of the model parameters, p_M, we marginalize the PDF. A Gaussian is fitted to the marginalized PDF and the peak $\pm\ 4\sigma$ become the boundaries of the next grid with 3^{N_p} grid points. The grid might still be too coarse or badly positioned to fit Gaussians. In that case, we either zoom into the grid point with the highest probability

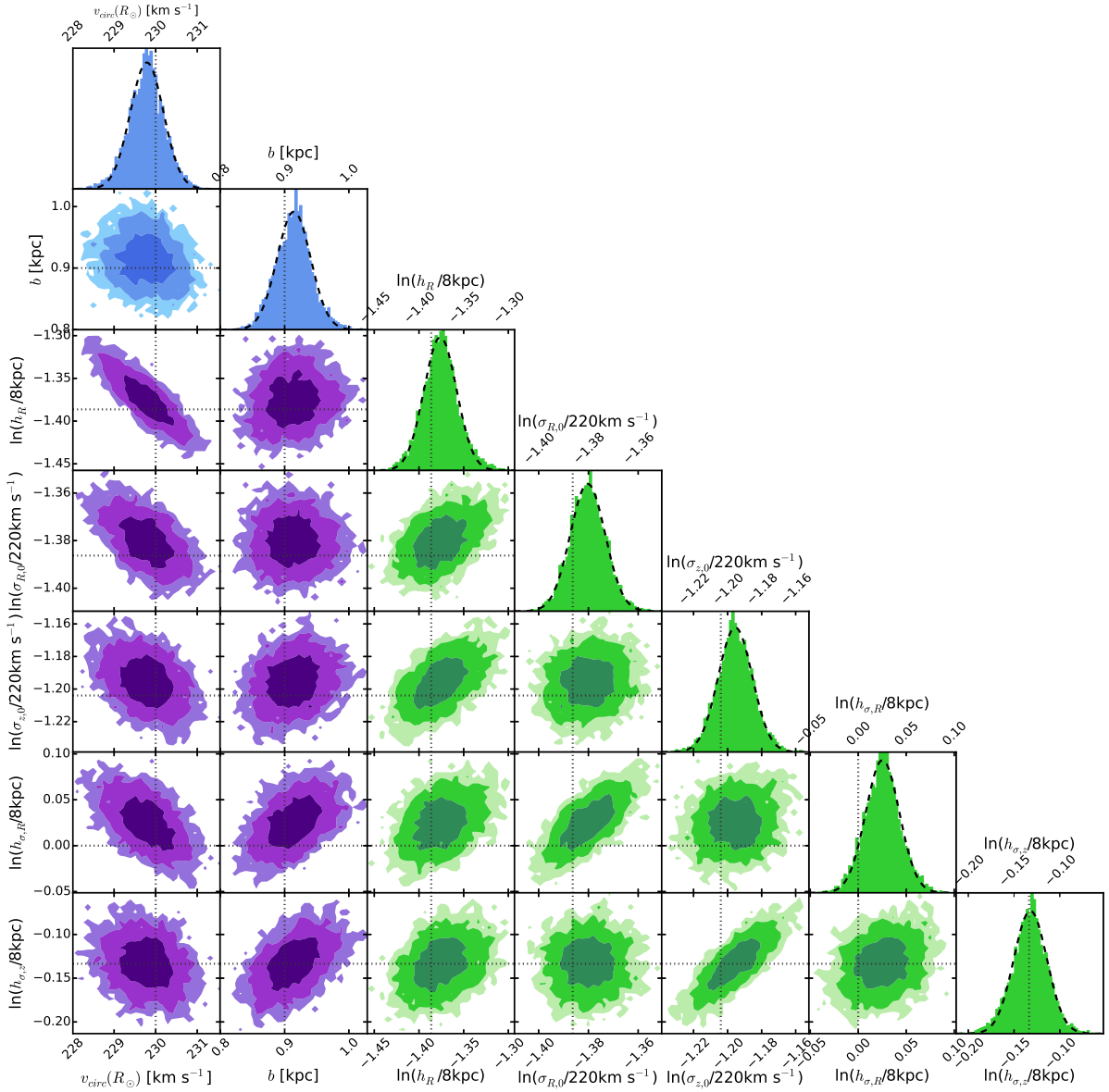


Figure 2.6: The PDF in the parameter space $p_M = \{p_\Phi, p_{\rm DF}\}$ for one example mock data set (see Test 3.1 in Table 2.3). Blue indicates the PDF for the potential parameters, p_Φ, and green indicates the qDF parameters, $p_{\rm DF}$. The true parameters are marked by dotted lines. The dark, medium, and bright contours in the 2D distributions represent 1σ, 2σ and 3σ confidence regions, respectively. The parameters are weakly to moderately covariant, but their level of covariance depends on the actual choice of the mock data's p_M. The PDF here was sampled using MCMC. The dashed lines in the 1D distributions are Gaussian fits to the histogram of MCMC samples. This demonstrates that, for such a large number of stars, the PDF approaches the shape of a multivariate Gaussian that also projects into Gaussians when considering the marginalized PDF for all the individual p_M, as expected for a maximum likelihood estimator.

or shift the current range to find new approximate grid boundaries. We proceed with iteratively evaluating the PDF on finer and finer grids, until we have found a reliable 4σ fit range in each of the p_M dimensions. The central grid point is then very close to the best fit p_M, and the grid range is of the order of the PDF width.

- *The fiducial qDF.* To save time by pre-calculating actions, they have to be independent of the choice of $p_{\rm DF}$. However, the normalization in Equation (2.5) requires actions on a $N_x^2 \times N_v^3$ grid and the grid ranges in velocity space *do* depend on the current $p_{\rm DF}$ (see Equation (2.1)). To relax this, we follow Bovy & Rix (2013) and use a fixed set of qDF parameters (the *fiducial qDF*) to set the velocity grid boundaries in Equation (2.1) globally for a given p_Φ. Choosing a fiducial qDF that is very different from the true DF can, however, lead to large biases in the p_M recovery. Bovy & Rix (2013) did not account for that. RoadMapping avoids this as follows: to successively get closer to the optimal fiducial qDF—with the (yet unknown) best fit $p_{\rm DF}$—we use in each iteration step of the nested-grid search the central grid point of the current p_M grid as the fiducial qDF's $p_{\rm DF}$. As the nested-grid search approaches the best fit values, the fiducial qDF approaches its optimum as well.

- *Computational expense.* Overall, the computation speed of this nested-grid approach is dominated (in descending order of importance) by a) the complexity of potential and action calculation, b) the $N_* \times N_{\rm samples} + N_x^2 \times N_v^3$ actions required to be calculated per p_Φ, c) the number of potential parameters, and d) the number of DF parameters.

The second step samples the shape of the PDF using MCMC. Formally, calculating the PDF on a fine grid like Bovy & Rix (2013) (e.g., with $K = 11$ grid points in each dimension) would provide the same information. However, the number of expensive PDF evaluations scales as K^{N_p}. For a high-dimensional p_M ($N_p > 4$), an MCMC approach might sample the PDF much faster: we use *emcee* by Foreman-Mackey et al. (2013) and release the walkers very close to the best fit p_M found by the nested-grid search, which assures fast convergence in much less than K^{N_p} PDF evaluations. We also use the best fit p_M of the grid search as fiducial qDF for the whole MCMC. In doing so, the normalization varies smoothly with different p_M and is less sensitive to the accuracy in Equation (2.1).

2.3 Results

We are now in a position to examine the limitations of action-based modeling posed in the introduction using our RoadMapping machinery. We explore: (i) whether the parameter estimates are unbiased, (ii) the role of the survey volume, (iii) imperfect SFs, (iv) measurement uncertainties, and what happens if the true (v) DF or (vi) potential are not included in the space of models. We will rely on mock data as input to explore the limitations of the modeling. The mock data is generated directly from the fiducial potential and DF models introduced in Sections 2.2.3 and 2.2.4, following the procedure described in Section 2.2.6. With the exception of the test suite on measurement uncertainties in Section 2.3.4, we assume that phase-space uncertainties are negligible. All tests are also summarized in Table 2.3.
We do not explore the breakdown of the assumption that the system is axisymmetric and in steady state nor the impact of resonances, which is not possible in the current setup using mock data drawn from axisymmetric galaxy models. We plan, however, to investigate this in a future paper, where we will apply RoadMapping to N-body simulations of disk galaxies.

Table 2.3. Summary of Test Suites in this Chapter.

Test		Model for Mock Data	Model in Analysis	Figures and Results
Test 1 : Numerical accuracy in calculating the likelihood normalization	*Potential:*	`DHB-Pot`	-	Figure 2.4 Suitable accuracy for our tests: $N_x = 16, N_v = 24, n_\sigma = 5$. Higher spatial resolution is required for cooler populations.
	DF:	`hot` or `cool` qDF		
	Survey volume:	sphere around Sun, $r_{\rm max} = 0.2, 1, 2, 3$, or 4 kpc		
	Numerical accuracy:	$N_x \in [5, 32]$, $N_v \in [4, 48]$, $n_\sigma \in [3, 7]$		
Test 2 : Numerical convergence of convolution with measurement uncertainties	*Potential:*	`Iso-Pot`	`Iso-Pot`, all parameters free	Figure 2.5 The number of required MC samples scales as $N_{\rm samples} \propto (\delta v_{\rm max})^2$, with $\delta v_{\rm max}$ being the largest δv in the data set. If N_* is smaller, less accuracy, i.e., less $N_{\rm samples}$, is needed to reach a given accuracy.
	DF:	`hot` qDF	qDF, all parameters free	
	Survey Volume:	sphere around Sun, $r_{\rm max} = 3$ kpc	(fixed and known)	
	Uncertainties:	δR.A. $= \delta$decl. $= \delta(m - M) = 0$	(fixed and known)	
		$\delta v_{\rm los} = 2$ km s^{-1}		
		$\delta\mu = 2$,3,4, or 5 mas yr^{-1}		
	Numerical accuracy:		$N_{\rm samples} \in [25, 1200]$	
	N_**:*	10,000		
Test 3.1 : Shape of the model parameters' PDF for large data sets	*Potential:*	`Iso-Pot`	`Iso-Pot`, all parameters free	Figure 2.6 The PDF becomes a multivariate Gaussian in the limit of large data. The width of the PDF scales as $1/\sqrt{N_*}$.
	DF:	`hot` qDF	qDF, all parameters free	
	Survey Volume:	sphere around Sun, $r_{\rm max} = 2$ kpc	(fixed and known)	
	N_**:*	20,000		
Test 3.2 : Parameter estimates are unbiased; Influence of survey volume size	*Potential:*	`Iso-Pot`	`Iso-Pot`, free parameter: b	Figure 2.7 RoadMapping behaves like an unbiased maximum likelihood estimator. Larger survey volumes lead to tighter constraints, even at the same N_*.
	DF:	`hot` or `cool` qDF	qDF, free parameters: $\ln h_R, \ln \sigma_{z,0}, \ln h_{\sigma,z}$	
	Survey volume:	sphere around Sun, $r_{\rm max} = 0.2, 1, 2, 3$, or 4 kpc	(fixed and known)	
	N_**:*	20,000		
Test 4 : Influence of position and shape of survey volume on parameter recovery	*Potential:*	(i) `Iso-Pot` or (ii) `DHB-Pot`	(i) `Iso-Pot`, all parameters free (ii) `DHB-Pot`, free parameters: $v_{\rm circ}(R_\odot)$, $a_{\rm disk}$, $f_{\rm halo}$	Figure 2.8 The exact position and shape of the survey volume plays only a minor role. Having both large radial and vertical extent should give the tightest constraints.
	DF:	`hot` qDF	qDF, all parameters free	
	Survey volume:	four different wedges, see Figure 2.8, upper panel	(fixed and known)	
	N_**:*	20,000		
Test 5 : Influence of wrong assumptions about the spatial SF on parameter recovery	*Potential:*	`DHB-Pot`	`DHB-Pot`, free parameters: $v_{\rm circ}(R_\odot), a_{\rm disk}, f_{\rm halo}$	Figures 2.9 For minor misjudgements of a radially symmetric SF (i.e., $\epsilon_r \lesssim 0.15$ for `hot` and $\epsilon_r \lesssim 0.2$ for `cool` populations) the potential recovery is still robust.
	DF:	`hot` or `cool` qDF	qDF, all parameters free	
	Survey volume:	sphere around Sun, $r_{\rm max} = 3$ kpc	(fixed and known)	
	Completeness:	Equation (2.11) with $\epsilon_r \in [0, 0.7]$	completeness$(\boldsymbol{x}) = 1$, i.e., $\epsilon_r = 0$	
	N_**:*	20,000		

Table 2.3 (cont'd)

Test		Model for Mock Data	Model in Analysis	Figures and Results
Test 6.1 :	*Potential:*	`Iso-Pot`	`Iso-Pot`, all parameters free	Figure 2.10
Effect of proper motion	*DF:*	`hot` or `cool` qDF	qDF, all parameters free	If $\delta\mu$ is perfectly known, the
uncertainties on	*Survey volume:*	sphere around Sun, $r_{\rm max} = 3$ kpc	(fixed and known)	precision of the potential para-
precision of	*Uncertainties:*	(i) δR.A. = δdecl. = $\delta(m-M) = 0$,	(fixed and known)	meter recovery is only a
potential recovery		$\delta v_{\rm los} = 2$ km s^{-1},		factor $\sim 1.15 - 2$ worse
		$\delta\mu = 1, 2, 3, 4,$ or 5 mas yr^{-1}		for $\delta\mu \sim 1 - 5$ mas yr^{-1}
		(ii) no measurement uncertainties		than for a data set without
	N_*:	10,000		proper motion uncertainties.
Test 6.2 :	*Potential:*	`Iso-Pot`	`Iso-Pot`, all parameters free	Figure 2.11
Testing the convolution	*DF:*	`hot` qDF	qDF, all parameters free	The approximate likelihood
with measurement	*Survey Volume:*	sphere around Sun, $r_{\rm max} = 3$ kpc	(fixed and known)	in Equation (2.10) is the true
uncertainties in Equation (2.10)	*Uncertainties:*	δR.A. = δdecl. = 0,	(fixed and known)	likelihood in the absence of
with and without		$\delta v_{\rm los} = 2$ km s^{-1},		position errors. For distance
distance uncertainties		$\delta\mu$ = 1,2,3,4, or 5 mas yr^{-1},		uncertainties $\lesssim 10\%$ at
		(i) $\delta(m-M) = 0$ or		$r_{\rm max} \sim 3$ kpc, which is the case
		(ii) $\delta(m-M) \neq 0$ (see Figure 2.11)		for *Gaia* and $G \lesssim 15$, the intro-
	N_*:	10,000		duced bias is still less than 2σ.
Test 6.3 :	*Potential:*	`Iso-Pot`	`Iso-Pot`, all parameters free	Figure 2.12
Underestimation	*DF:*	`hot` or `cool` qDF	qDF, all parameters free	The bias introduced by under-
of proper motion	*Survey volume:*	sphere around Sun, $r_{\rm max} = 3$ kpc	(fixed and known)	estimating $\delta\mu$ grows with $\delta\mu$.
uncertainties	*Uncertainties:*	only proper motion uncertainties	proper motion uncertainties	If $\delta\mu \lesssim 3$ mas yr^{-1} and $\delta\mu$
		1, 2, or 3 mas yr^{-1}	10% or 50% underestimated	is underestimated by 10%
	N_*:	10,000		the bias is only $\sim 2\sigma$ or less.
Test 7 :	*Potential:*	`DHB-Pot`	`DHB-Pot`, free parameters:	Figures 2.13, 2.14, 2.15, and 2.16
Deviations of the			$v_{\rm circ}(R_\odot), a_{\rm disk}, f_{\rm halo}$	*Example 1:*
assumed DF	*DF:*	mix of two qDFs...	single qDF, all parameters free	Hot qDF-like populations polluted
from the		(i) *Example 1:*		by up to $\sim 30\%$ of stars from a much
stars' true DF		... with different mixing rates		cooler population give still reliable
		and fixed qDF parameters		potential constraints. This is not
		(`hot` and `cooler` qDF from Table 2.2)		true for the opposite case.
		(ii) *Example 2:*		*Example 2:*
		... with 50/50 mixing rate		Differences of $\lesssim 20\%$ in qDF
		and varying qDF parameters (by $X\%$):		parameters do not matter
		a) `hot` and `colder` qDF or		when mixing sub-populations,
		b) `cool` and `warmer` qDF (see Table 2.2)		e.g., due to finite binning in
	Survey volume:	sphere around Sun, $r_{\rm max} = 2$ kpc	(fixed and known)	abundance space and abundance
	N_*:	20,000		errors.

Table 2.3 (cont'd)

Test		Model for Mock Data	Model in Analysis	Figures and Results
Test 8 : Deviations of the assumed potential model from the stars' true potential	*Potential:*	`MW14-Pot`	`KKS-Pot`, all parameters free, only $v_{\rm circ}(R_\odot) = 230$ km s^{-1} fixed	Figures 2.17, 2.18, and 2.19 We find a good approximation for the true potential, especially where most stars are located, given the limitations of the wrong assumed potential model family.
	DF:	`hot` or `cool` qDF	qDF, all parameters free	
	Survey volume:	sphere around Sun, $r_{\rm max} = 4$ kpc	(fixed and known)	
	N_**:*	20,000		
	Numerical accuracy:		$N_x = 20, N_v = 28, n_\sigma = 5.5$	

Note. — Reference potentials and qDFs are introduced in Tables 2.1 and 2.2, respectively. Parameters that are not left free in the analysis, are always fixed to their true value. Unless stated otherwise, all mock data sets have SFs with completeness($\boldsymbol{x}$) = 1 and no measurement uncertainties, and we use $N_x = 16, N_v = 24, n_\sigma = 5$ as numerical accuracy for calculating the likelihood normalization (see Section 2.2.8).
The first column indicates the test suite; the second column the potential, DF, and SF model, etc., used for the mock data creation; the third column the corresponding model assumed in the RoadMapping analysis; and the last column lists the figures belonging to the test suite and summarizes the results.

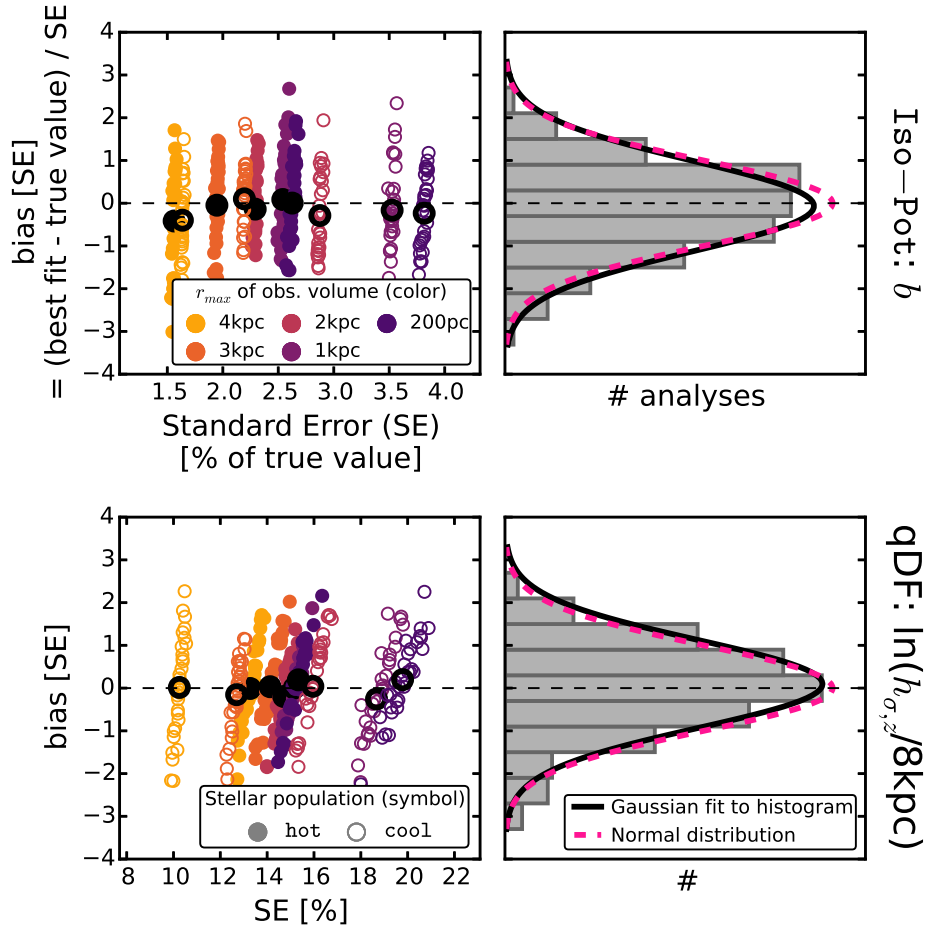


Figure 2.7: Lack of bias in the parameter estimates. Maximum likelihood estimators converge to the true parameter values for large numbers of data points and have a Gaussian spread—if the model assumptions are fulfilled. To test that these conditions are satisfied for RoadMapping, we create 320 mock data sets, which come from two different stellar populations and five spherical observation volumes (see legends). (All model parameters are summarized in Table 2.3 as Test 3.2.) Bias and relative standard error (SE) are derived from the marginalized PDF for two model parameters (isochrone scale length b in the first row, and qDF parameter $h_{\sigma,z}$ in the second row). The second column displays a histogram of the 320 bias offsets. As it closely follows a normal distribution, our modeling method is therefore well-behaved and unbiased. The black dots denote the PDF expectation value for the 32 analyses belonging to the same p_M.

2.3.1 Model Parameter Estimates in the Limit of Large Data Sets

The MAPs in Bovy & Rix (2013) contained between 100 and 800 objects, which implied broad PDFs for the model parameters, p_M. Several consequences arise in the limit of much larger samples, say $N_* = 20,000$: (i) As outlined in Section 2.2.8 and investigated in Figure 2.4 (Test 1 in Table 2.3), higher numerical accuracy is needed due to the likelihood normalization requirement, $\delta M_{\rm tot} \lesssim 1/N_*$ (see Equation (2.8)), which drives the computing time. (ii) The PDFs of the p_M become Gaussian, with a PDF width (i.e., the standard error (SE) on the parameter estimate) that scales as $1/\sqrt{N_*}$. The former is demonstrated in Figure 2.6 (Test 3.1 in Table 2.3) and we also verified that the latter is true. (iii) Any bias in the PDF expectation value has to be considerably less than the SE. Figure 2.7 (Test 3.2 in Table 2.3) illustrates that RoadMapping behaves like an unbiased maximum likelihood estimator; the average parameter estimates from many mock data sets are very close to the input p_M, and the distribution of the actual parameter estimates are Gaussian around it.

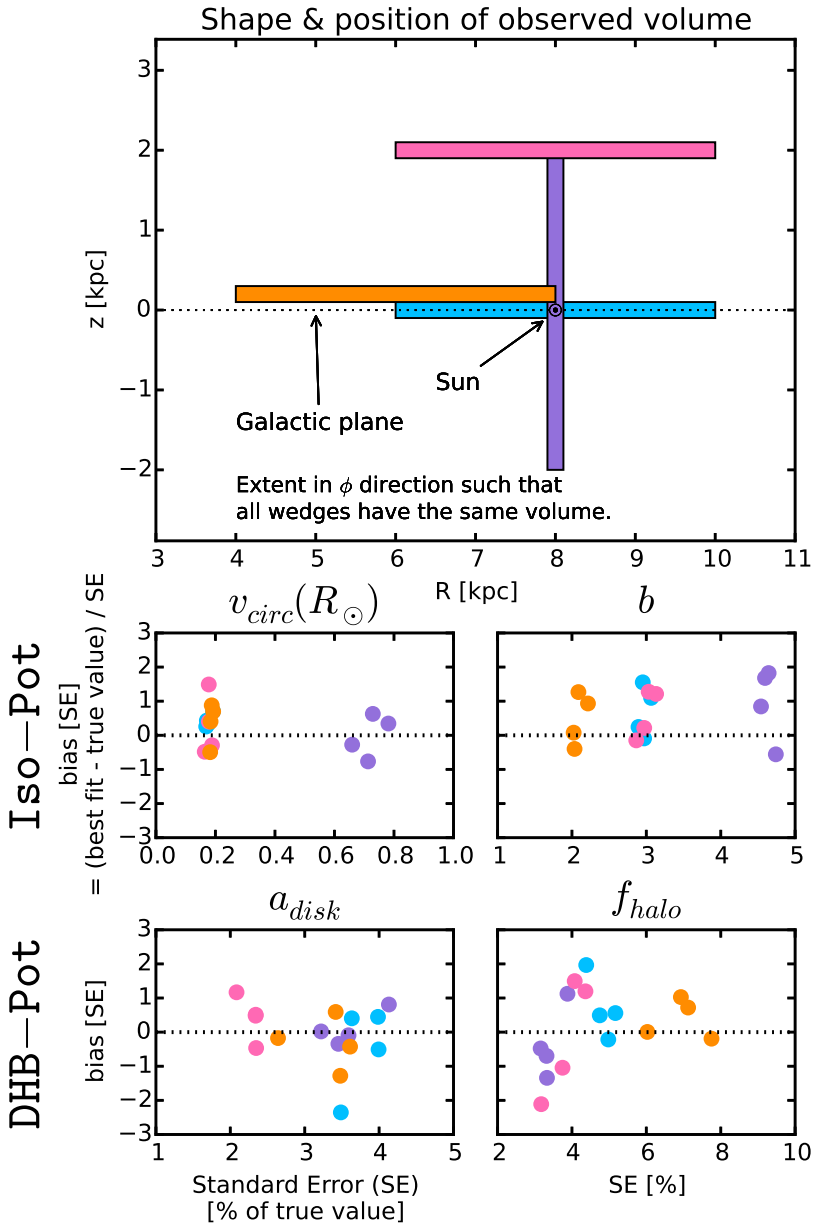


Figure 2.8: Bias vs. standard error in recovering the potential parameters for mock data sets drawn from four different wedge-shaped test observation volumes within the Galaxy (illustrated in the upper panel; the corresponding analyses are color-coded) and two different potentials (`Iso-Pot` and `DHB-Pot` from Table 2.1; see also Test 4 in Table 2.3 for all model parameters used). Standard error and offset were determined from a Gaussian fit to the marginalized PDF. The angular extent of each wedge-shaped observation volume was adapted such that all have a volume of 4.5 kpc^3, even though their extent in (R, z) is different. (The recovery of the free potential parameter $v_{\rm circ}(R_\odot)$ in the different wedges is very similar for both potentials and therefore only shown for the `Iso-Pot`). Minor expected differences can be seen (e.g., $v_{\rm circ}(R_\odot)$ and tracer density scale lengths requiring larger radial extent), but overall there is no clear trend that an observation volume around the Sun, above the disk, or at smaller Galactocentric radii should give remarkably better constraints on the potential than the other volumes.

2.3.2 The Role of the Survey Volume Geometry

To explore the role of the survey volume at a given sample size, we devise two suites of mock data sets.
The first suite draws mock data for two different potentials (`Iso-Pot` and `DHB-Pot`) and four volume wedges (see Section 2.2.5) with different extent and at different positions within the Galaxy, illustrated in the upper panel of Figure 2.8. Otherwise the data sets are generated from the same p_M (see Test 4 in Table 2.3). To isolate the role of the survey volume geometry and position, the mock data sets all have the same number of stars ($N_* = 20,000$), and are drawn from identical total survey volumes (4.5 kpc^3, achieved by adjusting the angular width of the wedges). We investigate these rather unrealistic survey volumes to test (i) if there are regions in the Galaxy where stars are on intrinsically more informative orbits and (ii) if spatial cuts applied to the survey volume (e.g., to avoid regions of large dust extinction or measurement uncertainties) would therefore strongly affect the precision of the potential constraints. To make this effect—if it exists—noticeable, we choose some extreme but illustrative examples.
The results are shown in Figure 2.8. The wedges all have the same volume and all give results of similar precision. There are some minor and expected differences, e.g., $v_{\rm circ}(R_\odot)$ and radial scale lengths (b and $a_{\rm disk}$) are slightly better recovered for a large radial extent and the halo fraction at the Sun, $f_{\rm halo}$, for volumes centered around $R_\odot$. In the case of an axisymmetric model galaxy, the extent in ϕ direction is not expected to matter. Overall radial extent and vertical extent seem to be equally important to constrain the potential. Figure 2.8 implies, therefore, that volume offsets or spatial cuts of the survey volume in the radial or vertical direction have at most a modest impact—even in case of the very large sample size at hand.
The second suite of mock data sets was already introduced in Section 2.3.1 (see also Test 3.2 in Table 2.3), where mock data sets were drawn from five spherical volumes around the Sun with different $r_{\rm max}$, for two different stellar populations. The results of this second suite are shown in Figure 2.7 and exemplify the effect of the size of the survey volume.
Figure 2.7 demonstrates that, given a choice of $p_{\rm DF}$, a larger volume always results in tighter constraints. There is no obvious trend that a hotter or cooler population will always give better results; it depends on the survey volume and the model parameter in question.
While it appears that the argument for significant radial and vertical extent is generic, we have not done a full exploration of all combinations of p_M and volumes.
In reality, different regions in the Galaxy have different stellar number densities and different measurement uncertainties, which should therefore be the major factor to drive the precision of the potential recovery when choosing a survey volume.

2.3.3 Impact of Misjudging the Selection Function of the Data Set

The survey SF (see also Section 2.2.5) is sometimes not perfectly determined. While the pattern of the survey area on the sky may be complex, it is usually precisely known. It is the uncertainties (e.g., in completeness) in the line-of-sight directions that is most prone to systematic misassessment. We therefore focus here on completeness misjudgements (with a simplified angular pattern) and investigate how much this could affect the recovery of the potential. We do this by creating mock data in the `DHB-Pot` within a spherical survey volume with radius $r_{\rm max}$ around the Sun (see Test 5 in Table 2.3) and a spatially varying completeness function

$$\text{completeness}(r) \equiv 1 - \epsilon_r \frac{r}{r_{\rm max}}, \tag{2.11}$$

which drops linearly with the distance r from the Sun. The completeness function can be understood as the probability of a star at distance r to be detected (see also Equation 2.2). In the RoadMapping analysis, on the other hand, we assume constant completeness ($\epsilon_r = 0$). The

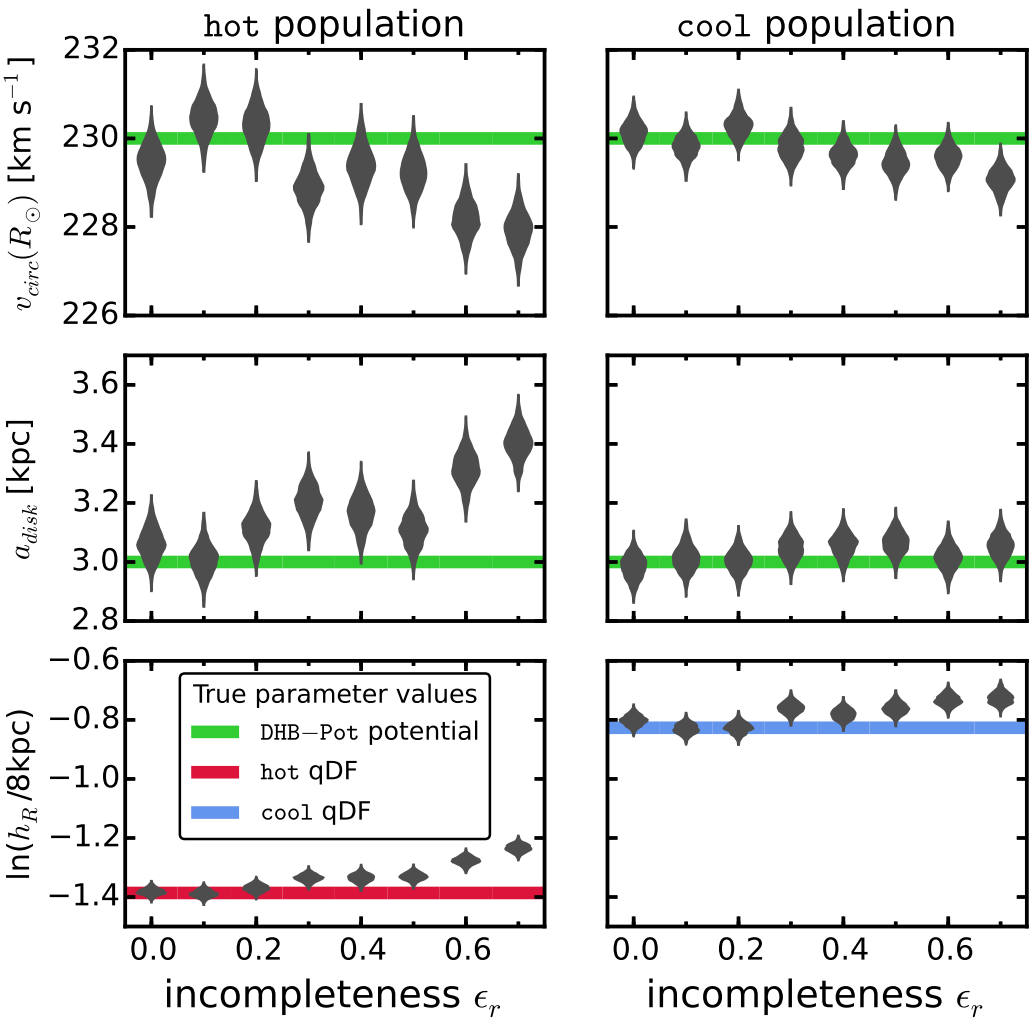


Figure 2.9: Impact of misjudging the completeness of the data on the parameter recovery with RoadMapping. Each mock data set was created with a different incompleteness parameter, ϵ_r, (shown on the x-axis, see Equation (2.11)). (The model parameters are given as Test 5 in Table 2.3.) The analysis, however, assumed that all data sets had constant completeness within the survey volume ($\epsilon_r = 0$). The violins show the full shape of the projected PDFs for each model parameter, and the solid lines their true values. The RoadMapping method seems to be robust against modest deviations between the true and the assumed data incompleteness. (The potential parameter $f_{\rm halo}$, and the other qDF parameters are recovered to a comparable accuracy and are therefore not shown here.)

incompleteness parameter ϵ_r of the mock data therefore quantifies how much we misjudge the SF. This mock test captures the relevant case of stars being less likely to be observed (than assumed) the further away they are (e.g., due to unknown dust obscuration).

Figure 2.9 demonstrates that the potential recovery with RoadMapping is quite robust against somewhat wrong assumptions about the completeness of the data, i.e., $\epsilon_r \lesssim 0.15$ for the `hot` and $\epsilon_r \lesssim 0.2$ for the `cool` population. The `cool` population is more robust, because it is less affected by the SF misjudgement at high $|z|$ than the `hot` population. Our simple model SF affects stars at large and small radii in equal proportion. As long as the misjudgement is small, the tracer scale length parameter, h_R, can still be reliable recovered, and, with it, the potential.

We have also investigated several test suites using the `Iso-Pot`. The recovery of $v_{\rm circ}(R_\odot)$ and the qDF parameters at different ϵ_r is qualitatively and quantitatively similar to Figure 2.9 for

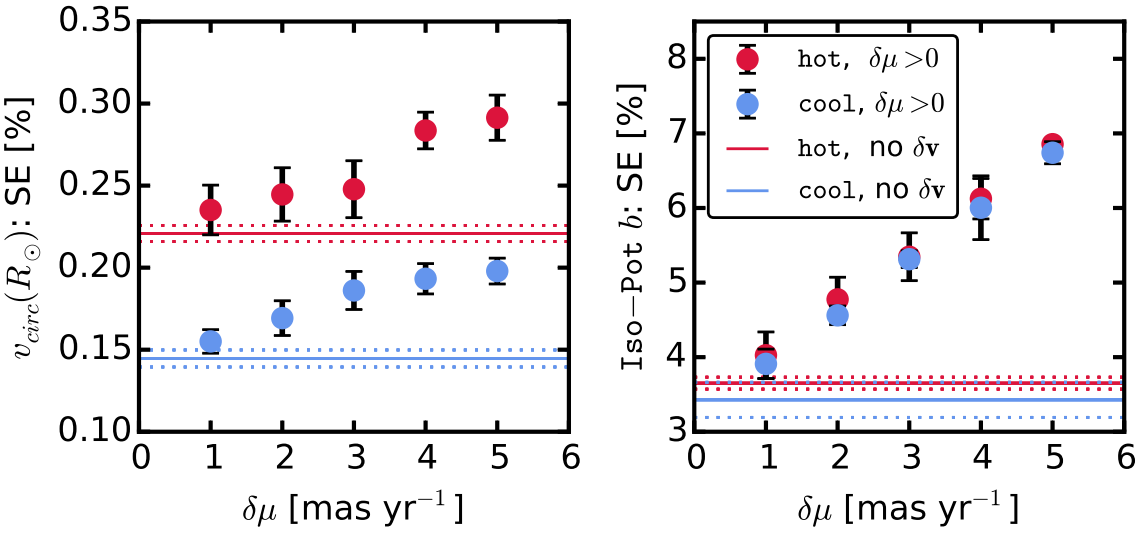


Figure 2.10: Effect of proper motion uncertainties, $\delta\mu$, on the precision of potential parameter recovery for two stellar populations of different kinematic temperature (see Test 6.1 in Table 2.3 for all model parameters). The relative standard error (SE) derived from the marginalized PDF for each model parameter was determined for precise data sets without measurement uncertainties (solid lines, with dotted lines indicating the error) and for data sets affected by different proper motion uncertainties $\delta\mu$ and $\delta v_{\rm los} = 2$ km s^{-1} (data points with error bars), but no uncertainties in position. The errors come from taking the mean over several data sets.

the `DHB-Pot`. The isochrone scale length, b, however, is recovered independently of ϵ_r—probably because rotation curve measurements in the plane alone, which are not affected by the SF cuts, give reliable constraints on b. When not including tangential velocity measurements in the analysis (which is done by marginalizing the likelihood in Equation (2.3) over v_T), the parameters are well recovered only for $\epsilon_r \lesssim 0.15$ and $\epsilon_r \lesssim 0.2$ for the `hot` and `cool` population, respectively. As this is in concordance with our findings for the `DHB-Pot`, this result seems to be valid for different choices of potentials.
For spatial completeness functions varying with the distance from the plane $|z|$ only, the `Iso-Pot` potential recovery is similarly robust as long as v_T measurements are included.

2.3.4 Measurement Uncertainties and their Effect on the Parameter Recovery

Measurement uncertainties in proper motions and distance dominate over uncertainties in position on the sky (R.A., decl.) and line-of-sight velocity, which can be more accurately determined.
The range of proper motion uncertainties in this section, $1 - 5$ mas yr^{-1}, is the approximate measurement accuracy that can be achieved by combining catalogs from ground-based surveys like the Sloan Digital Sky Survey (SDSS; Abazajian et al. 2003), the USNO-B catalog (Monet et al. 2003), 2MASS (Skrutskie et al. 2006), and the Pan-STARRS1 photometric catalog (PS1; Kaiser et al. 2010).[10] Space-based surveys can do even better. The Hipparcos (ESA 1997) and

[10]Combining observations from the SDSS DR1 with the USNO-B catalog based on the Palomar Observatory Sky Survey's (POSS) photographic plates from the 1950s lead to proper motion measurements precise to $\delta\mu \sim 3$ or 5 mas yr^{-1} depending on magnitude $r < 18$ or $r < 20$, respectively (Gould & Kollmeier 2004; Munn et al. 2004, 2008). The same accuracy can be achieved when using the four years of measurements by the PS1 only. By careful calibration of USNO-B and 2MASS with PS1, Sesar et al. (2015) even got proper motions as accurate as $\delta\mu \sim 1.5$ mas yr^{-1} for $r \lesssim 18$. The Large Synoptic Survey Telescope (LSST, Ivezic et al. 2008) planned for 2021 might even achieve $\delta\mu \lesssim 1$ mas yr^{-1} during its 10 years of scanning the sky (Ivezić et al. 2008).

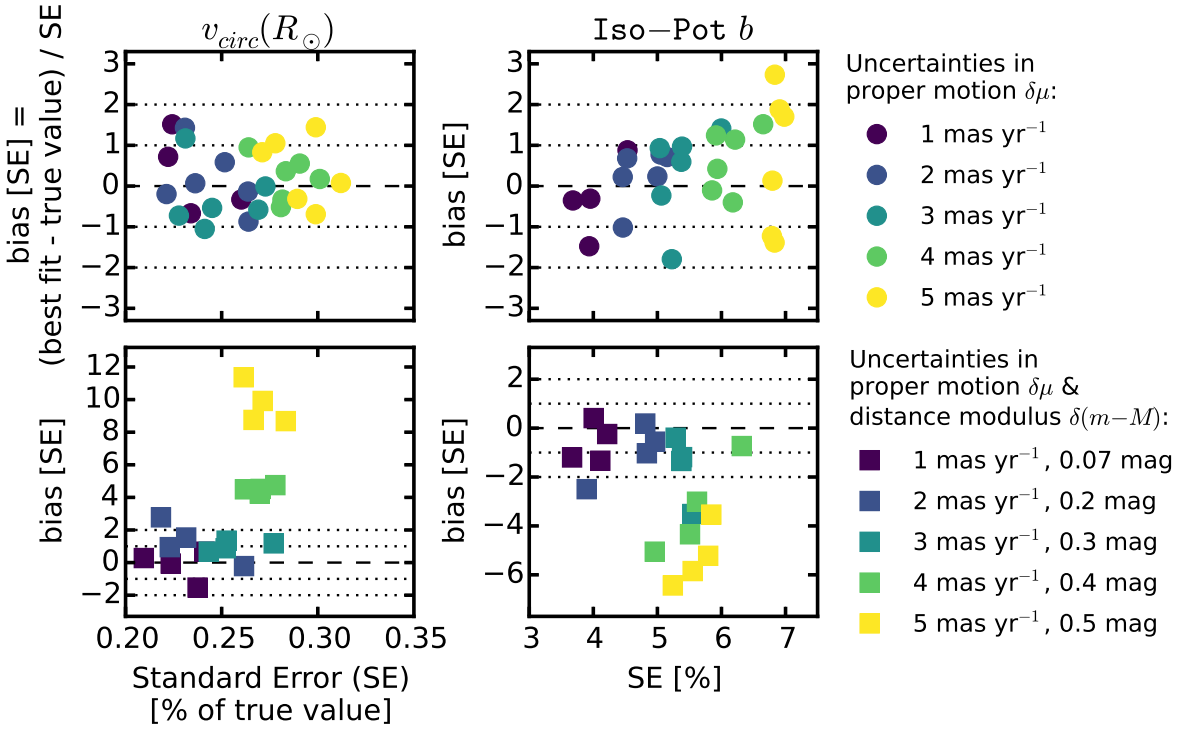


Figure 2.11: Potential parameter recovery using the approximation for the model probability convolved with measurement uncertainties in Equation (2.10). We show the PDF offset and relative width (i.e., standard error SE) for potential parameters recovered from mock data sets (which were created according to Test 6.2 in Table 2.3). The data sets in the upper panels are affected only by proper motion uncertainties, $\delta\mu$, (and $\delta v_{\rm los} = 2$ mas yr^{-1}), while the data sets in the lower panels also have distance (modulus) uncertainties, $\delta(m-M)$, as indicated in the legend. For data sets with $\delta\mu \leq 3$ mas yr^{-1} Equation (2.10) was evaluated with $N_{\rm samples} = 800$, for $\delta\mu > 3$ mas yr^{-1} we used $N_{\rm samples} = 1200$. In the absence of distance uncertainties, Equation (2.10) gives unbiased results. For $\delta(m-M) > 0.2$ mag (i.e., $\delta r/r > 0.1$; for $r \sim 3$ kpc), however, biases of several σ are introduced, as Equation (2.10) is only an approximation for the true likelihood in this case.

Tycho-2 (Høg et al. 2000) catalogs achieve $\delta\mu \sim 2.5$ mas yr^{-1} (and even $\delta\mu \lesssim 1$ mas yr^{-1} for all stars with $V < 12$), which will be soon superseded by *Gaia* with only $\delta\mu \sim 0.3$ mas yr^{-1} at its faint end at magnitude $G \sim 20$ (de Bruijne et al. 2014).

We first investigate the impact of (perfectly known) proper motion uncertainties on the precision of the potential parameter recovery (see Test 6.1 in Table 2.3). Figure 2.10 demonstrates that for data sets with $\delta\mu$ as high as 5 mas yr^{-1} the precision degrades by a factor of no more than ~ 2 as compared to a data set without measurement uncertainties. The precision gets monotonically better for smaller $\delta\mu$, being larger only by a factor of ~ 1.15 at $\delta\mu = 1$ mas yr^{-1}. With relative standard errors on the recovered parameters of only a few percent at most for 10,000 stars, this means we still get quite precise constraints on the potential, as long as we know the proper motion uncertainties perfectly.

We also note that, in this case, the relative and absolute difference in recovered precision between the precise and the uncertainty-affected data sets does not seem to depend strongly on the kinematic temperature of the stellar population.

Secondly, we investigate the impact of additional measurement uncertainties in distance (modulus). In the absence of distance uncertainties, the uncertainty-convolved model probability given in Equation (2.10) is unbiased (see upper left panel in Figure 2.11). When including distance (modulus) uncertainties, Equation (2.10) is just an approximation for the true likelihood; the

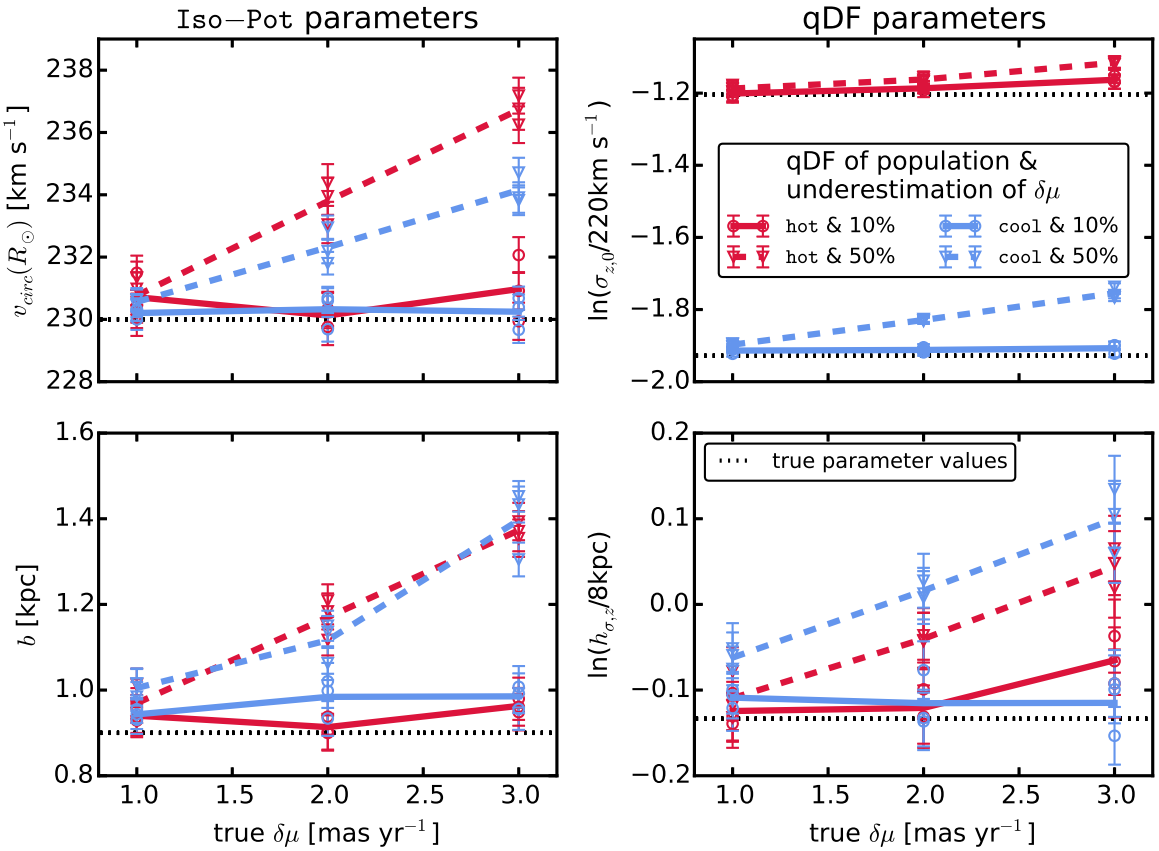


Figure 2.12: Effect of a systematic underestimation of proper motion uncertainties, $\delta\mu$, on the recovery of the model parameters. (The true model parameters used to create the mock data are summarized as Test 6.3 in Table 2.3, four of them are indicated as black dotted lines in this figure.) The mock data was perturbed according to proper motion uncertainties, $\delta\mu = \delta\mu_{\mathrm{decl.}} = \delta\mu_{\mathrm{R.A.}}$, as indicated on the x-axis. In the RoadMapping analysis (see likelihood in Equation (2.10)), however, we underestimated the true $\delta\mu$ by 10% (circles, solid lines) and 50% (triangles, dashed lines). The symbols denote the best fit parameters with 1σ error bars of several mock data sets. The lines connect the mean of corresponding data realizations to guide the eye.

systematic bias thus introduced in the parameter recovery gets larger with the size of $\delta(m-M)$, as demonstrated in Figure 2.11, lower panels (see also Test 6.2 in Table 2.3). We find, however, that in the case of $\delta(m-M) \lesssim 0.2$ mag (if also $\delta\mu \lesssim 2$ mas yr^{-1} and a maximum distance of $r_{\mathrm{max}} = 3$ kpc, see Test 6.2 in Table 2.3) the potential parameters can still be recovered within 2σ. This corresponds to a relative distance uncertainty of $\sim 10\%$. The overall precision of the potential recovery is also not degraded much by introducing distance uncertainties of less than 10%.

How does this compare with the distance uncertainties expected for *Gaia*? For a typical red clump giant star with $M_I \sim 0$ mag and $(V-I) \sim 1$ mag at a distance of $r = 3$ kpc we estimate (using the magnitude transformation by Jordi et al. (2010) and the uncertainty parameterization by de Bruijne et al. (2014)) a parallax uncertainty of $\delta\pi \sim 11$ μas, which is consistent with a distance uncertainty of less than 5%, and a proper motion uncertainty of $\delta\mu \sim 6$ μas yr^{-1}, which is negligible. For the case that the modeling is not restricted to giant stars only, a quick investigation by Rene Andrae (private communication) of stars at $r = 3$ kpc $\pm$ 5 pc from the *Gaia* Universe model snapshot catalog (GUMS; Robin et al. 2012) revealed that a magnitude cut at $G \sim 15$ mag in the overall *Gaia* data set should keep all distance uncertainties within 3 kpc

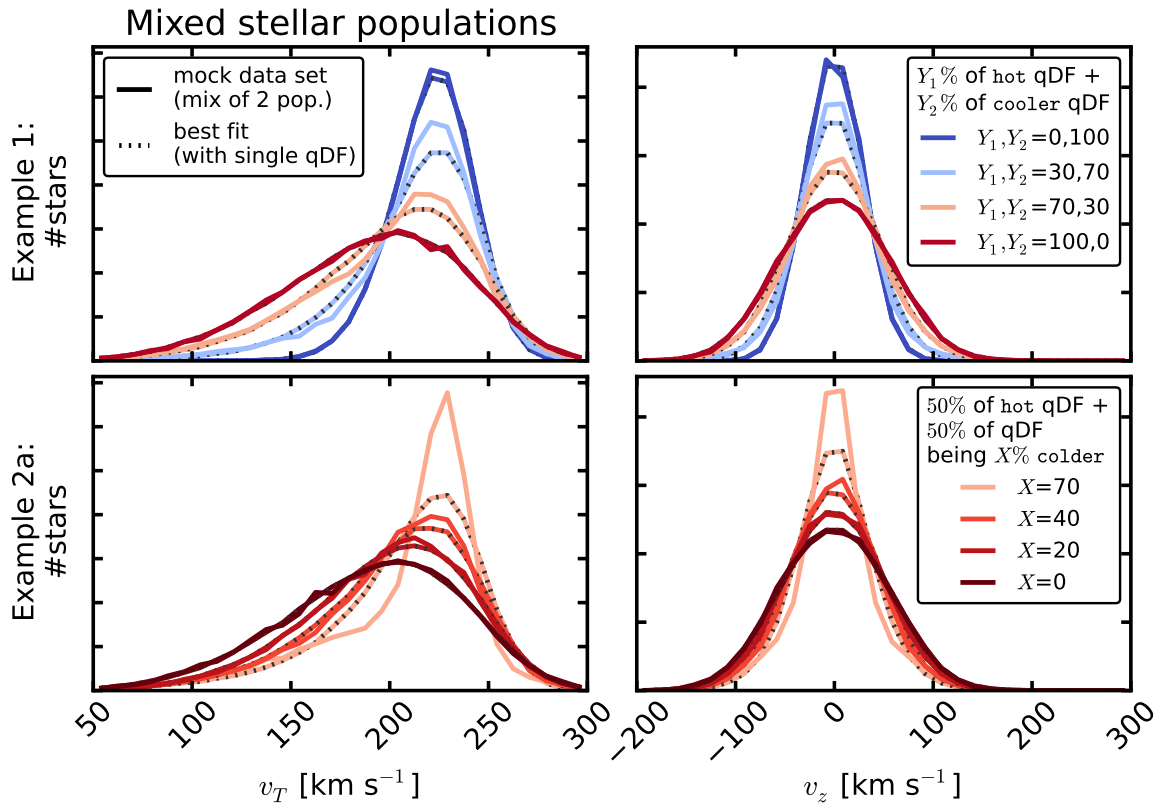


Figure 2.13: Distribution of mock data v_T and v_z created by mixing stars drawn from two different qDFs (solid lines), and the distribution predicted by the best fit of a single qDF and potential to the data (dotted lines). (The model parameters used to create the mock data are given in Table 2.3 as Test 7, *Example 1 and 2a*, with the qDF parameters referred to in the legend given in Table 2.2.) The corresponding single qDF best fit curves were derived from the best fit parameters found in Figures 2.15 and 2.16. (The data sets are color-coded in the same way as the corresponding analyses in Figures 2.15 and 2.16.) We use the mixtures of two qDFs to demonstrate how RoadMapping behaves for data sets following DFs with shapes differing slightly from a single qDF. For large deviations it might already become visible from directly comparing the mock data and best fit distribution, that a single qDF is a bad assumption for the stars' true DF.

below $\sim 10\%$, while also preserving *Gaia*'s simple SF.

We therefore found that in case we perfectly know the measurement uncertainties (and the distance uncertainty is negligible or of the order of the uncertainties expected from *Gaia* within ~ 3 kpc), the convolution of the model probability with the measurement uncertainties gives *precise and accurate* constraints on the model parameters—even if the measurement uncertainty itself is quite large.

Lastly, Figure 2.12 investigates the effect of a systematic underestimation of the true proper motion uncertainties, $\delta\mu$, by 10% and 50% (see also Test 6.3 in Table 2.3). We find that this causes a bias in the parameter recovery that grows seemingly linear with $\delta\mu$. For an underestimation of only 10%, however, the bias becomes $\lesssim 2\sigma$ for 10,000 stars—even for $\delta\mu \sim 3$ mas yr^{-1}.

The size of the bias also depends on the kinematic temperature of the stellar population and the model parameter considered (see Figure 2.12). The qDF parameters are, for example, better recovered by hotter populations. This is because the relative difference between the *true* $\sigma_i(R)$ (with $i \in \{R, z\}$) and *measured* $\sigma_i(R)$ (which comes from the deconvolution with an underestimated velocity uncertainty) is smaller for hotter populations.

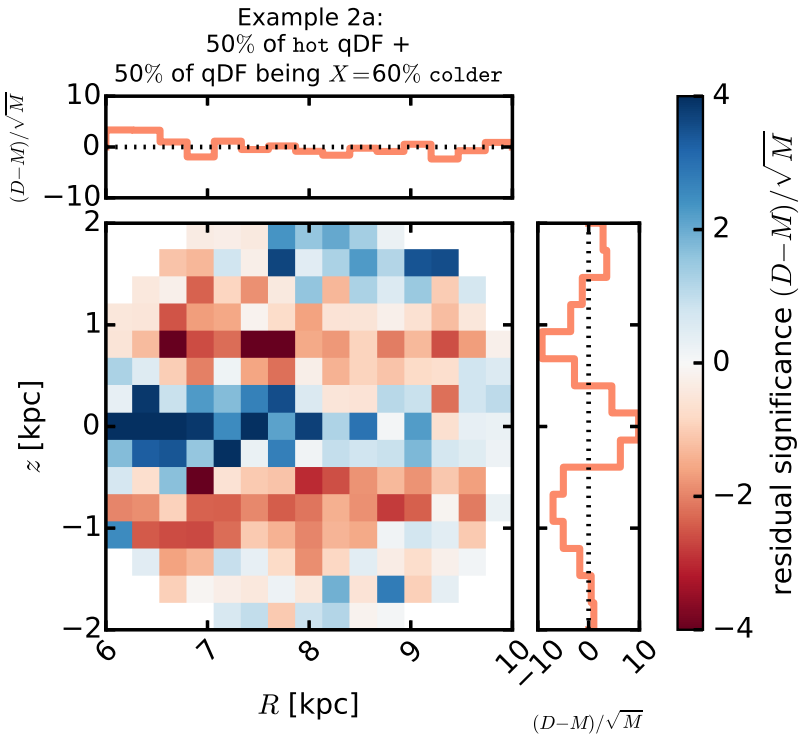


Figure 2.14: Residual significance, $(D-M)/\sqrt{M}$, of one example mock data set, D, and its best fit single qDF model, M, in the (R,z) plane. The mock data set, D, was created by mixing a `hot` and an X% `colder` population in equal proportion (see also Table 2.3, Test 7, *Example 2a*, with $X = 60\%$). The best fit distribution, M, was derived analogously to the ones shown for the velocity components in Figure 2.13. This is an extreme example where the best fit single qDF is not a good fit anymore (see Figure 2.16, *Example 2a*, $X = 60\%$), but it illustrates how we constructed mock data distributions with radial and vertical density profiles differing from a single qDF by mixing two different qDF populations for the Test suite 7.

2.3.5 The Impact of Deviations of the Data from the Idealized DF

Our modeling approach assumes that each stellar population follows a simple DF; here we use the qDF. In this section, we explore what happens if this idealization does not hold. We investigate this issue by creating mock data sets that are drawn from *two* distinct qDFs of different temperatures[11] (see Table 2.2 and Test 7 in Table 2.3) in the `DHB-Pot`, and analyze the composite mock data set by fitting a *single* qDF to it. The velocity distributions of some mock data sets and their best fit qDFs are illustrated in Figure 2.13, and Figure 2.14 shows the tracer density residuals between data and best fit in the (R,z) plane. Figures 2.15 and 2.16 compare the input and best fit parameters. In *Example 1*, we choose qDFs of widely different temperatures and vary their relative fraction of stars in the composite mock data set (Figure 2.15); in *Example 2* we always mix mock data stars from two different qDFs in equal proportion, but vary by how much the qDFs' temperatures differ (Figure 2.16).

The first set of tests mimics a DF that has wider wings or a sharper core in velocity space than a qDF (see Figure 2.13), and slightly different radial and vertical tracer density profiles (similar to Figure 2.14). The second test could be understood as mixing neighboring MAPs in the [α/Fe]-[Fe/H] plane due to large bin sizes or abundance measurement errors (cf. Bovy & Rix (2013)).

[11] Following the observational evidence, our mock data populations with cooler qDFs also have longer tracer scale lengths.

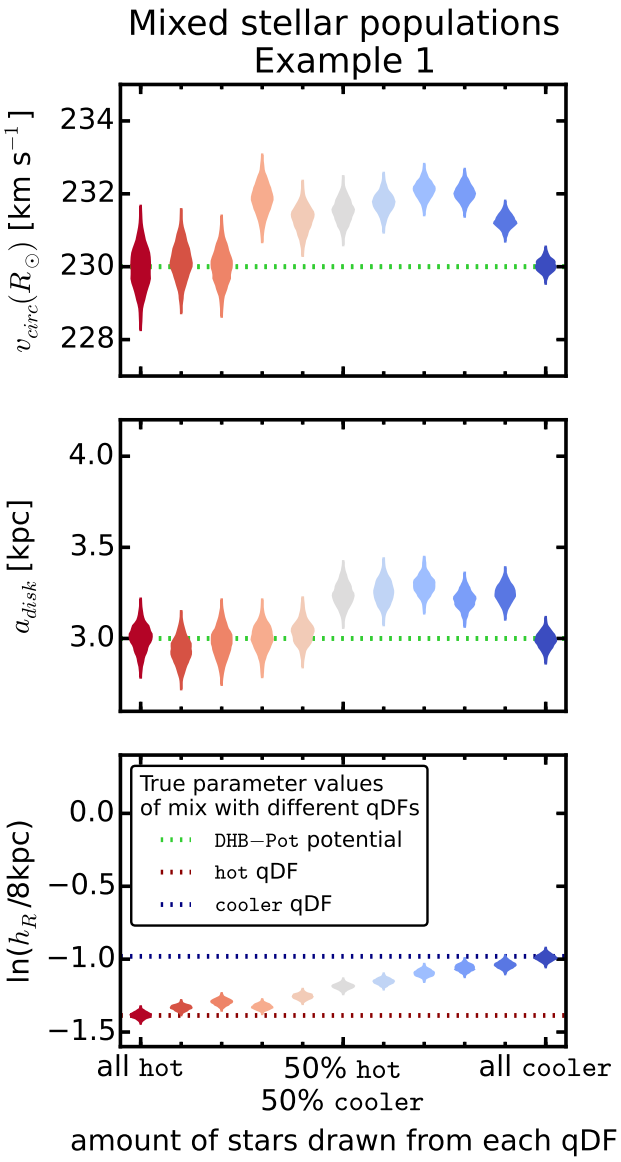


Figure 2.15: The dependence of the parameter recovery on the degree of pollution and temperature of the stellar population. We mix (i.e., "pollute") varying amounts of stars from a `hot` stellar population with stars from a very different `cooler` population (see Table 2.2), as indicated on the x-axis. (All model parameters used to create the mock data are given as Test 7, *Example 1*, in Table 2.3.) The composite polluted mock data set follows a true DF that has a slightly different shape than the qDF. We then analyze it using RoadMapping and fit a *single* qDF only. The violins represent the marginalized PDFs for the best fit model parameters. Some mock data sets are shown in Figure 2.13, first row, in the same colors as the violins here. We find that a hot population is much less affected by pollution with stars from a cooler population than vice versa. (The potential parameter $f_{\rm halo}$ is recovered to a similar or even slightly better accuracy than $a_{\rm disk}$ at each given mixing rate and is therefore not shown here.)

We consider the impact of the DF deviations on the recovery of the potential and the qDF parameters separately.

We find from *Example 1* that the potential parameters can be more robustly recovered if a mock data population is polluted by a modest fraction ($\lesssim 30\%$) of stars drawn from a much cooler qDF, as opposed to the same pollution of stars from a hotter qDF. When considering the case of a 50/50 mix of contributions from different qDFs in *Example 2* there is a systematic, but mostly small, bias in recovering the potential parameters, monotonically increasing with the qDF parameter difference. In particular, for fractional differences in the qDF parameters of $\lesssim 20\%$, the systematics are insignificant even for sample sizes of $N_* = 20,000$, as used in the mock data. Overall, the circular velocity at the Sun is very reliably recovered to within 2% in all these tests. But the best fit $v_{\rm circ}(R_\odot)$ is not always unbiased at the implied precision.

The recovery of the effective qDF parameters, in light of non-qDF mock data, is quite intuitive (in Figures 2.15 and 2.16 we therefore show only h_R): the effective qDF temperature lies between the two temperatures from which the mixed DF of the mock data was drawn; in all cases the scale lengths of the velocity dispersion fall-off, $h_{\sigma,R}$ and $h_{\sigma,z}$, are shorter than the true scale

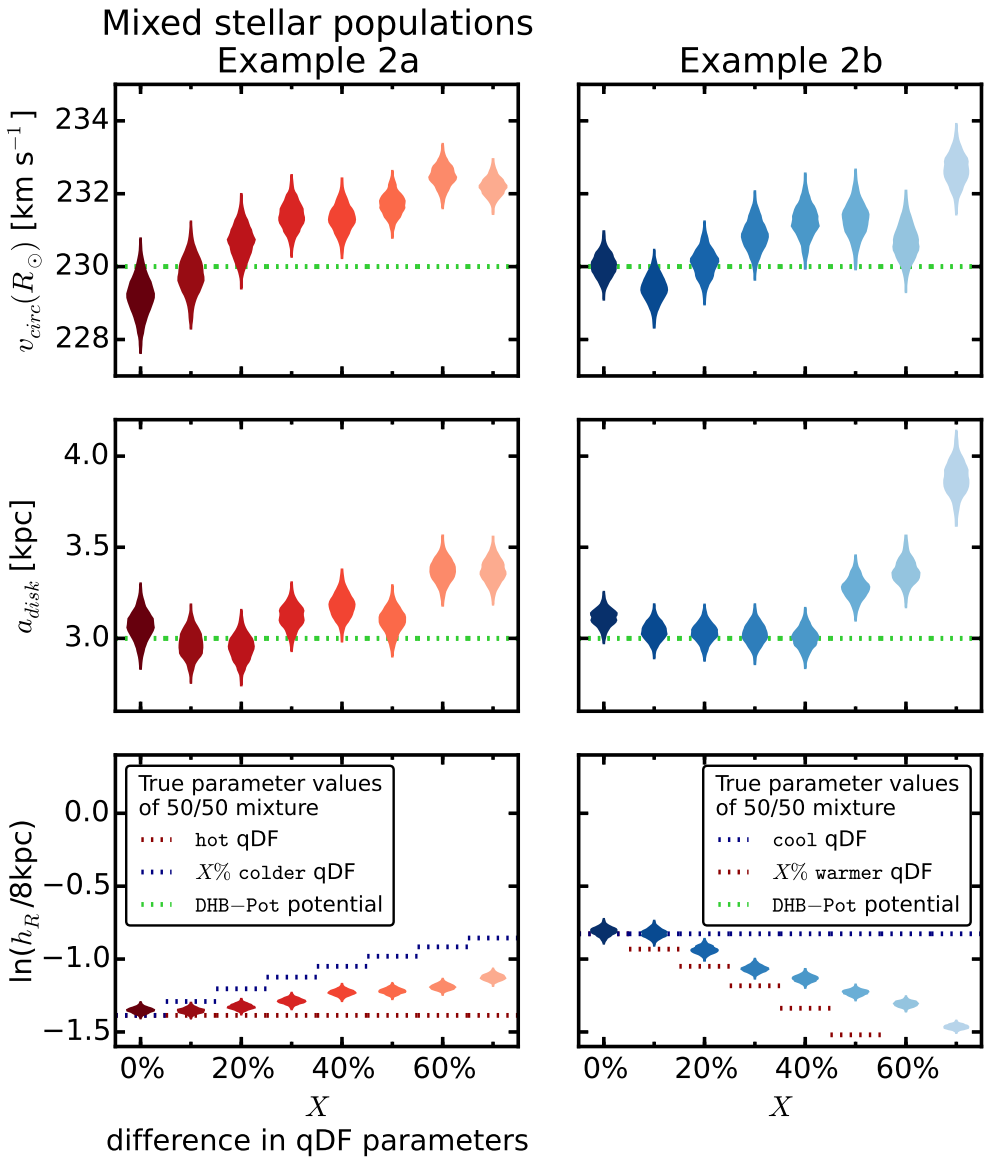


Figure 2.16: The dependence of the parameter recovery on the difference in qDF parameters of a 50/50 mixture of two stellar populations and their temperature. The two qDFs from which the stars in each mock data set were drawn are indicated in the legend, with the qDF parameters $\sigma_{R,0}, \sigma_{z,0}$ and h_R differing by X% (see also Table 2.2 and Section 2.2.4) as indicated on the x-axis. (The model parameters used for the mock data creation are given as Test 7, *Example 2a and b*, in Table 2.3.) Each composite mock data set is fitted with a *single* qDF, and the marginalized PDFs are shown as violins. Some mock data sets of *Example 2a* and their best fit distributions are shown in Figure 2.13, last row (color-coded analogous to the violins here), and Figure 2.14 shows the corresponding residuals in the (R, z) plane. By mixing populations with varying difference in their qDF parameters, we model the effect of finite bin size or abundance errors when sorting stars into different MAPs in the [α/Fe]-[Fe/H] plane and assuming that they follow single qDFs (cf. Bovy & Rix (2013)). We find that the bin sizes should be chosen such that the difference in qDF parameters between neighboring MAPs is less than 20%. (The potential parameter $f_{\rm halo}$ is recovered to a similar or even slightly better accuracy than $a_{\rm disk}$ at each given X and is therefore not shown here.)

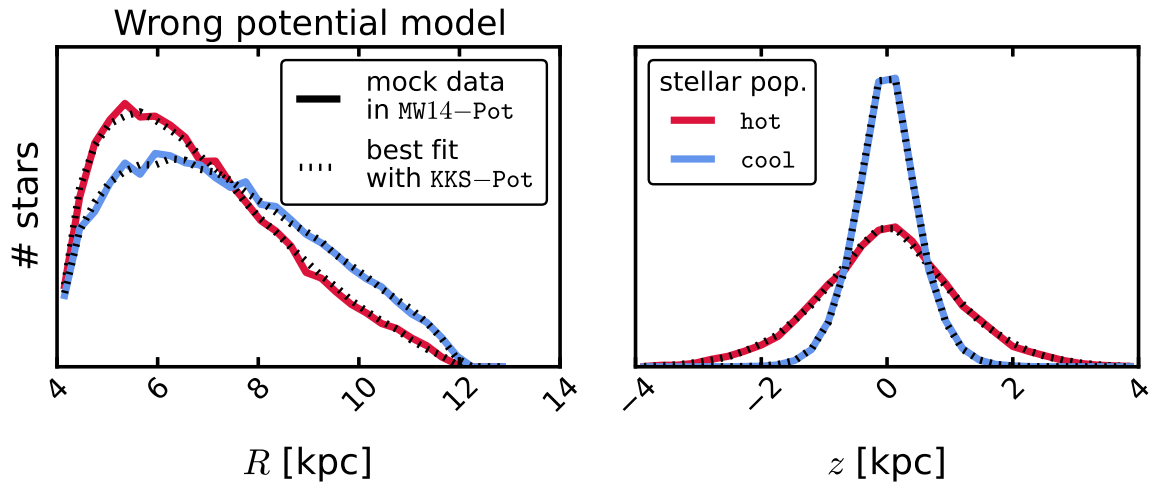


Figure 2.17: Comparison of the spatial distribution of mock data in R and z created in the `MW14-Pot` potential and with two different stellar populations (see Test 8 in Table 2.3 for all mock data model parameters), and the best fit distribution recovered by fitting the family of `KKS-Pot` potentials to the data. The best fit potentials are shown in Figure 2.18 and the corresponding best fit qDF parameters in Figure 2.19. The data is very well recovered, even though the fitted potential family did not incorporate the true potential.

lengths because the stars drawn form the hotter qDF dominate at small radii, while stars from the cooler qDF (with its longer tracer scale length) dominate at large radii; the recovered tracer scale lengths, h_R, vary smoothly between the input values of the two qDFs that entered the mix of mock data. The latter is also demonstrated in Figure 2.14: the radial tracer density profile of the mock data is steeper than a single qDF in the mid-plane and more shallow at higher $|z|$; overall the best fit h_R therefore lies in between.
We note that in the cases where the systematic bias in the potential parameter recovery becomes as big as several σ, a direct comparison of the true mock data set and best fit distribution (see Figure 2.13) can sometimes already reveal that the assumed DF is not a good model for the data.
We also performed the same tests for the spherical `Iso-Pot` instead of the galaxy-like `DHB-Pot` and for a much higher sampling of the mixing rate and qDF difference, X. The results are qualitatively and quantitatively very similar and therefore independent of the exact choice of potential.
Overall, we find that the potential inference is quite robust to modest deviations of the data from the assumed DF.

2.3.6 The Implications of a Gravitational Potential not from the Space of Model Potentials

We now explore what happens when the mock data were drawn from one axisymmetric potential family, here `MW14-Pot`, and are then modeled considering potentials from another axisymmetric family, here `KKS-Pot` (see Table 2.1 and Figure 2.1). In the analysis, we assume the circular velocity at the Sun to be fixed and known, and we only fit the parametric potential form.[12]
We analyze a mock data set from a `hot` and `cool` stellar population each (see Test 8 in Table 2.3) with high numerical accuracy. The distributions generated from the best fit parameters reproduce the data in configuration space very well (see Figure 2.17 for the spatial distribution and the circles in Figure 2.19 for the velocity distribution).

[12]We made sure that $v_{\rm circ}(R_\odot)$ can be very well recovered when included in the fit of a `cool` population. The model assumption that $v_{\rm circ}(R_\odot)$ is known does therefore not affect the discussion qualitatively.

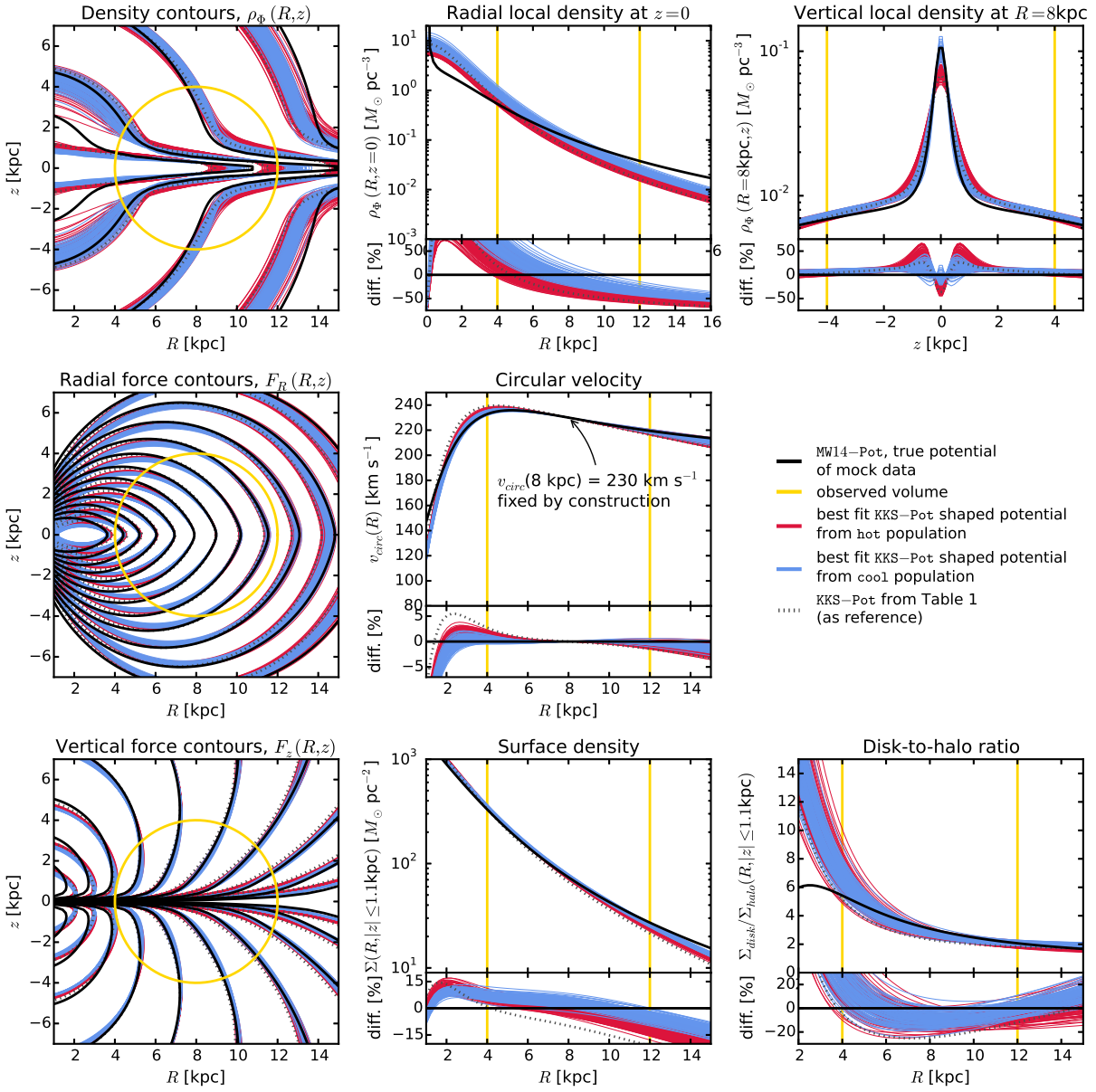


Figure 2.18: Recovery of the gravitational potential if the assumed potential model family (`KKS-Pot` with fixed $v_{\rm circ}(R_\odot)$) and the true potential of the (mock data) stars (`MW14-Pot` in Table 2.1) have slightly different parametric forms. In addition to contours of equal density, ρ_Φ, radial and vertical force, F_R and F_z, in the (R,z) plane (left column), we show local density profiles, $\rho_\Phi(R,z=0)$ and $\rho_\Phi(R=8\ {\rm kpc},z)$, as well as the circular velocity curve, $v_{\rm circ}(R)$, the total surface density profile within $|z|\leq 1.1$ kpc, $\Sigma(R)\equiv\int_{-1.1{\rm kpc}}^{1.1{\rm kpc}}\rho_\Phi(R,z)\,{\rm d}z$, and the ratio of the disk and halo contributions to the total surface density, $\Sigma_{\rm disk}(R)/\Sigma_{\rm halo}(R)$. We compare the true potential (black lines) with 100 sample potentials (red and blue lines) drawn from the PDF found with MCMC for a `hot` (red) and a `cool` (blue) stellar population, and also display the relative difference as a percentage of the true value. (All mock data model parameters are given as Test 8 in Table 2.3.) Overall, the true potential is well recovered—especially in regions where most of the observed stars are located.

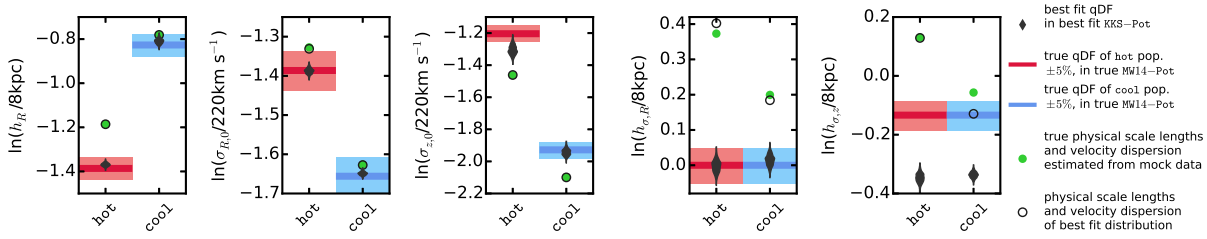


Figure 2.19: Recovery of the qDF parameters for the case where the true and assumed potential deviate from each other (see Test 8 in Table 2.3). The thick red (blue) lines represent the true qDF parameters of the `hot` (`cool`) qDF in Table 2.2 used to create the mock data, surrounded by a 5% error region. The gray violins are the marginalized PDFs for the qDF parameters found simultaneously with the potential constraints shown in Figure 2.18. We compare the qDF parameters with the actual physical scale lengths and velocity dispersion at the Sun estimated from the mock data and the best fit distribution by fitting exponential functions to the data. First, this shows that—apart from some small deviations in the velocity dispersion scale lengths—the velocity distribution of the mock data is very well reproduced by the best fit. Second, this demonstrates how the qDF parameters in different potentials do not necessarily agree with each other or with the actual physical velocity distribution.

The comparison between true and best fit potentials are shown in Figure 2.18. We find that the potential recovered by RoadMapping is in good agreement with the true potential inside of the observed volume of mock tracers. Outside of it we can make predictions, at least to a certain extent. Especially the potential forces, to which the stellar orbits are sensitive, are recovered and tightly constrained. This robust recovery of the radial and vertical forces leads to small errors on the estimated circular velocity curve ($\lesssim 5\%$) and surface density within $|z| = 1.1$ kpc ($\lesssim 10\%$), respectively. We get the best results for the local density, the surface density, and disk-to-halo ratio between $R \sim 4$ kpc and $R \sim 8$ kpc, i.e., where most of the tracer stars used in the analysis are located (see Figure 2.17).

The local density distribution is, in general, less reliably constrained than the forces, but we still capture the essentials. Exceptions are the inner regions $R \lesssim 3$ kpc, where the `KKS-Pot` model is missing a bulge by construction, and the local radial density profile, which is somewhat misjudged by the `KKS-Pot` model. The `cool` population, where most stars are confined to regions close to the mid-plane, recovers the flatness of the disk better than the `hot` population, but overall the best fit disk is slightly less dense in the mid-plane than the true disk. While it is generally possible to generate very flattened density distributions from Stäckel potentials, it might be difficult to simultaneously have a roundish halo and to require that both Stäckel components have the same focal distance (see Table 2.1).

The disk-to-halo surface density fraction within $|z| = 1.1$ kpc is not tightly constrained ($\gtrsim 20\%$), but recovered within the errors inside of the survey volume. Using a wrong potential model does therefore not necessarily lead to biases in local DM measurements.

Overplotted in Figure 2.18 is also the `KKS-Pot` with the parameters from Table 2.1, which were fixed based on a (by-eye) fit directly to the force field (within $r_{\max} = 4$ kpc from the Sun) and the rotation curve of the `MW14-Pot`. The potential found with the RoadMapping analysis is an even better fit. This demonstrates that RoadMapping fitting infers a potential that, in its actual properties, resembles the input potential for the mock data in regions of large tracer density as closely as possible, given the differences in functional forms.

Figure 2.19 compares the true qDF parameters with the best fit qDF parameters belonging to the best fit potentials from Figure 2.18, and we also overplot the actual physical scale lengths and velocity dispersion as estimated directly from the mock data. While we recover h_R, $\sigma_{R,0}$, and $h_{\sigma,R}$ within the errors, we misjudge the parameters of the vertical velocity dispersion ($\sigma_{0,z}$ and especially $h_{\sigma,z}$), even though the actual mock data distribution is well reproduced. This

discrepancy could be connected to the `KKS-Pot` not being able to reproduce the flatness of the disk. Also, σ_z and σ_R in Equations (1.64)-(1.65) are scaling profiles for the qDF (cf. Bovy & Rix (2013)), and how close they are to the actual velocity profile depends on the choice of potential; that is, the *physical* velocity dispersion is well recovered even if the qDF velocity dispersion parameters are not. Figure 2.19 stresses once more that the actual parameter values of action-based DFs have always to be considered together with the potential in which they were derived. This is of importance in studies that use a fiducial potential to fit action-based DFs to stellar data, like, e.g., Sanders & Binney (2015b) and Das & Binney (2016).

2.3.7 The Influence of the Stellar Population's Kinematic Temperature

Overall, we found that it does not make a big difference if we use hot or cool stellar populations in our modeling.

How precise and reliable model parameters can be recovered does, to a certain extent, depend on the kinematic temperature of the data, as well as on the model parameter in question and on the observation volume. But there is no easy rule of thumb as to what combination would give the best results (see Figure 2.7). There are two exceptions.

First, the circular velocity at the Sun, $v_{\rm circ}(R_\odot)$, is always best recovered with cooler populations (see Figures 2.10, 2.12, 2.15, and 2.16, and for the recovery of $v_{\rm circ}(R)$ at $R \neq R_\odot$ see Figure 2.18) because more stars are on near-circular orbits (see Figure 2.2). Cooler populations are also less sensitive to misjudgements of (spatial) SFs at large $|z|$ (see Figure 2.9). There is, however, the caveat that cool populations are more susceptible to non-axisymmetric streaming motions in the disk.

Second, hotter populations seem to be less sensitive to misjudgements of proper motion measurement uncertainties (see Figure 2.12) and pollution with stars from a cooler population (see Figures 2.15 and 2.16) because of their higher intrinsic velocity dispersion (see Figure 2.3).

In addition, we find indications in Figure 2.18 that different regions within the Galaxy are best probed by populations of different kinematic temperature; the `hot` population gives the best constraints on the radial local and surface density profiles at a smaller radius than the `cool` population because of its smaller tracer scale length. The `cool` population with most stars close to the mid-plane recovers the flatness of the disk more reliably.

2.4 Summary and Discussion

Recently, implementations of action DF-based modeling of 6D data in the Galactic disk have been put forth, in part to lay the ground-work for *Gaia* (Bovy & Rix 2013; McMillan & Binney 2013; Piffl et al. 2014; Sanders & Binney 2015b).

This chapter has presented RoadMapping, an improved implementation of the dynamical modeling machinery of Bovy & Rix (2013), to recover the MW's gravitational potential by fitting an orbit DF to stellar populations within the Galactic disk. In this chapter, we investigated the capabilities, strengths, and weaknesses of RoadMapping by testing its robustness against the breakdown of some of its assumptions—for well-defined, isolated test cases using mock data. Overall, the method works very well and is robust, even when there are small deviations of the model assumptions from the "real" Galaxy.

RoadMapping applies a full likelihood analysis and is statistically well-behaved. It goes beyond Bovy & Rix (2013) by allowing for a straightforward and flexible implementation of different model families for potential and DF. It also accounts for selection effects by using full 3D SFs (given some symmetries).

2.4.1 Computational Speed

Large data sets in the age of *Gaia* require increasingly accurate likelihood evaluations and flexible models. To be able to deal with these computational demands we sped up the RoadMapping code by combining a nested-grid approach with MCMC and by faster action calculation using the Stäckel (Binney 2012a) interpolation grid by Bovy (2015). Our approach therefore allows us to explore the full PDF, while similar studies (e.g., Piffl et al. 2014; Sanders & Binney 2015b; Das & Binney 2016) focus more on the model with maximum likelihood only. This also makes RoadMapping slower; fitting three `DHB-Pot` and five qDF parameters in each of the analyses in Tests 5 and 7 (see Table 2.3) takes, for example, $\sim 25 - 30$ hr on 25 CPUs. This is still a feasible computational effort as long as we restrict ourselves to potentials with a closed-form expression for $\Phi(R, z)$ (as done in this work). An equivalent analysis using, e.g., a double exponential disk (requiring integrals over Bessel functions) would take several days to weeks for $N_* = 20,000$. In any case, the application of RoadMapping to millions of stars will be a task for supercomputers, and calls for even more improvements and speed-up in the fitting machinery.

2.4.2 Properties of the Data Set

We could show that RoadMapping can provide potential and DF parameter estimates that are very accurate (i.e., unbiased) and precise in the limit of large data sets as long as the modeling assumptions are fulfilled.

In case the data set is affected by substantive measurement uncertainties, the potential can still be recovered to high precision, as long as these uncertainties are perfectly known and distance uncertainties are negligible. For large proper motion uncertainties, e.g., $\delta\mu \sim 5$ mas yr^{-1}, the formal errors on the parameters are only twice as large as in the case of no measurement uncertainties. However, properly accounting for measurement uncertainties is computationally expensive.

For the results to be accurate within 2σ (for 10,000 stars), we need to know to within 10% both the true stellar distances (at $r_{\max} \leq 3$ kpc and $\delta\mu \lesssim 2$ mas yr^{-1}) and the true proper motion uncertainties (with $\delta\mu \lesssim 3$ mas yr^{-1}).

The distance condition is an artifact of the likelihood approximation (Equation (2.10)) that RoadMapping uses to save computation time, and the reason why we will have to restrict the RoadMapping modeling to stars with small distance uncertainties.

Fortunately, the measurement uncertainties of the final *Gaia* data release with $\delta\mu \lesssim 0.3$ mas yr^{-1} at $G \lesssim 20$ mag, and $\delta r/r \lesssim 5\%$ at $r \sim 3$ kpc for $G < 15$ mag (see Section 2.3.4 and de Bruijne et al. 2014) will be well below these limits and promise accurate potential constraints. Before the final *Gaia* data release, however, we might have to restrict the modeling to suitable giant tracers with small uncertainties.

The main caveat of Tests 2 and 6.1-6.3 in Section 2.3.4 (see Table 2.3) concerning measurement uncertainties is the use of the `Iso-Pot`, which we chose for computational speed reasons. However, Tests 5 and 7, which we run for both `DHB-Pot` and `Iso-Pot`, gave qualitatively and quantitatively very similar results for both potentials. This makes us confident that also our results about measurement uncertainties are independent of the actual choice of potential.

We also found that the location of the survey volume within the Galaxy matters little. At a given sample size, a larger survey volume with large coverage in both radial and vertical direction will give the tightest constraints on the model parameters.

The potential recovery with RoadMapping seems to be robust against minor misjudgements of the spatial data SF, in particular to a completeness overestimation of $\lesssim 15 - 20\%$ at the edge of a survey volume with $r_{\max} = 3$ kpc.

We found indications that populations of different scale lengths and temperatures probe different regions of the Galaxy because the best potential constraints are achieved where most of the

stellar tracers are located. This supports the approach by Bovy & Rix (2013), who measured for each MAP the surface mass density only at one single best radius to account for missing flexibility in their potential model.
While cooler populations probe the Galaxy rotation curve better and hotter populations are less sensitive to pollution, overall stellar populations of different kinematic temperature seem to be equally well suited for dynamical modeling.

2.4.3 Deviations from the DF Assumption

RoadMapping assumes that stellar sub-populations can be described by simple DFs. We investigated how much the modeling would be affected if the assumed family of DFs would differ from the stars' true DF.
In *Example 1* in Section 2.3.5 we considered true stellar DFs being (i) hot with more stars with low velocities and less stars at small radii than assumed (reddish data sets in Figures 2.13 and 2.15), or (ii) cool with broader velocity dispersion wings and less stars at large radii than assumed (bluish data sets). We find that case (i) would give more reliable results for the potential parameter recovery.
Binning of stars into MAPs in $[\alpha/\mathrm{Fe}]$ and [Fe/H], as done by Bovy & Rix (2013), could introduce systematic errors due to abundance uncertainties or too large bin sizes—always assuming MAPs follow simple DF families (e.g., the qDF). In *Example 2* in Section 2.3.5, we found that, in the case of 20,000 stars per bin, differences of $\lesssim 20\%$ in the qDF parameters of two neighboring bins can still give quite good constraints on the potential parameters.
The relative differences in the qDF parameters $\sigma_{R,0}$ and $\sigma_{z,0}$ of neighboring MAPs in Figure 6 of Bovy & Rix (2013) (which have bin sizes of $\Delta[\mathrm{Fe/H}] = 0.1$ dex and $\Delta[\alpha/\mathrm{Fe}] = 0.05$ dex) are indeed smaller than 20%. For the h_R parameter, however, the bin sizes in Figure 6 of Bovy & Rix (2013) might not yet be small enough to ensure no more than 20% of difference in neighboring bins.
The qDF is a specific example for a simple DF for stellar sub-populations that we used in this paper, but it is not essential for the RoadMapping approach. Future studies might apply slight alternatives or completely different DFs to the data.
The RoadMapping approach could be considered as a means to constrain the Galactic potential treating the DF parameters as nuisance parameters. That we were able to show in this chapter that RoadMapping results are quite robust to the form of the DF not being entirely correct motivates this approach further.

2.4.4 Gravitational Potential Beyond the Parameterized Functions Considered

In addition to the DF, RoadMapping also assumes a parametric model for the gravitational potential. We test how using a potential of Stäckel form (`KKS-Pot`, Batsleer & Dejonghe 1994) affects the RoadMapping analysis of mock data from a different potential family with halo, bulge, and exponential disk (`MW14-Pot`, Bovy 2015). The potential recovery is quite successful; we properly reproduce the mock data distribution in configuration space, and the best fit potential is—within the limits of the model—as close as it gets to the true potential, even outside of the observation volume of the stellar tracers.
For as many as 20,000 stars, constraints already become so tight that it should presumably be possible to distinguish between different parametric MW potential models (e.g., the `DHB-Pot` and the `KKS-Pot`).
Fitting parameterized potentials of Stäckel form to MW data (see, e.g., Batsleer & Dejonghe 1994; Famaey & Dejonghe 2003) has the advantage of allowing action calculations that are accurate and fast. It does, however, limit the space of potentials that can be investigated, as different potential components are all required to have the same focal distance. Using the *Stäckel*

fudge (Binney 2012a) together with parameterized potentials made up from physically motivated building blocks (exponential disks, power-law DM halo etc.), as was done by Bovy & Rix (2013), seems to be the most promising approach—even though there still remain several challenges concerning computational speed to be solved.

2.4.5 Outlook

We know that real galaxies, including the MW, are not axisymmetric. Using N-body models, we will explore in the next chapter how the recovery of the gravitational potential with RoadMapping will be affected when data from a non-axisymmetric disk galaxy system with spiral arms gets interpreted through axisymmetric models.

CHAPTER 3

The Influence of Spiral Arms in RoadMapping Modeling

3.1 Preface

Introduction. In Chapter 2 we showed that RoadMapping is, in the axisymmetric case, robust under modest breakdowns of its modeling assumptions.

The MW is, however, not axisymmetric, as we have laid out in Section 1.2.4 (and references therein): The Galactic bar and the spiral arms are strong non-axisymmetric perturbations that are expected to trigger moving groups via resonances in the stellar disk. As RoadMapping and related approaches can only build axisymmetric models, this is an important breakdown of modeling assumptions, which was not investigated in Chapter 2. In this chapter, we want to understand in which respects RoadMapping will still give reliable constraints on the MW's gravitational potential in the presence of spiral arms.

Our investigation makes use of an N-body simulation snapshot of a spiral galaxy with strong spiral arms presented in (D'Onghia et al. 2013; see their Figure 8, top left panel). From this snapshot, we draw mock data in regions with different spiral arm strengths. We then apply the RoadMapping machinery to these data sets and test how well we recover the local and overall gravitational potential.

In Chapter 2, Sections 2.3.5-2.3.6, we confirmed and tested separately the robustness of RoadMapping in the case that the data came from a different model family—for either the potential or DF—than assumed in the dynamical modeling. What would happen if both potential and DF model families were slightly wrong at the same time? The setup of this study will automatically cover this important test case. The potential and orbit DF model that we are using were picked as a pragmatic compromise between (1) being a reasonable choice given the initial axisymmetric setup of the galaxy simulation and (2) because of their simplicity, computational advantages, and—in case of the qDF—because that is what we are planning to use in the MW. Given that the simulation has evolved away from its axisymmetric beginnings, we expect our chosen model to be reasonable, but not particularly well-suited to model this galaxy.

Spiral arms introduce another—but possibly minor—breakdown of the modeling assumptions: in a non-axisymmetric gravitational potential, the three actions will not be strict integrals of motions anymore (Binney & Tremaine 2008; Minchev et al. 2011, 2012; Grand et al. 2012; Solway et al. 2012; Vera-Ciro & D'Onghia 2016). It will be interesting to see if the action-based DF in RoadMapping modeling is still informative, even if it only uses the approximate actions estimated

in an axisymmetric potential.
In the MW, we expect the central bar to introduce additional non-axisymmetries in the Galactic disk—but presumably not stronger ones than the spiral arms. Because the galaxy simulation in this work does not have a central bar, we do not investigate specific bar effects here.
Though non-axisymmetry could be a severe problem for RoadMapping, we show in this paper that RoadMapping potential estimates are still surprisingly accurate, which makes us optimistic that they will also be so for the MW.

Structure of this chapter. This chapter is organized as follows. Section 3.2 describes the N-body simulation snapshot of a spiral galaxy that we are going to model in this study, explains how we extract 6D stellar phase-space data from it, and how we quantify the spiral arm strength. There, we also review similarities and differences between the simulation in this work and what we know about the MW. Section 3.3 summarizes the RoadMapping dynamical modeling framework, and introduces the DF and potential model that we will fit to the data. Section 3.4 is dedicated to presenting the results. In Section 3.4.1, we discuss in detail the RoadMapping modeling results derived from one data set within a survey volume with radius $r_{\max} = 4$ kpc around the Sun. Section 3.4.2 then investigates a whole suite of RoadMapping analyses, corresponding to survey volumes of different sizes and different positions within the galaxy and with respect to the spiral arms. In Section 3.5, we discuss the results and give an outlook to the application of RoadMapping to *Gaia* data.

Attribution. This chapter is based on the published and peer-reviewed paper by Trick et al. (2017). My main accomplishments in this paper are:

- Selecting a suite of 22 data sets of stellar phase-space coordinates from a N-body spiral galaxy simulation, covering a wide range of cases with respect to location within the galaxy and size of the survey volume.
- Applying RoadMapping successfully to all 22 data sets, demonstrating how well RoadMapping performs even if several modeling assumptions break down at the same time, including the fundamental assumption of axisymmetry in the model.
- Developing all diagnostic tools used in this chapter to quantify the performance quality of RoadMapping.
- Creating all tests, analyses, text, and figures in this chapter.

My collaborators contributed the following:

- **Hans-Walter Rix (MPIA)** suggested the project, the use of the specific N-body simulation in this work, and the simple outlier model for the likelihood.
- **Elena D'Onghia (University of Wisconsin)** provided the snapshots of the N-body simulation D'Onghia et al. (2013).
- **Jo Bovy (University of Toronto)**, as well as the other co-authors, gave valuable feedback on the work, including some suggestions for diagnostic plots, the structure and language of the paper.

In addition, I thank **Stephen Pardy (University of Wisconsin)** for helpful advice on handling N-body simulations.

3.2 Data from a Galaxy Simulation

RoadMapping requires 6D phase-space coordinates $(\boldsymbol{x}_i, \boldsymbol{v}_i)$ for a large set of stars that move independently in a collisionless galactic potential. If we want to test RoadMapping on a simulated galaxy, it is most convenient to apply it to an N-body simulation with a huge number of low-mass "star" particles. In that way, we can directly take the positions and the motions of individual particles as independent tracers of the potential, just as with the stars in the MW, without having to use an error-prone prescription to turn a single particle into many stars. The high-resolution simulations with its millions of particles by D'Onghia et al. (2013) satisfy this requirement.

3.2.1 Description of the Galaxy Simulation Snapshot

The high-resolution N-body simulation snapshot of a disk galaxy by D'Onghia et al. (2013), which we use in this chapter, was carried out with the `GADGET-3` code, and set up in the manner described in Springel et al. (2005). In this simulation, overdensities with properties similar to giant molecular clouds induced four prominent spiral arms—and therefore a non-axisymmetric substructure—via the swing amplification mechanism. This galaxy simulation was also investigated by D'Onghia et al. (2013; their Figure 8, top left panel) and D'Onghia (2015; their Figure 2, top left panel). For details, see D'Onghia et al. (2013). Here, we summarize the essential characteristics.

The simulation has a gravitationally evolving stellar disk within a static/rigid analytic Dark Matter (DM) halo.

The analytic halo follows a Hernquist (1990) profile

$$\rho_{\mathrm{DM}}(r) = \frac{M_{\mathrm{DM}}}{2\pi} \frac{a_{\mathrm{halo}}}{r(r + a_{\mathrm{halo}})^3} \tag{3.1}$$

with total halo mass $M_{\mathrm{DM}} = 9.5 \times 10^{11}$ $\mathrm{M}_\odot$ and scale length $a_{\mathrm{halo}} = 29$ kpc.

The disk consists of 10^8 "disk star" particles, each having a mass of ~ 370 $\mathrm{M}_\odot$, and 1000 "giant molecular cloud" particles with masses of $\sim 9.5 \times 10^5$ $\mathrm{M}_\odot$. The initial vertical mass distribution of the stars in the disk is specified by the profile of an isothermal sheet with a radially constant scale height $z_{\mathrm{s,init}}$, i.e.,

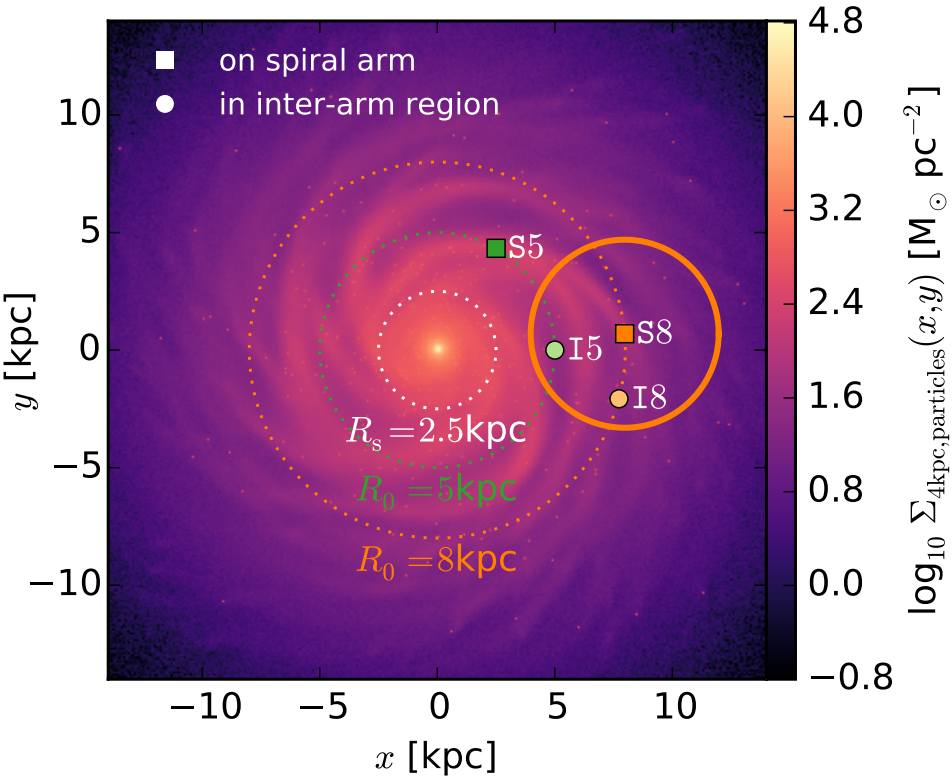
$$\rho_*(R, z) = \frac{M_*}{4\pi z_{\mathrm{s,init}} R_{\mathrm{s}}^2} \operatorname{sech}^2\left(\frac{z}{z_{\mathrm{s,init}}}\right) \exp\left(-\frac{R}{R_{\mathrm{s}}}\right), \tag{3.2}$$

with a total disk mass of $M_* = 0.04 M_{\mathrm{DM}} = 3.8 \times 10^{10}$ $\mathrm{M}_\odot$. The scale length R_{s} is assumed to be 2.5 kpc and $z_{\mathrm{s,init}} = 0.1 R_{\mathrm{s}}$. In this model, the disk fraction within 2.2 scale lengths is 50% of the total mass, leading to a formation of approximately four arms (D'Onghia 2015) (see Figure 3.2(b)).

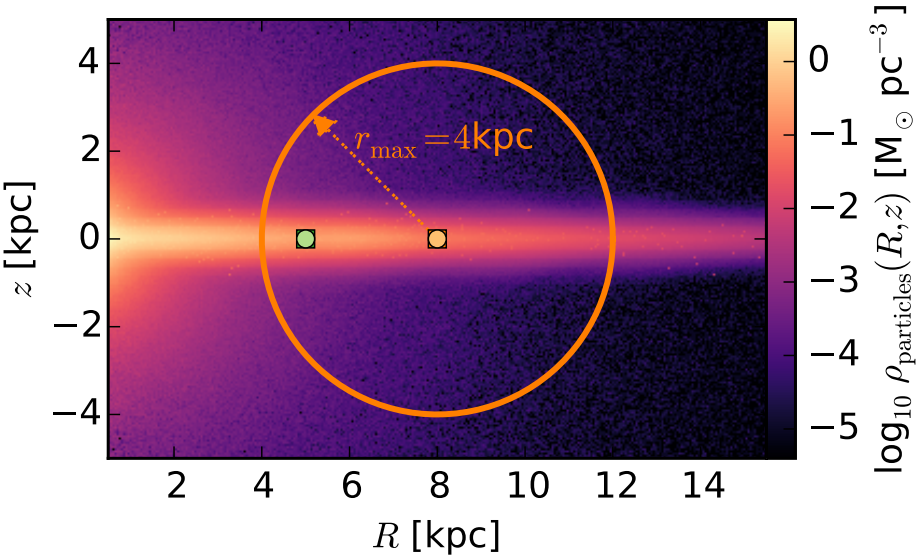
The bulge consists of 10^7 "bulge star" particles with masses of ~ 950 $\mathrm{M}_\odot$ and they are distributed following a spherical Hernquist profile analogous to Equation (3.1), with a total mass of $M_{\mathrm{bulge}} = 0.01 M_{\mathrm{DM}} = 9.5 \times 10^9$ $\mathrm{M}_\odot$ and scale length of $a_{\mathrm{bulge}} = 0.1 R_{\mathrm{s}} = 0.25$ kpc.

The initial velocity setup of the "disk star" particles assumes for simplicity Gaussian velocity dispersion profiles (Springel et al. 2005).

The simulation snapshot, which we are using in this work, has evolved from these initial conditions in isolation for ~ 250 Myr, which corresponds to approximately one orbital period at $R \sim 8$ kpc. The mass density of simulation particles (without the DM halo) at this snapshot time is shown in Figure 3.1. The "molecular cloud perturbers", which caused the formation of the four pronounced spiral arms, can be seen in Figure 3.1 as small overdensities in the disk. The spherical bulge and very flattened disk are shown in Figure 3.1(b).



(a) Surface mass density of particles.



(b) Mass density of particles.

Figure 3.1: Simulation snapshot by D'Onghia et al. (2013). Shown are the surface mass density (in the (x, y)-plane, panel 3.1(a)) and mass density (in the (R, z)-plane, panel 3.1(b)) of the "star" particles belonging to disk, bulge, and giant molecular clouds (the DM halo in this simulation is static and analytic and not shown here). Overplotted are the disk's scale length $R_s = 2.5$ kpc (see Section 3.2.1) and the radii at which we center our test survey volumes in this investigation, $R_0 = 8$ kpc and $R_0 = 5$ kpc (see Section 3.2.2). The centers of the different survey volumes are marked with a square, if the survey volume is centered on a spiral arm (`S8` and `S5`), or with a circle, if the volume is centered on an inter-arm region (`I8` and `I5`). The coordinates are summarized in Table 3.1. The orange circle with a radius of $r_{\mathrm{max}} = 4$ kpc marks the survey volume in which we conduct the analysis discussed in detail in Section 3.4.1.

Table 3.1. Vantage Points for Survey Volumes within the Galaxy Simulation Snapshot

Name	Position	R_0 [kpc]	ϕ_0 [°]	Legend
`S8`	on spiral arm	8	5	
`I8`	in inter-arm region	8	-15	
`S5`	on spiral arm	5	60	
`I5`	in inter-arm region	5	0	

Note. — All spherical volumes have a radius $r_{\max}$ and are centered on $z_0 = 0$ in the plane of the disk at the position given in this table. ϕ_0 is measured counter-clockwise from the positive x-coordinate axis.

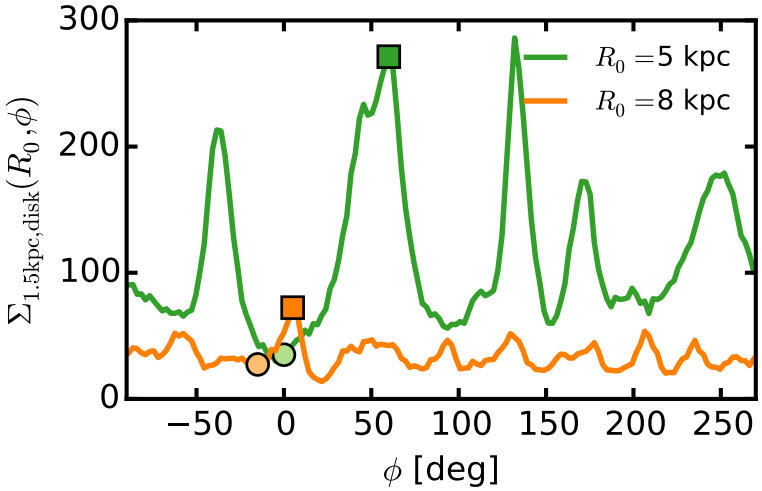
We have confirmed that the gravitational center of the particles corresponds to the coordinate origin.

3.2.2 Data Selection and Survey Volume

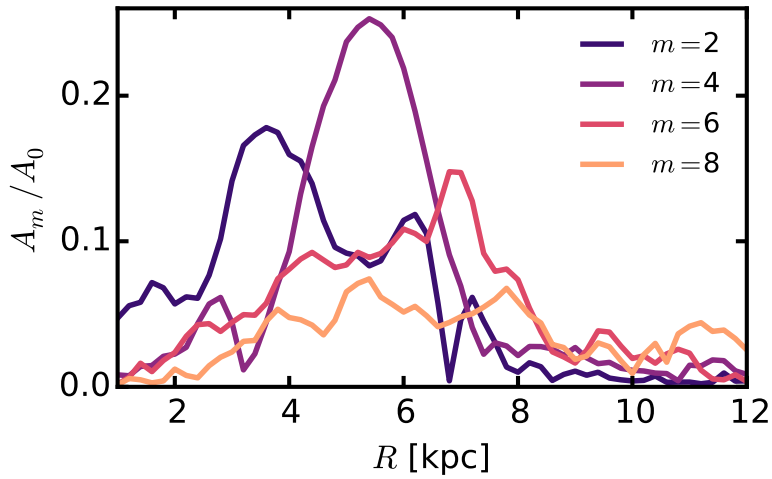
The selection function of all-sky surveys like *Gaia*, that are only limited by the brightness of the tracers, are contiguous and—when ignoring anisotropic effects like dust obscuration—spherical in shape. For simplicity, we will use spherical survey volumes centered on different vantage points, and with sharp edges at a distance $r_{\max}$ around it (see Equation (3.8)), which corresponds to a magnitude cut for stellar tracers all having the same luminosity.
Figure 3.1 illustrates the different survey volume positions analyzed in this study. We selected volumes with $r_{\max} = [0.5, 1, 2, 3, 4, 5]$ kpc centered on a spiral arm (`S`) and on an inter-arm region (`I`) at both the equivalent of the solar radius, $R_0 = 8$ kpc, and at $R_0 = 5$ kpc, where the disk strongly dominates (see Figure 3.3), and the spiral arms are more pronounced than at $R_0 = 8$ kpc (see Figure 3.2). The exact positions of the vantage points `S8`, `I8`, `S5`, and `I5` are summarized in Table 3.1.
From within each volume, we drew $N_* = 20,000$ random "disk star" particles, and used their phase-space positions $(\boldsymbol{x}_i, \boldsymbol{v}_i)$ within the simulated galaxy's rest-frame as data.
To make the data sample more realistic, one would actually have to add measurement uncertainties, especially to the distances from the survey volume's central vantage point and the proper motions measured from there. We decided not to include measurement uncertainties. First, their effect on RoadMapping modeling has already been investigated in Chapter 2, Sections 2.2.9 and 2.3.4, and we found that the measurement uncertainties of the last data release of *Gaia* should be small enough to not significantly disturb the modeling. Second, in this study, we want to isolate and investigate the deviations of the data from axisymmetry and the assumed potential and DF model independently of other effects.

3.2.3 True Symmetrized Potential

For a galaxy with pronounced spiral arms, an axisymmetric model matter distribution per se cannot reproduce the true matter distribution globally. We therefore obtain an "overall best-fit symmetrized" potential model from the distribution of particles to be able (1) to quantify the non-axisymmetries in the simulation snapshot better and (2) to compare how close our axisymmetric RoadMapping results can get to it.
We derive this model by fitting axisymmetric analytical functions to the density distribution of each of the galaxy components' particles. The bulge and halo follow Hernquist profiles by construction (see Section 3.2.1).



(a) Surface density along ϕ at $R_0 = [5, 8]$ kpc.



(b) Spiral strength from Fourier mode analysis.

Figure 3.2: Demonstrating the spiral arm strength at different radii. Panel 3.2(a) shows the surface density along the azimuth angle at the radii R_0 on which we center our survey volumes. We also mark the corresponding ϕ_0 from Table 3.1. The difference between the surface density at `S5` (■) and `I5` (●) is 200% of the mean surface density at R_0; for `S8` (■) and `I8` (●) the difference is 130%. Panel 3.2(b) shows the Fourier model amplitudes for $m = 2, 4, 6, 8$ calculated as $A_m/A_0 = |\sum_l M_l \exp(im\phi_l)| / \sum_l M_l$ for all disk particles at a given radius with mass M_l and azimuth position ϕ_l. As can be seen, the simulation has overall four strong spiral arms, dominating between $R = 4$ kpc and $R = 7$ kpc. Inside of that there are two, and outside of that six or more arms. As the particle density increases with smaller radius, most tracers in the analysis will come from regions with only a few strong spiral arms.

The disk in this simulation snapshot deviates from its initial conditions in Equation (3.2): after 250 Myr pronounced spiral arms have formed, also causing some in-plane heating. Except of an overdensity around $R \sim 6$ kpc (see Figure 3.6(d) in Section 3.4.1.2), the overall radial surface density profile (i.e., the azimuthal average) did not change by much, and appears smooth and exponential. We therefore chose a double exponential disk model to fit the particle distribution in the disk. The fit assumes the total disk mass to be known and to be equal to the total mass of all disk particles. The best-fit double exponential disk profile is found by maximizing the likelihood for all disk particles to be drawn from this axisymmetric density profile. In this way, the fit does not depend on binning choices and is driven by the number and location of the

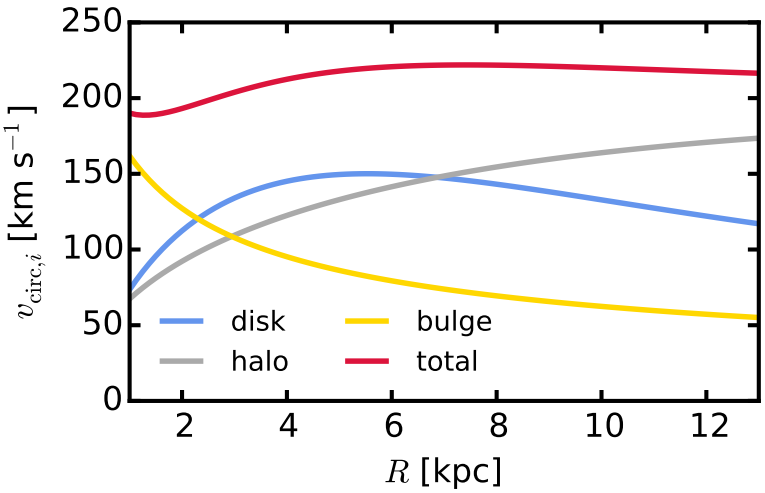


Figure 3.3: Circular velocity curve of the `DEHH-Pot`, i.e., the symmetrized best fit to the N-body simulation, and its disk, halo, and bulge components. The rotational support at 2.2 scale lengths is $(v_{\mathrm{circ,disk}}/v_{\mathrm{circ,total}})^2 \sim 47\%$. This demonstrates that the simulation is a disk-dominated spiral galaxy.

Table 3.2. Best-fit Parameters of the `DEHH-Pot`.

Potential parameter		Best-fit value
Circular velocity	$v_{\mathrm{circ}}(R_\odot)$	222 km s^{-1}
Disk scale length	R_{s}	2.5 kpc
Disk scale height	z_{s}	0.17 kpc
Halo fraction	f_{halo}	0.54
Halo scale length	a_{halo}	29 kpc
Bulge mass	M_{bulge}	0.95×10^{10} M$_\odot$
Bulge scale length	a_{bulge}	0.25 kpc

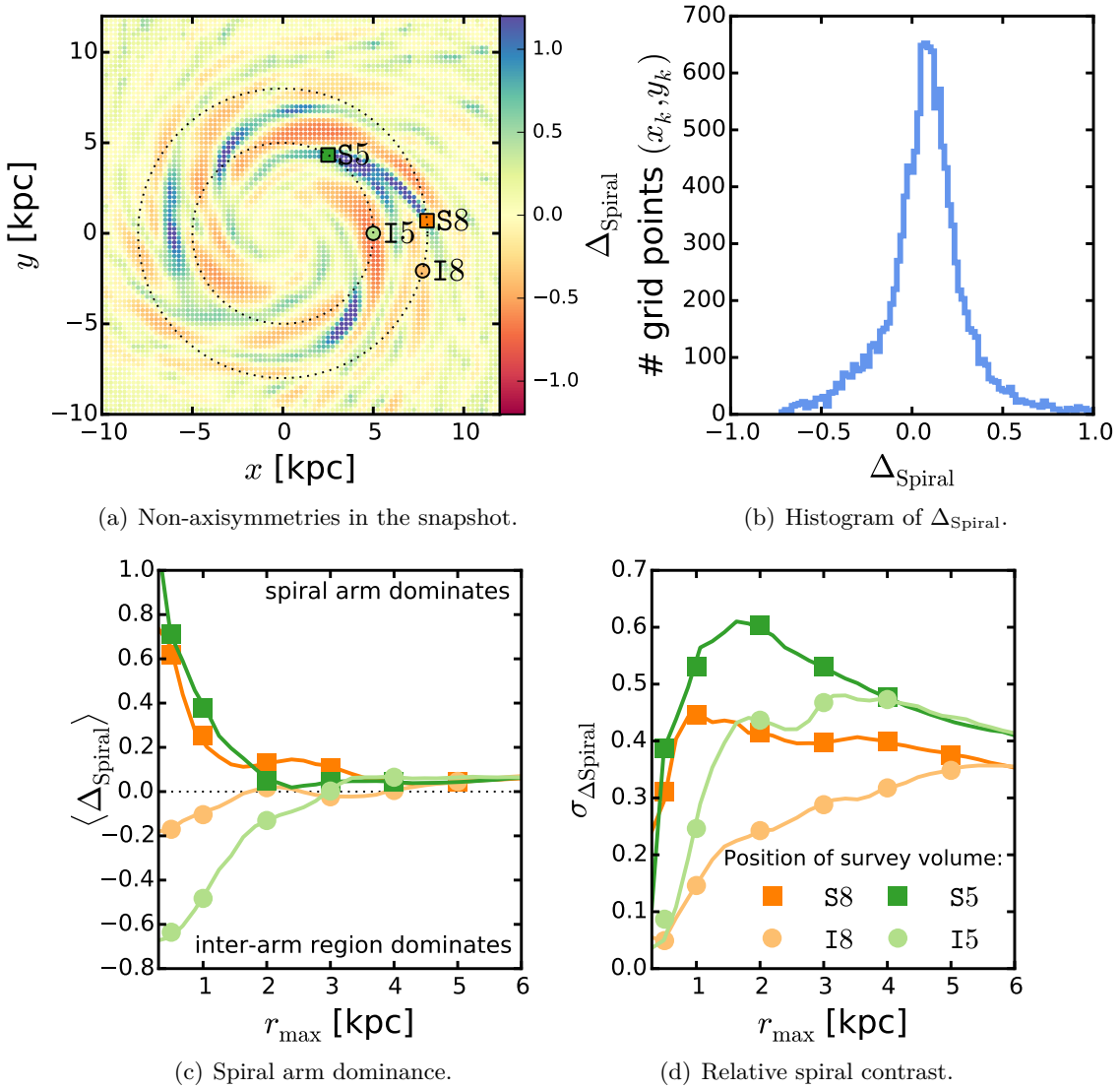
Note. — The `DEHH-Pot` is introduced in Section 3.2.3, and we use it as the global best-fit symmetrized potential model for the simulated galaxy. The halo fraction, f_{halo}, and circular velocity at the "solar" radius, $v_{\mathrm{circ}}(R_\odot)$, which scales the total mass of the model, are defined in Equations (3.18) and (3.19), with $R_\odot = 8$ kpc.

stars—analogous to our RoadMapping procedure.

The best-fit parameters for this reference potential, which we will refer to as the `DEHH-Pot` (Double-Exponential disk + Hernquist halo + Hernquist bulge) in the remainder of this chapter, are given in Table 3.2. As can be seen in Figure 3.6 in Section 3.4.1.2 below, the `DEHH-Pot` fits the overall true density distribution very well, with the exception of $z \sim 0$, where the particle distribution is not as cuspy as the exponential disk. Figure 3.3 shows the circular velocity curve of the `DEHH-Pot`, and its decomposition into disk, halo, and bulge contribution. The disk clearly dominates between $R \sim 2$ kpc and $R \sim 7$ kpc.

3.2.4 Quantifying the Strength of Spiral Arms

Depending on the size and position of the survey volume, spiral arms and inter-arm regions dominate the stellar distribution within the volume to different degrees. To quantify the strength



(a) Non-axisymmetries in the snapshot.

(b) Histogram of Δ_{Spiral}.

(c) Spiral arm dominance.

(d) Relative spiral contrast.

Figure 3.4: Dominance and contrast of the spiral arms. Panel 3.4(a) shows the local spiral strength Δ_{Spiral} (calculated according to Equation (3.3) as described in Section 3.2.4) at regular grid points (x_k, y_k) with bin width 0.25 kpc. Marked are the centroids of the four test survey volumes of this study analogous to Figure 3.1. The histogram in panel 3.4(b) demonstrates the number of different Δ_{Spiral} values in the region $x, y \in [-14, 14]$ kpc. The panels 3.4(c) and 3.4(d) then show the dominance and relative contrast of spiral arms and inter-arm regions within each survey volume, depending on the volumes' size, r_{max}, and position (color-coded). As measure for the spiral dominance, we use the mean $\langle \Delta_{\mathrm{Spiral}} \rangle$, and for the relative spiral contrast the standard deviation $\sigma_{\Delta\mathrm{Spiral}}$, calculated on the basis of all Δ_{Spiral} measurements within the given survey volume. We chose two volumes in which the spiral arms dominate, and two in which an inter-arm region dominates. The dominance and contrast of spiral arms and inter-arm regions is stronger at $R_0 = 5$ kpc than at $R_0 = 8$ kpc. Also, inter-arm regions appear larger and smoother than spiral arms, as already inside a small volume centered on a spiral arm the contrast is quite large. The larger the volume the more the overall effect of spiral arms and inter-arm regions average out.

of the spiral arms, we introduce the quantity

$$\Delta_{\mathrm{Spiral}}(x_k, y_k) \equiv \frac{\Sigma_{1.5\mathrm{kpc,disk},T}(x_k, y_k)}{\Sigma_{1.5\mathrm{kpc,disk},S}(x_k, y_k)} - 1 \tag{3.3}$$

where $\Sigma_{1.5\mathrm{kpc,disk},\alpha}$ is the true surface density of the disk component of the simulation snapshot ($\alpha = T$ for "true"), or of the symmetrized snapshot model `DEHH-Pot` in Section 3.2.3 ($\alpha = S$ for "symmetrized"),

$$\Sigma_{1.5\mathrm{kpc,disk},\alpha}(x_k, y_k) \equiv \int_{-1.5\ \mathrm{kpc}}^{1.5\ \mathrm{kpc}} \rho_{\mathrm{disk},\alpha}(x_k, y_k, z)\ \mathrm{d}z. \tag{3.4}$$

(x_k, y_k) are the coordinates of regular grid points with a spacing of $\delta = 0.25$ kpc.[13] ($x_c = R_{0,c} \cdot \cos\phi_{0,c}, y_c = R_{0,c} \cdot \sin\phi_{0,c}, z_c = 0$) is the position of the survey volume's center within the simulation, with $c \in \{\mathtt{S8}, \mathtt{I8}, \mathtt{S5}, \mathtt{I5}\}$ and $(R_{0,c}, \phi_{0,c})$ given in Table 3.1. We consider all $n \simeq \pi r_{\mathrm{max}}^2/\delta^2$ values of $\Delta_{\mathrm{Spiral}}(x_k, y_k)$ inside a given survey volume of radius r_{max} around position c and calculate the mean and standard deviation,

$$\langle\Delta_{\mathrm{Spiral}}\rangle \equiv \frac{1}{n}\sum_{k=1}^{n}\Delta_{\mathrm{Spiral}}(x_k, y_k) \tag{3.5}$$

$$\sigma_{\Delta\mathrm{Spiral}} \equiv \sqrt{\frac{1}{n}\sum_{k=1}^{n}\left[\Delta_{\mathrm{Spiral}}(x_k, y_k) - \langle\Delta_{\mathrm{Spiral}}\rangle\right]^2} \tag{3.6}$$

$$\text{with} \quad (x_k - x_c)^2 + (y_k - y_c)^2 \leq r_{\mathrm{max}}^2. \tag{3.7}$$

These quantities tell us if and how much a spiral arm or an inter-arm region dominates the survey volume ($\langle\Delta_{\mathrm{Spiral}}\rangle > 0$ for spiral arms, $\langle\Delta_{\mathrm{Spiral}}\rangle < 0$ for inter-arm regions) and how large the relative contrast between spiral arms and inter-arm regions is ($\sigma_{\Delta\mathrm{Spiral}}$). For example, volumes will have a smaller relative spiral contrast $\sigma_{\Delta\mathrm{Spiral}}$, if they are either small and sitting completely within an inter-arm region, or if they are large volumes that contain—in addition to some spiral arms and depleted inter-arm regions—large areas of unperturbed disk.
Figure 3.4 shows Δ_{Spiral} as a function of (x_k, y_k), a histogram over all $\Delta_{\mathrm{Spiral},k}$ within the galaxy, and $\langle\Delta_{\mathrm{Spiral}}\rangle$ and $\sigma_{\Delta\mathrm{Spiral}}$ calculated for all test survey volumes in this chapter (see Section 3.2.2), depending on position and size.[14]

3.2.5 Comparison of the N-Body Simulation to the Milky Way

The N-body simulation in this work is not a perfect match to the MW. However, to be suitable for this study it is most important that it satisfies our main requirements: it is a disk-dominated spiral galaxy with a very high number of star particles and strong spiral arms. To set the context, we discuss in the following the differences between the simulation at hand and what we know about the MW (see Section 1.2 and Table 1.1).

Bar. The MW has a central bar; our simulation does not. The Galactic bar can introduce non-axisymmetries around the co-rotation and Lindblad resonances (see Section 1.4.5 and references therein). Bar effects are consequently not part of this study and remain a possible source of

[13] We average the particle surface density of the true simulation potential over area element sizes of $\delta \times \delta$ around (x_k, y_k), when calculating $\Sigma_{1.5\mathrm{kpc,disk},T}(x_k, y_k)$.

[14] When considering the whole galaxy or a large survey volume, $\langle\Delta_{\mathrm{Spiral}}\rangle$ in Figures 3.4(b) and 3.4(c) is not exactly at 0, but slightly larger (< 0.05). We account this small bias to the different functional forms of the `DEHH-Pot` and the initial axisymmetric disk in the simulation, Equation (3.2). This bias will, however, not affect our results.

uncertainty in RoadMapping. We suspect, however, that its influence is modest outside of the bar region, except close to the outer Lindblad resonance.

Mass components. The MW's dark halo is estimated to have a mass $M_{\mathrm{DM,200,MW}} \approx 10^{12}$ $\mathrm{M}_\odot$ (Bland-Hawthorn & Gerhard 2016); the simulation's halo is only slightly less massive with $M_{\mathrm{DM,200}} \approx 7 \times 10^{11}$ $\mathrm{M}_\odot$. The stellar bulge mass of $M_{\mathrm{bulge}} = 9.5 \times 10^9$ $\mathrm{M}_\odot$ in this simulation is a bit smaller than the estimated stellar mass of the MW bulge with $M_{\mathrm{bulge,MW}} = (1.4 - 1.7) \times 10^{10}$ $\mathrm{M}_\odot$ (Portail et al. 2015). The total stellar mass of the MW is estimated to be $M_{\mathrm{stars,MW}} = (5 \pm 1) \times 10^{10}$ $\mathrm{M}_\odot$ (Bovy & Rix 2013; Bland-Hawthorn & Gerhard 2016), consistent with the total baryonic mass in the simulation of $M_{\mathrm{stars}} = 4.75 \times 10^{10}$ $\mathrm{M}_\odot$. The fraction of bulge mass to total stellar mass of the MW is $M_{\mathrm{bulge,MW}}/M_{\mathrm{stars,MW}} = 0.3 \pm 0.06$; in our simulation, it is 0.25.

Disk. The best estimate for the MW's thin disk scale length from combining several measurements in the literature is $R_{\mathrm{s}} = 2.6 \pm 0.5$ kpc (Bland-Hawthorn & Gerhard 2016). Bovy & Rix (2013), for example, found a stellar disk scale length of $R_{\mathrm{s}} = 2.15 \pm 0.14$ kpc. The disk of this N-body simulation has a similar scale length, $R_{\mathrm{s}} = 2.5$ kpc. The stellar disk, however, is thinner than in the MW. Jurić et al. (2008) found scale heights 300 pc and 900 pc (with 20% uncertainty) for the thin and thick disk of the MW, respectively; Bovy et al. (2012d), who considered the disk to be a superposition of many exponential MAPs, measured scale heights from ≈ 200 pc up to 1 kpc, continuously increasing with the age of the sub-population. In our simulation, there is only a very thin stellar disk component with a scale height of $z_{\mathrm{s}} = 170$ pc, and no gas and thick disk component as compared to the MW. This discrepancy does, however, not affect the objective of this study. The thick disk has a much higher velocity dispersion and is less prone to spiral perturbations. When we will apply RoadMapping to real MW data, the gas disk can be included as an additional component in the mass model, analogous to Bovy & Rix (2013), and if needed also the thick disk. By slicing data according to MAPs, the thick disk will be accounted for implicitly in the tracer selection.

Disk fraction. The strength of the perturbations in the disk depends on the disk fraction (e.g., D'Onghia 2015). The total disk mass in the simulation might be only 4% of the halo mass, but within $2.2R_{\mathrm{s}}$ it is already 50% of the total mass. It is still under debate if the MW disk is maximal (see Sections 1.1.5 and 1.3.3). Bovy & Rix (2013) found, for example, a maximum disk with rotational support $(v_{\mathrm{circ,disk}}/v_{\mathrm{circ,total}})^2 = (69 \pm 6)\%$ at $2.2R_{\mathrm{s}}$ with $R_{\mathrm{s}} = 2.15 \pm 0.14$ kpc. The `DEHH-Pot` (and therefore our N-body model) has $(v_{\mathrm{circ,disk}}/v_{\mathrm{circ,total}})^2 \approx 47\%$ at $2.2R_{\mathrm{s}}$ and is therefore slightly sub-maximal (see Figure 3.3). The decomposed rotation curve in Figure 3.3 is qualitatively similar to the MW models by McMillan (2011; their Figure 5) and by Barros et al. (2016; their Figure 5, left panels, model MI). Their disks dominate between $R \approx 2.5 - 8.5$ kpc and $R \approx 5 - 9.5$ kpc, respectively, while our simulation's disk dominates between $R \approx 2 - 7$ kpc. We account for this slight difference by also drawing mock data sets from regions around $R_0 = 5$ kpc (see Section 3.2.2).

Number of spiral arms. If the MW is a two-armed or a four-armed spiral galaxy is still not known (see Section 1.2.5). Our simulation is overall a four-armed spiral galaxy, with the mode $m = 4$ dominating between $R = 4 - 7$ kpc, $m = 2$ at smaller and $m \geq 6$ at larger radii (see Figure 3.2(b)).

Strength of spiral arms. As demonstrated in Figure 3.2(a), the spiral arms introduce strong peak-to-peak differences in the stellar surface density (e.g., $\sim 200\%$ at $R = 5$ kpc). The disk dominates inside $R = 8$ kpc (see Figure 3.3), and in Section 3.4.1.2, Figure 3.7, we will see that

the spiral arms introduce relative perturbations in the total gravitational forces of up to 30%. Figure 3.5(c), which we will discuss in Section 3.4.1.1, suggests an excess of stars with radial velocities of up to 50 km s^{-1} in our simulated galaxy as compared to an axisymmetric model. This can be compared to Reid et al. (2014), who measured, for example, that typical peculiar non-circular motions in the MW spiral arms were around $10-20$ km s^{-1} for $R \gtrsim 4$ kpc. Siebert et al. (2012) and Bovy et al. (2015) found velocity fluctuations of the same order in the solar neighborhood. We therefore expect the strength of the perturbation to the total potential due to the spiral arms in this simulation—especially inside $R = 8$ kpc where most of our mock data is drawn from—to be similar or even stronger than those in the MW.

3.3 Ingredients for the RoadMapping Modeling

In this section, we summarize the mathematical ingredients of RoadMapping, and motivate the DF and potential model that we are going to fit to the data. RoadMapping makes extensive use of the `galpy` python library by Bovy (2015)[15]. For full details on the RoadMapping machinery, see also Chapter 2, Sections 2.2.7-2.2.10.

3.3.1 Likelihood

As already laid out in Section 3.2.2, we use as data the 6D $(\boldsymbol{x}_i, \boldsymbol{v}_i)$ coordinates of N_* stars within a spherical survey volume. The corresponding, purely spatial selection function $\mathrm{SF}(\boldsymbol{x})$ is

$$\mathrm{SF}(\boldsymbol{x}) \equiv \begin{cases} 1 & \text{if } |\boldsymbol{x} - \boldsymbol{x}_0| \leq r_{\max} \\ 0 & \text{otherwise} \end{cases}, \tag{3.8}$$

with $\boldsymbol{x}_0 = (R_0, \phi_0, z_0 = 0)$ from Table 3.1.
Given a parametrized axisymmetric potential model $\Phi(R, z)$ with parameters p_Φ, the probability that the ith star is on an orbit with the actions

$$\boldsymbol{J}_i \equiv \boldsymbol{J}[\boldsymbol{x}_i, \boldsymbol{v}_i \mid p_\Phi], \tag{3.9}$$

is proportional to the given orbit distribution function $\mathrm{DF}(\boldsymbol{J})$ with parameters p_{DF},

$$\mathrm{DF}(\boldsymbol{J}_i \mid p_{\mathrm{DF}}) \equiv \mathrm{DF}(\boldsymbol{x}_i, \boldsymbol{v}_i \mid p_\Phi, p_{\mathrm{DF}}). \tag{3.10}$$

The joint likelihood of a star being within the survey volume and on a given orbit is therefore

$$\begin{aligned} \mathscr{L}_i &\equiv \mathscr{L}(\boldsymbol{x}_i, \boldsymbol{v}_i \mid p_\Phi, p_{\mathrm{DF}}) \\ &= \frac{\mathrm{DF}(\boldsymbol{x}_i, \boldsymbol{v}_i \mid p_\phi, p_{\mathrm{DF}}) \cdot \mathrm{SF}(\boldsymbol{x_i})}{\int \mathrm{DF}(\boldsymbol{x}, \boldsymbol{v} \mid p_\Phi, p_{\mathrm{DF}}) \cdot \mathrm{SF}(\boldsymbol{x}) \ \mathrm{d}^3x\, \mathrm{d}^3v}. \end{aligned} \tag{3.11}$$

The details of how we numerically evaluate the likelihood normalization to a sufficiently high precision are discussed in Section 2.2.8.[16]
In the scenario considered in this chapter, it can happen that there are a few (~ 1 in 20,000) stars entering the catalog that are for some reason on rather extreme orbits, e.g., moving radially directly toward the center. These kinds of orbits do not belong to the set of orbits that we classically expect to make up an overall smooth galactic disk. To avoid such single stars with

[15] The `galpy` python package by Bovy (2015) can be downloaded from `http://github.com/jobovy/galpy`.
[16] In the terminology of Chapter 2, we use the high numerical accuracy of $N_x = 20$, $N_v = 28$, $n_\sigma = 5.5$ to calculate the likelihood normalization, or in other words, to evaluate the spatial and velocity integrals over the qDF within the survey volume.

very low likelihoods having a strong impact on the modeling, we employ here a simple strategy to ensure a robust likelihood,

$$\mathscr{L}_i \longrightarrow \max\left(\mathscr{L}_i, \epsilon \times \mathrm{median}(\mathscr{L})\right), \tag{3.12}$$

where $\epsilon = 0.001$ for $N_* = 20,000$ stars and median($\mathscr{L}$) is the median of all the N_* stellar likelihoods $\mathscr{L}_i$ with the given p_Φ and p_{DF}. This robust likelihood method was not used in Chapter 2. For future applications to real MW data, an outlier model similar to the one in Bovy & Rix (2013) could be added: a stellar outlier distribution with constant spatial number density and velocities following a broad Gaussian.
Following Chapter 2, we assume for now uninformative flat priors on the model parameters p_Φ and p_{DF} and find the maximum and width of the Probability Distribution Function $\mathrm{PDF}(p_\Phi, p_{\mathrm{DF}} \mid \mathrm{data}) \propto \prod_{i=1}^{N_*} \mathscr{L}_i \times \mathrm{prior}(p_\Phi, p_{\mathrm{DF}})$ using a nested-grid approach and then explore the full shape of the PDF using a MCMC[17]. Full details on this procedure are given in Section 2.2.10.

3.3.2 Distribution Function Model

The most simple action-based orbit DF is the quasi-isothermal Distribution Function (qDF) introduced by Binney (2010) and Binney & McMillan (2011), which has been a successful ingredient in Chapter 2 and many disk modeling approaches (Bovy & Rix 2013; Piffl et al. 2014; Sanders & Binney 2015b). The exact functional form of the $\mathrm{qDF}(J_R, L_z, J_z \mid p_{\mathrm{DF}})$ is given, for example, in Binney & McMillan (2011), or in Equations (1.58)-(1.60).
The qDF is expressed in terms of actions, frequencies, and scaling profiles for the radial stellar tracer density $n(R_g)$, and velocity dispersion profiles $\sigma_z(R_g)$ and $\sigma_R(R_g)$. The latter are functions of the guiding-center radius R_g, i.e., the radius of a circular orbit with given angular momentum L_z in a given potential. We set the scaling profiles to

$$n(R_g \mid p_{\mathrm{DF}}) \propto \exp\left(-\frac{R_g}{h_R}\right) \tag{3.13}$$

$$\sigma_R(R_g \mid p_{\mathrm{DF}}) = \sigma_{R,0} \times \exp\left(-\frac{R_g - R_\odot}{h_{\sigma,R}}\right) \tag{3.14}$$

$$\sigma_z(R_g \mid p_{\mathrm{DF}}) = \sigma_{z,0} \times \exp\left(-\frac{R_g - R_\odot}{h_{\sigma,z}}\right). \tag{3.15}$$

The free model parameters of the qDF are

$$p_{\mathrm{DF}} \equiv \{\ln h_R, \ln \sigma_{R,0}, \ln \sigma_{z,0}, \ln h_{\sigma,R}, \ln h_{\sigma,z}\}. \tag{3.16}$$

In an axisymmetric potential superimposed with non-axisymmetric perturbations, the actions are not exact integrals of motion. However, if one simply considers action-angles as phase-space coordinates, a DF that is a function of the actions only might still be a good model for $\mathrm{DF}(\boldsymbol{x}, \boldsymbol{v}, t) = \mathrm{DF}(\boldsymbol{J}, \boldsymbol{\theta}, t)$ at a given point in time and if the system is well-mixed in phase.
We motivate the use of the qDF as specific action-based DF for the simulation snapshot in this chapter as follows. There is no stellar abundance or age information in the simulation. We therefore cannot define stellar sub-populations, as we normally would for the MW (see Section 1.5.3 and Bovy & Rix 2013). However, the disk of the galaxy simulation was originally set up as a single axisymmetric flattened particle population whose density decreases exponentially with radius (see Section 3.2.1). This is actually very similar to the stellar distribution generated by a single qDF (see, e.g., Ting et al. 2013). Because all particles in the disk have evolved for the

[17] We use the MCMC software `emcee` by Foreman-Mackey et al. (2013).

same ~ 250 Myr since its axisymmetric setup, we can consider the disk essentially as a mono-age population. All of this motivates us therefore to use one single qDF to model the whole disk. Locally, the current particle distribution in the snapshot at hand might be dominated by non-axisymmetries, which evolved later in the simulation. We have no indication if for small survey volumes the qDF is still a good model for the data. We will use it anyway—to see how far we can get with the simplest model possible and to test if actions are still informative in this case.

3.3.3 Potential Model

In all RoadMapping analyses in this chapter, we will fit an axisymmetric gravitational potential model consisting of a (fixed and known) Hernquist bulge, a free Hernquist halo and a free Miyamoto-Nagai disk (Miyamoto & Nagai 1975),

$$\Phi_{\rm disk}(R,z) = -\frac{GM}{\sqrt{R^2 + (a_{\rm disk} + \sqrt{z^2 + b_{\rm disk}^2})^2}}, \tag{3.17}$$

where $a_{\rm disk}$ and $b_{\rm disk}$ are the equivalents of a disk scale length and scale height. Using Hernquist profiles for halo and bulge is motivated by our knowledge of the snapshot galaxy, and we fix the bulge's total mass and scale length to the true values (see Section 3.2.1 and Table 3.3). As the bulge contribution to the total radial force at $R_\odot \equiv 8$ kpc is only $\sim 9-10\%$, this will not give the modeling an unfair advantage. The free model parameters of the halo are the halo scale length $a_{\rm halo}$ and the halo fraction, i.e., the relative halo-to-disk contribution to the radial force at $R_\odot$, defined as

$$f_{\rm halo} \equiv \left.\frac{F_{R,\rm halo}}{F_{R,\rm disk} + F_{R,\rm halo}}\right|_{\substack{R=R_\odot \\ z=0}}. \tag{3.18}$$

As a parameter that scales the total mass of the galaxy model, we use the circular velocity at the "solar" radius $R_\odot$,

$$v_{\rm circ}(R_\odot = 8\ {\rm kpc}) \equiv \left.\sqrt{R\frac{\partial\Phi}{\partial R}}\right|_{\substack{R=R_\odot \\ z=0}}. \tag{3.19}$$

The total set of free potential model parameters is therefore

$$p_\Phi \equiv \{v_{\rm circ}(R_\odot), a_{\rm disk}, b_{\rm disk}, a_{\rm halo}, f_{\rm halo}\}\,. \tag{3.20}$$

We will call this potential model the `MNHH-Pot` (Miyamoto-Nagai disk + Hernquist halo + Hernquist bulge) in the remainder of this chapter.
To estimate the stellar actions $\boldsymbol{J} = (J_R, L_z, J_z)$ in the axisymmetric `MNHH-Pot`, we use the *Stäckel fudge* algorithm by Binney (2012a) with fixed focal length $\Delta = 0.45$, and interpolate the actions on a grid (Binney 2012a; Bovy 2015). We made sure that the accuracy of the parameter estimates are not degraded by interpolation errors.[18]
Galaxy disks, in general, as well as the simulated disk in this work, have exponential radial density profiles. A single Miyamoto-Nagai disk is more massive at large radii than an exponential disk (see, e.g., Smith et al. 2015). By construction, the `DEHH-Pot` introduced in Section 3.2.3 is therefore better suited to reproduce the overall density distribution in the simulation than the `MNHH-Pot` (as we will see in Figure 3.6 in the next section). However, the closed form expression of the Miyamoto-Nagai potential in Equation (3.17) has the crucial advantage of allowing much faster force and therefore action calculations. In addition, by using a potential model of which we already know that it is not the optimal model for the galaxy's disk, we challenge RoadMapping

[18] For the action interpolation grid following Bovy (2015), we use $R_{\rm max} = 40$ kpc, $n_E = 70$, $n_\psi = 40$, $n_{L_z} = 50$ in their terminology.

Table 3.3. Best-fit `MNHH-Pot` and qDF Parameters as Recovered from the RoadMapping Analysis of a Survey Volume with $r_{\max} = 4$ kpc Centered on a Spiral Arm at $R_0 = 8$ kpc (Position `S8`).

Potential parameter		Best-fit value
Circular velocity at $R_\odot = 8$ kpc	$v_{\rm circ}(R_\odot)$	(223.0 ± 0.1) km s^{-1}
Miyamoto-Nagai disk scale length	$a_{\rm disk}$	$(3.62^{+0.06}_{-0.05})$ kpc
Miyamoto-Nagai disk scale height	$b_{\rm disk}$	(0.26 ± 0.02) kpc
Halo fraction at $R_\odot = 8$ kpc	$f_{\rm halo}$	(0.53 ± 0.02)
Halo scale length	$a_{\rm halo}$	(21 ± 2) kpc
Bulge mass	$M_{\rm bulge}$	0.95×10^{10} M$_\odot$ (fixed)
Bulge scale length	$a_{\rm bulge}$	0.25 kpc (fixed)
qDF tracer scale length	h_R	$(3.34^{+0.05}_{-0.04})$ kpc
qDF radial velocity dispersion	$\sigma_{R,0}$	(15.91 ± 0.08) km s^{-1}
qDF vertical velocity dispersion	$\sigma_{z,0}$	$(14.0^{+0.2}_{-0.1})$ km s^{-1}
qDF radial velocity dispersion scale length	$h_{\sigma,R}$	(4.6 ± 0.5) kpc
qDF vertical velocity dispersion scale length	$h_{\sigma,z}$	$(5.65^{+0.06}_{-0.07})$ kpc

Note. — The bulge mass and scale length were fixed in the analysis to their true values, see Sections 3.2.1 and 3.3.3.

even further.

3.4 Results

At the core of this chapter is a suite of 22 data sets consisting of the phase-space coordinates of stellar tracer particles, drawn from the spiral galaxy simulation snapshot introduced in Section 3.2.1. Each data set comes from a different survey volume within the galaxy's disk (see Section 3.2.2). We modeled all data sets with RoadMapping as described in Section 3.3, by fitting to it a single qDF (see Section 3.3.2) and the potential model `MNHH-Pot` (introduced in Section 3.3.3). This resulted in 22 independent measurements of the simulated galaxy's potential and DF.
We present our results in two steps. In Section 3.4.1, we look at one of these RoadMapping models in detail. In Section 3.4.2, we then compare all 22 RoadMapping results, and discuss their differences in the context of spiral arms.

3.4.1 An Axisymmetric Galaxy Model from RoadMapping

In this section, we will discuss all aspects of a RoadMapping model for one single data set. This data set has $N_* = 20,000$ stars that were drawn from the spherical volume with $r_{\max} = 4$ kpc centered on a spiral arm at the "solar" radius $R_0 = 8$ kpc. This volume is shown in orange in Figure 3.1 (position `S8`). We chose this volume because of its position centered on a smaller spiral arm, similar to our Sun being located in the Orion spiral arm. It is a bit larger than our conservative guess for the survey volume size for which we currently expect unbiased potential estimates from the final *Gaia* data release, $r_{\max} = 3$ kpc from the Sun (see Sections 2.4.2 and 3.5.3). However, improvements in RoadMapping with respect to the treatment of measurement errors might ultimately allow us to also model larger volumes. This example survey volume ranging from $R = 4 - 12$ kpc also has the advantage of being crossed by several spiral arms of different spiral strength (see Figures 3.1(a) and 3.2). In this section, our goal is now to investigate the ability of the best-fit RoadMapping model to serve as an overall axisymmetric model for the galaxy.

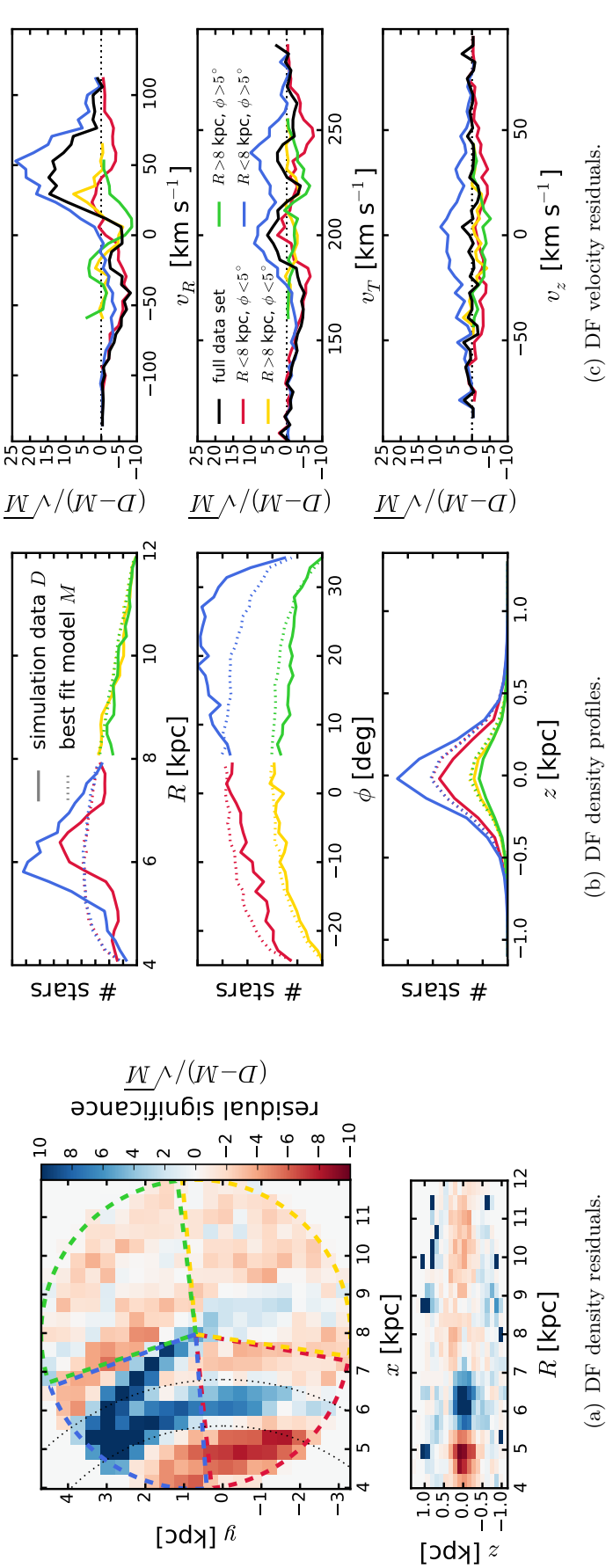
The parameters of the best-fit `MNHH-Pot` and qDF recovered with RoadMapping from this data set are summarized in Table 3.3. The circular velocity $v_{\rm circ}(R_\odot)$ and halo fraction $f_{\rm halo}$ are especially well recovered (compare to Table 3.2).

3.4.1.1 Recovering the Stellar Distribution

The RoadMapping fit itself takes place in action space. However, an important sanity check to decide if the fit was successful, is to test if the best-fit RoadMapping model (i.e., best-fit action-based DF in best-fit potential and in given selection function) generates a stellar distribution that reproduces the distribution of data points in observable phase-space, $(\boldsymbol{x}, \boldsymbol{v})$. This comparison is shown in Figure 3.5.
We note that the spiral arms introduce very strong non-axisymmetries in the data, both in the spatial and the velocity distribution (especially in v_R, where a significantly larger number of stars move outward than inward as compared to an axisymmetric model). We therefore compare the data and fit separately for different spatial regions, $R > 8$ kpc and $R < 8$ kpc, and $\phi < 5°$ and $\phi > 5°$. In the region where the spiral arm dominates (blue in Figure 3.5), the best-fit RoadMapping model is actually a very poor model. However, what the model underestimates in the spiral arm, it slightly overestimates in the other regions and is therefore indeed something like a good average model for the overall distribution. The region at $R > 8$ kpc and $\phi < 5°$ (yellow in Figure 3.5), where neither the spiral arm nor the inter-arm regions dominate strongly, is especially well described by the model.
Overall, the qDF appears to be a good model for unperturbed regions of the disk and averages over spiral arms.

3.4.1.2 Recovering the Gravitational Potential

As shown in the previous section, the best-fit RoadMapping model seems to reproduce the average stellar phase-space distribution quite well. But is the corresponding potential close to the true potential?
Figure 3.6 compares the true potential from the simulation snapshot (symmetrized by averaging over the whole $\Delta\phi = 2\pi$) and the axisymmetric reference `DEHH-Pot` from Table 3.2 with the best-fit `MNHH-Pot` from the RoadMapping analysis. In particular, Figure 3.6 illustrates the overall matter density distribution, the rotation curve and the surface density profile. Figure 3.7 compares the true and recovered (median) gravitational forces at the position of each star in the data set.
The recovery of density, surface density, and circular velocity curve is especially good in the region where most of the stars are located, around $R \sim 6$ kpc and in the plane of the disk. In large regions inside the survey volume, and even outside, the density is recovered to within 15%. The circular velocity curve is recovered to within 5%, which is also approximately the extent of perturbation that the spiral arms cause with respect to a smooth rotation curve. There is, however, a very small ($< 1.5\%$) underestimation of $v_{\rm circ}$ at larger radii. We suspect that this bias is introduced by the spiral arms (see the discussions in Sections 3.4.2.6 and 3.4.1.4). The overall surface density profile is a bit overestimated ($\sim 15\%$) at smaller radii; this is clearly due to the local spiral arm at $R \sim 6$ kpc with its higher surface density and many stars entering the analysis, which bias the result and which was to be expected. At larger radii ($R \sim 10 - 12$ kpc), the fit of the local density in the disk and surface density profile starts to flare due to the choice of the Miyamoto-Nagai disk family with its shallow profile (see Section 3.4.1.4). But again, where most of the stars are located, our RoadMapping model is a very good average model for the true galaxy.
The aspect of the potential to which the stellar orbits are actually sensitive is the gravitational forces. In Figure 3.7, we therefore compare the true force that each star in the data set feels (i.e.,

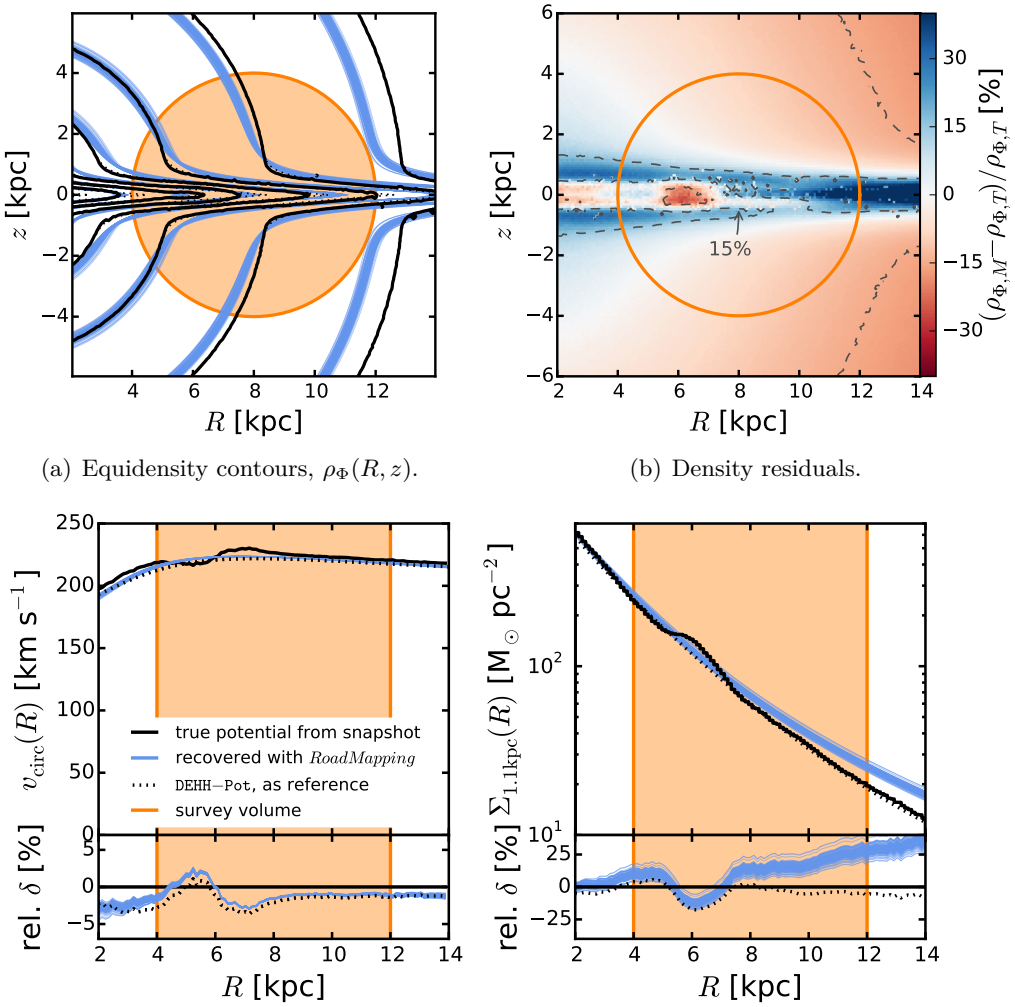


(a) DF density residuals.

(b) DF density profiles.

(c) DF velocity residuals.

Figure 3.5: Comparison of the true and best-fit stellar DF$(\boldsymbol{x}, \boldsymbol{v})$ in position-velocity space. The true DF$(\boldsymbol{x}, \boldsymbol{v})$ (i.e., the data D) is the distribution of all the stars in the data set drawn from the simulation snapshot (with $N_* = 20,000$ and $r_{\rm max} = 4$ kpc centered on position S8). The best-fit DF$(\boldsymbol{x}, \boldsymbol{v})$ (i.e., the model M) is generated by MC sampling of the best-fit qDF$(\boldsymbol{J})$ in the best-fit potential model from RoadMapping in Table 3.3, given the known selection function. Panel 3.5(a) shows the spatial density residual significance $(D - M)/\sqrt{M}$ of the projection to the (x, y) and (R, z) plane (the MC sampled M is the expected number of stars per bin, and $\sqrt{M}$ is the expected error due to Poisson statistics). In the (x, y) panel, the following regions are marked: $R < 8$ kpc, $\phi > 5°$ (blue), $R < 8$ kpc, $\phi < 5°$ (red), $R > 8$ kpc, $\phi > 5°$ (green), $R > 8$ kpc, $\phi < 5°$ (yellow). Panels 3.5(b) and 3.5(c) show the density profiles and velocity residual significance along each of the 6D phase-space coordinates separately for each of the four spatial regions. The blue region is very much dominated by the non-axisymmetric spiral arm. For the yellow region, the axisymmetric single-qDF model is a good description. Overall, the qDF is a good average axisymmetric model for the data. (In panel 3.5(a), we overplot the radii $R_{\rm spiral} \in [5.6, 6.8]$ kpc as black dotted lines to mark the approximate extent of the stronger spiral arm, to compare it with Figure 3.8.)



(a) Equidensity contours, $\rho_\Phi(R,z)$.

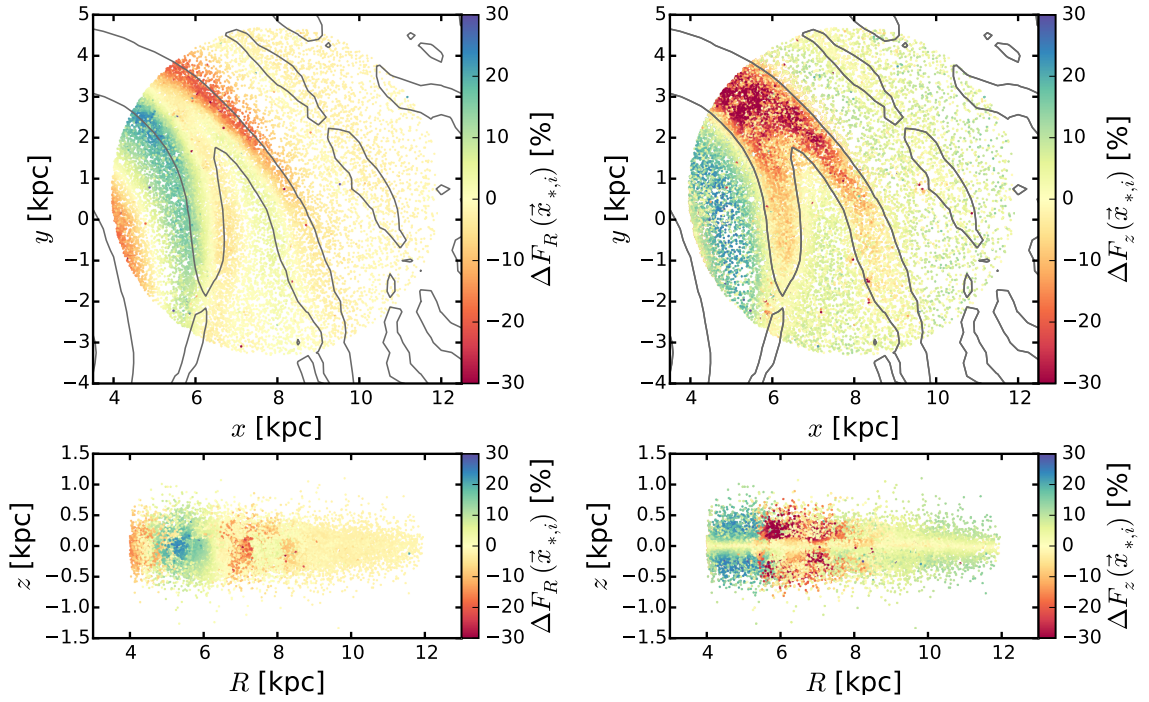
(b) Density residuals.

(c) Circular velocity curve.

(d) Surface density profile.

Figure 3.6: Comparison of the true and recovered gravitational potential Φ. Panel 3.6(a) compares in the (R,z) plane equidensity contours of the overall matter density distribution ρ_Φ, generating the potential. Panel 3.6(c) and 3.6(d) show the potential's circular velocity curve and radial surface density profile within $|z| = 1.1$ kpc, respectively. We compare the true ρ_Φ, $v_{\rm circ}$, and $\Sigma_{1.1\rm kpc}$, of the galaxy simulation (azimuthally averaged over the whole galaxy; black solid lines), with 100 `MNHH-Pot` potentials drawn from the PDF of the best-fit RoadMapping model (blue lines). This model was derived from $N_* = 20,000$ stars in the spherical survey volume at `S8` with $r_{\rm max} = 4$ kpc. (The extent of the survey volume is marked in orange.) The best-fit parameters are given in Table 3.3, and Panel 3.6(b) shows the residuals between the matter density corresponding to the median values in this table, $\rho_{\Phi,\rm M}$, and the true density $\rho_{\Phi,\rm T}$. Overplotted in Panels 3.6(a), 3.6(c), and 3.6(d) is also the reference `DEHH-Pot` (see Section 3.2.3; black dotted line). Over wide areas even outside of the survey volume the relative difference between true and recovered density is less than 15%. At $R \gtrsim 8$ kpc and $z \sim 0$ it becomes apparent that the chosen potential model cannot perfectly capture the structure of the disk. However, in the plane of the disk and at smaller radii within the survey volume, where most of the stars are located, the model gives good constraints on the density. The circular velocity curve is recovered to less than 5%.

the radial force, $F_{R,T}(\boldsymbol{x}_{*,i})$, and vertical force, $F_{z,T}(\boldsymbol{x}_{*,i})$, calculated as the sum of the individual contributions by each particle in the simulation and the analytic DM halo at the position of each star $\boldsymbol{x}_{*,i} \equiv (x_i, y_i, z_i)$) with the force that the RoadMapping median model predicts for each star ($F_{R,M}(\boldsymbol{x}_{*,i})$ and $F_{z,M}(\boldsymbol{x}_{*,i})$; M for "median model"). We scale the difference between truth and



(a) Recovery of the radial forces.

(b) Recovery of the vertical forces.

Figure 3.7: Recovery of the gravitational forces with RoadMapping. We compare the true gravitational forces with the forces estimated from the RoadMapping best-fit potential in Table 3.3 at the (x, y) positions (upper panels) and (R, z) positions (lower panels) of the stars that entered the analysis. In particular, we color-code the positions of the stars according to the radial (panel 3.7(a)) and vertical (panel 3.7(b)) force residuals scaled by a typical force, i.e., we show $\Delta F_R(\boldsymbol{x}_{*,i})$ and $\Delta F_z(\boldsymbol{x}_{*,i})$ in Equations (3.23) and (3.24). The overplotted gray contours correspond to $\Sigma_{1.5\mathrm{kpc,disk},T}/\Sigma_{1.5\mathrm{kpc,disk},S} = 1.15$, i.e., the true versus the symmetric disk surface density, and mark the position of the spiral arms. The red dots mark stars for which the best-fit model underestimates the (absolute value of the) force. This is the case for the radial force in the leading sides of the spiral arms and the vertical force within the spiral arms, which cannot be reproduced. Blue marks correspond to stars for which the force is overestimated. Overall, the radial forces are very well recovered, which is related to the good recovery of the circular velocity curve in Figure 3.6(c). There are more problems with the vertical force, which is related to the higher surface densities in spiral arms. This slightly biases the overall RoadMapping model.

model by a typical radial or vertical force at the given radius for which we use

$$F_{R,\mathrm{typ}}(R) \equiv v^2_{\mathrm{circ},S}(R)/R \tag{3.21}$$

$$F_{z,\mathrm{typ}}(R) \equiv F_{z,S}(R, z = z_\mathrm{s}), \tag{3.22}$$

where $v_{\mathrm{circ},S}$ and $F_{z,S}$ are the circular velocity and vertical force evaluated in the "true symmetric" reference potential `DEHH-Pot` in Table 3.2. As the typical vertical force at a given radius, we use $F_{z,S}$ evaluated at the scale height of the disk, $z_\mathrm{s} = 0.17$ kpc. This is of the same order as the true vertical force averaged over all stars at this radius. Figure 3.7 shows therefore

$$\Delta F_R(\boldsymbol{x}_{*,i}) \equiv \frac{|F_{R,M}(R_i, z_i)| - |F_{R,T}(x_i, y_i, z_i)|}{|F_{R,\mathrm{typ}}(R_i)|} \tag{3.23}$$

$$\Delta F_z(\boldsymbol{x}_{*,i}) \equiv \frac{|F_{z,M}(R_i, z_i)| - |F_{z,T}(x_i, y_i, z_i)|}{|F_{z,\mathrm{typ}}(R_i)|} \tag{3.24}$$

for each star ($R_i = \sqrt{x_i^2 + y_i^2}$). The recovery is as expected. The true vertical force is stronger in the spiral arms (due to the higher surface density) and weaker in the inter-arm regions as compared to the axisymmetric best-fit model. The radial force is well recovered where the majority of the stars are located, i.e., in the wide inter-arm regions and in the peaks of the spiral arms. Misjudgments happen in the wings of the spiral arms: The true radial force (i.e., the pull toward the galactic center) is stronger at the outer edge/leading side of the spiral arm because of the additional gravitational pull toward the massive spiral arm, and for the same reason weaker at the inner edge/trailing side. Overall, the recovered RoadMapping model appears to be a good mean model, averaging over spiral arms and inter-arm regions.

3.4.1.3 Recovering the Action Distribution

In Sections 3.4.1.1 and 3.4.1.2, we have demonstrated the goodness of the fit in the configuration space of the data, and of the recovered gravitational potential. What RoadMapping is actually fitting, however, is the distribution in action space. Figure 3.8 compares the data and the model action distribution (generated by the best-fit qDF) given the best-fit median `MNHH-Pot` in Table 3.3. (We use this axisymmetric potential to calculate the actions that lead to the best-fit model, and do not attempt to estimate the true actions in the true potential.)
We note that the radial and vertical action distribution fits quite well; the axisymmetric model, however, contains many more stars on close-to-circular orbits ($J_R \sim 0, J_z \sim 0$) than the simulation. In the data set, there is an excess of stars in the galactic plane ($J_z \sim 0$) that have more eccentric orbits than the axisymmetric model would predict. In Figure 3.5(a), we have marked the radial extent of the stronger spiral arm with dotted lines ($R_{\rm spiral} \in [5.6, 6.8]$ kpc), and overplotted the corresponding angular momenta $L_z = R_{\rm spiral} \times v_{\rm circ}(R_{\rm spiral})$ in Figure 3.8. This serves as a rough estimate for the region in action space, where we expect the stars of this spiral arm to be located. It is again obvious that this spiral arm contains (1) more stars in general and (2) more stars with eccentric orbits ($J_R > 0$), which are (3) mostly located close to the plane ($J_z \sim 0$), as compared to the axisymmetric model. All of this confirms our expectations for orbits in a spiral arm.
One of the open tasks that the Galactic dynamical modeling community faces is the description of the orbit distribution of spiral arms. The above exercise of comparing the data and the model actions in a best-fit axisymmetric potential should therefore be performed for any future application to data in the MW as well. It could help to learn more about the approximate orbits that stars move on in real spiral arms, and how spiral arms perturb axisymmetric action DFs.

3.4.1.4 Calibrating the Method by Modeling a Snapshot without Spiral Arms

To better understand the effect of spiral arms on the RoadMapping modeling in the previous sections, we also performed a calibration test run. For that, we applied RoadMapping to a mock data set drawn from the same spherical survey volume as in Section 3.4.1 (centered at `S8` with a maximum radius of $r_{\rm max} = 4$ kpc), but from the initial axisymmetric snapshot of the galaxy simulation ($t = 0$ Myr), in which no spiral arms had evolved yet.
The spatial tracer distribution was very well recovered by RoadMapping, as well as the distribution of tracers in v_R and v_z. There were some deviations in the v_T distribution between data and best-fit model, however, with our model predicting more asymmetric drift. We attribute this to the simple initial setup of the disk's stellar velocities, which does not follow a physical DF but only assumes a triaxial Gaussian velocity distribution (Springel et al. 2005). We expect the modeling in this calibration run to be slightly biased by this. Other than that, this confirms our expectation that the qDF is indeed a good model for the unperturbed disk, and that the deviations between data and model in Figure 3.5 were purely due to the spiral arms.
The surface density profile generated by the best-fit potential was a perfect fit inside of $R \sim 7$ kpc, but it started flaring further out. This was expected, because the Miyamoto-Nagai disk is known

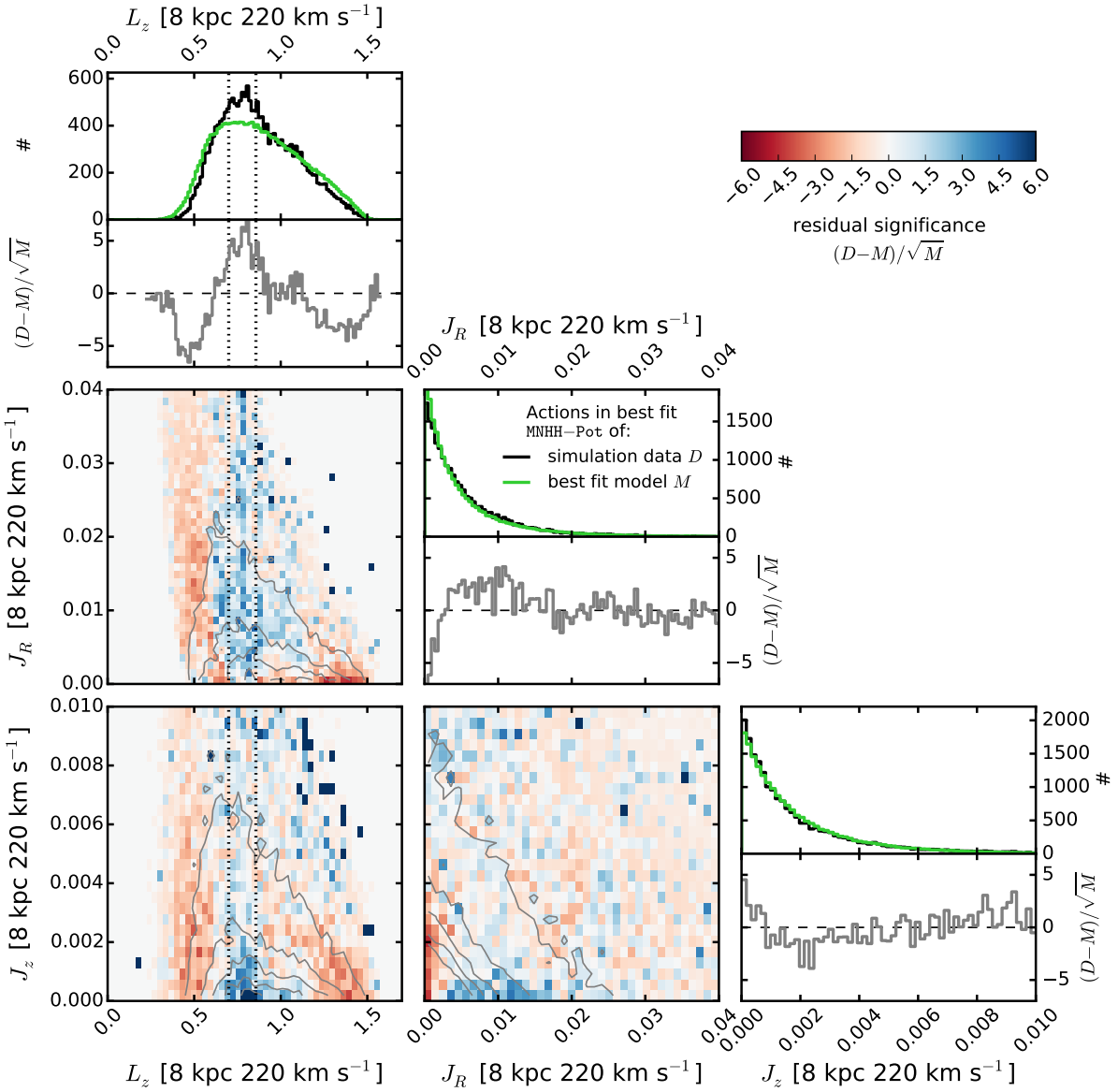


Figure 3.8: Comparison of the stellar action distribution of the data set D used in the analysis and the recovered axisymmetric action distribution M (see Figure 3.5 for the comparison in configuration space). All actions of the data set and best-fit distribution were calculated in the best-fit `MNHH-Pot` in Table 3.3. The upper panel in each column shows one-dimensional histograms of the D and M distribution of angular momentum, L_z, the radial action, J_R, and the vertical action, J_z. The other panels display the residual significance $(D-M)/\sqrt{M}$, as both one-dimensional and two-dimensional distribution (the model M was constructed by MC sampling the best-fit qDF, $\sqrt{M}$ is the expected noise due to Poisson statistics, and $(D-M)/\sqrt{M}$ therefore describes how significant any difference between D and M is). The two-dimensional residuals are overplotted with equidensity contours of the data D's two-dimensional action distribution (gray solid lines). In Figure 3.5(a), we have marked the approximate radial extent of the stronger spiral arm with black dotted lines ($R_{\rm spiral} \in [5.6, 6.8]$ kpc); the dotted lines in the L_z distributions in this figure correspond to $L_z = R_{\rm spiral} \times v_{\rm circ}(R_{\rm spiral})$. This comparison gives a first impression of how the approximate action distribution in spiral arms might look.

to be more massive at larger radii than an exponential disk (Smith et al. 2015). Also, if the fit is driven by the majority of stars, the fit is expected to be better at smaller radii with its higher tracer number density. The same flaring in the surface density also showed up in Figure 3.6. Overall, our chosen potential model will therefore systematically bias vertical forces at large radii to be too strong as compared to the truth.
The recovered potential parameters for the initial snapshot, $a_{\rm disk} = (3.73 \pm 0.04)$ kpc, $b_{\rm disk} = (0.34 \pm 0.03)$ kpc, $f_{\rm halo} = (0.52 \pm 0.03)$, and $a_{\rm halo} = (24 \pm 2)$ kpc, are consistent with the spiral arm-affected measurements in Table 3.3 to within 2, 4, 1, and 1.5 times the statistical error, respectively. While the halo scale length in this calibration run is also consistent with the truth within three times the error, we expect it to be underestimated to at least partly account for the Miyamoto-Nagai disk being too massive at large radii. As we will see later (in Section 3.4.2.3 and Figure 3.12), we seem to need an even larger survey volume to have enough radial coverage to constrain the halo scale length properly.
Compared to the truth, the circular velocity curve derived from the initial snapshot was underestimated by $\sim 2.5\%$; $v_{\rm circ}(R_\odot) = (219.36 \pm 0.08)$ km s^{-1}. For the snapshot with spiral arms (see Figure 3.6), we observe an underestimation of only $\sim 1.5\%$, so the bias in the initial snapshot might be related to its unphysical setup of the v_T velocities. One the one hand, because this underestimation is present for both snapshots, it could also be a systematic bias due to the chosen Miyamoto-Nagai disk model. On the other hand, the bias also shows up in the `DEHH-Pot`'s circular velocity curve in Figure 3.6. The `DEHH-Pot` has a more realistic shape and was found as a direct fit to the spiral-arm-affected disk particle distribution. The $\sim 1.5\%$ underestimate is therefore most likely introduced by the spiral arms (see Section 3.4.2.6).
Because the simulation immediately develops spiral arms after the initial snapshot, there was no snapshot that was still axisymmetric, yet already in a dynamical steady state.

3.4.2 The Influence of Spiral Arms in RoadMapping Modeling

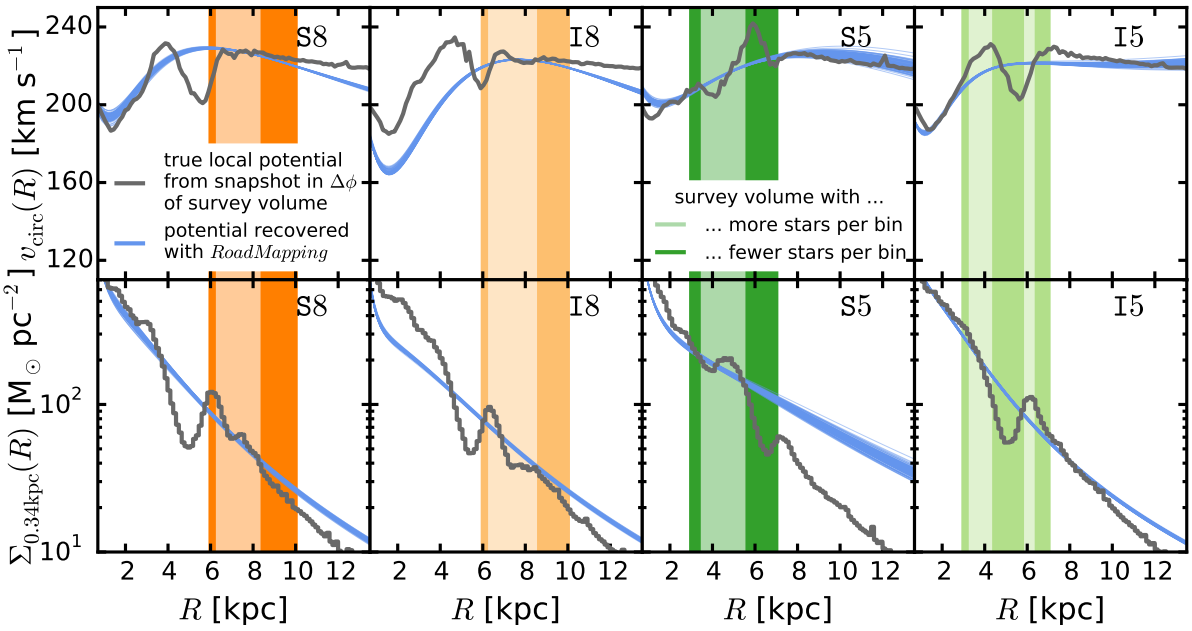
In the previous section, we showed that for a large survey volume ($r_{\rm max} = 4$ kpc) RoadMapping can construct a good average axisymmetric potential (and DF) model for a galaxy with spiral arms. In the following, we want to investigate how this modeling success depends on the position and the size of the survey volume within the galaxy and with respect to the spiral arms.

3.4.2.1 A Suite of Data Sets drawn from Spiral Arms and Inter-arm Regions

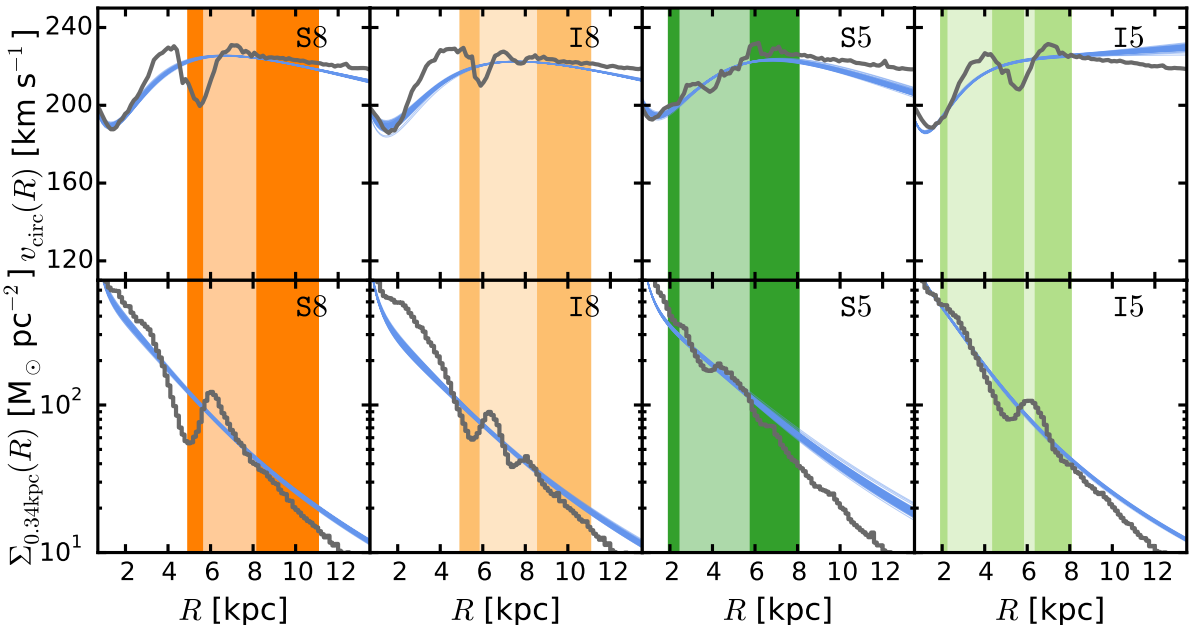
To investigate a range of data sets affected in different proportions by spiral arms, we center our test survey volumes at the positions marked in Figures 3.1 and 3.4 (see also Table 3.1) and consider volume sizes with $r_{\rm max} \in [0.5, 1, 2, 3, 4, 5]$ kpc for $R_0 = 8$ kpc and $r_{\rm max} \in [0.5, 1, 2, 3, 4]$ kpc for $R_0 = 5$ kpc (to avoid the galactic center). As demonstrated in Figure 3.4, the spiral arm strength is very different in these test volumes. Each data set that we draw from the simulation contains a random selection of $N_* = 20,000$ stars inside the given spherical volume and we fit a single qDF and `MNHH-Pot` to it.

3.4.2.2 Recovering the Circular Velocity Curves and Surface Density Profiles

It turns out that RoadMapping is successful in finding reasonable and even very good best-fit potential models for each one of the 22 test data sets independent of size and location—given the data and limitations of the model. To illustrate this and to make this encouraging result immediately obvious, we show the circular velocity curves and surface density profiles of all analyses in Figures 3.9-3.11.
In contrast to Figure 3.6, these Figures only show the true profiles for the region within the galaxy where the data comes from: averaged over the angular wedge covering the radial extent of

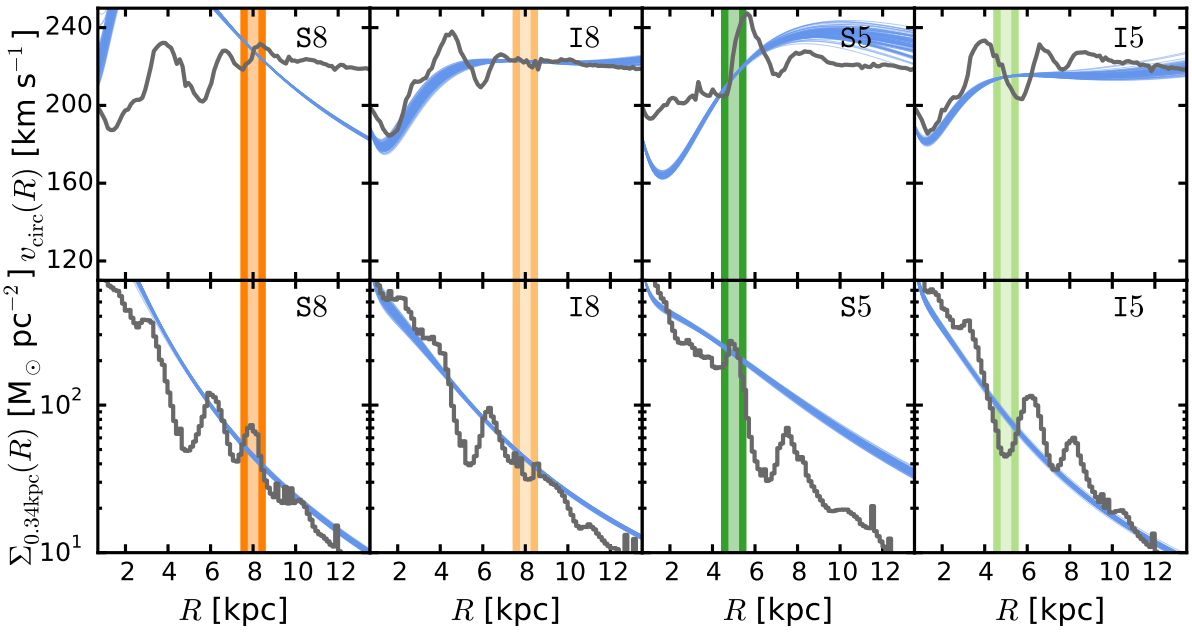


(a) $r_{\rm max} = 2$ kpc.

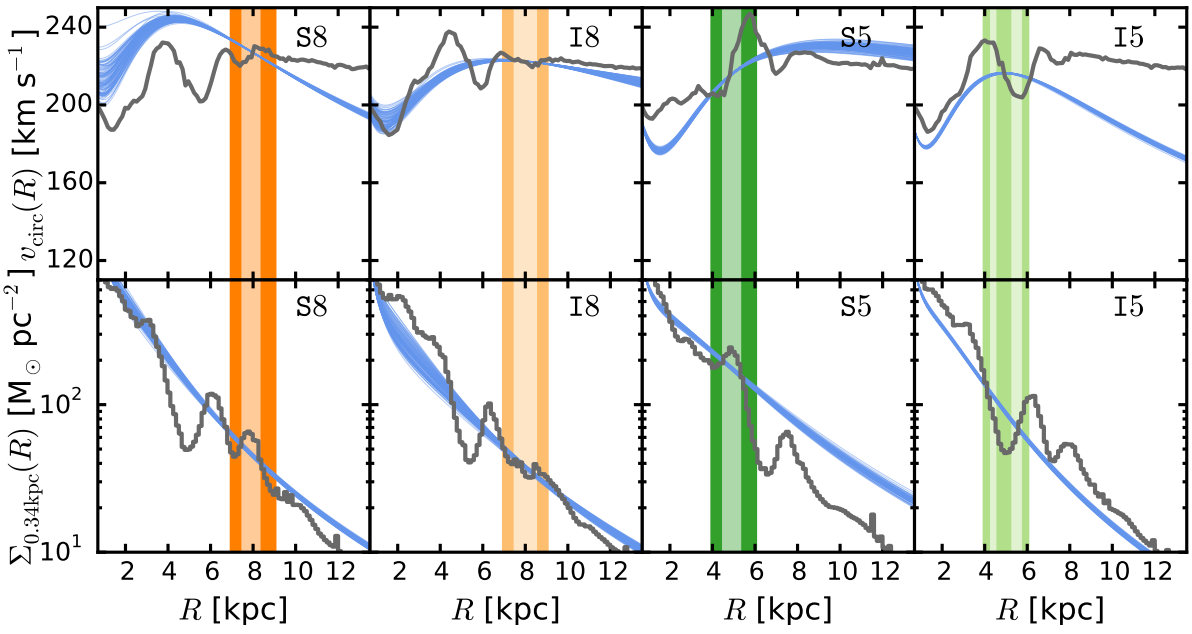


(b) $r_{\rm max} = 3$ kpc.

Figure 3.9: Comparison of the true local circular velocity curve and surface density within $|z| \leq 2z_{\rm s} = 0.34$ kpc with the recovered RoadMapping models from survey volumes of size $r_{\rm max} = 2$ kpc and $r_{\rm max} = 3$ kpc. The blue lines show the RoadMapping potential models recovered from these data sets (each line is one of 100 potentials drawn from the full PDF sampled with the MCMC). The gray curves show the true profiles as derived from the galaxy simulation snapshot, averaged over the angular wedge $\phi_0 \pm \arcsin(r_{\rm max}/R_0)$ that encloses the corresponding survey volume (see Table 3.1 for all R_0 and ϕ_0 values). In other words, we show the true profiles only for the region of the spiral galaxy that was actually probed by the data. The orange and green colored regions mark the radial extent of the survey volumes. We sorted the stars of each data set into radial bins of size $\Delta R = 200$ kpc. The radial bins with a higher than average number of stars are marked with a lighter shade of the corresponding color, and the bins with a lower than average number with a darker shade. It turns out that the constraints of highest accuracy and precision are always where most of the stars are located—within the survey volume and, in particular, at the peak of the distribution.



(a) $r_{\rm max} = 500$ pc.

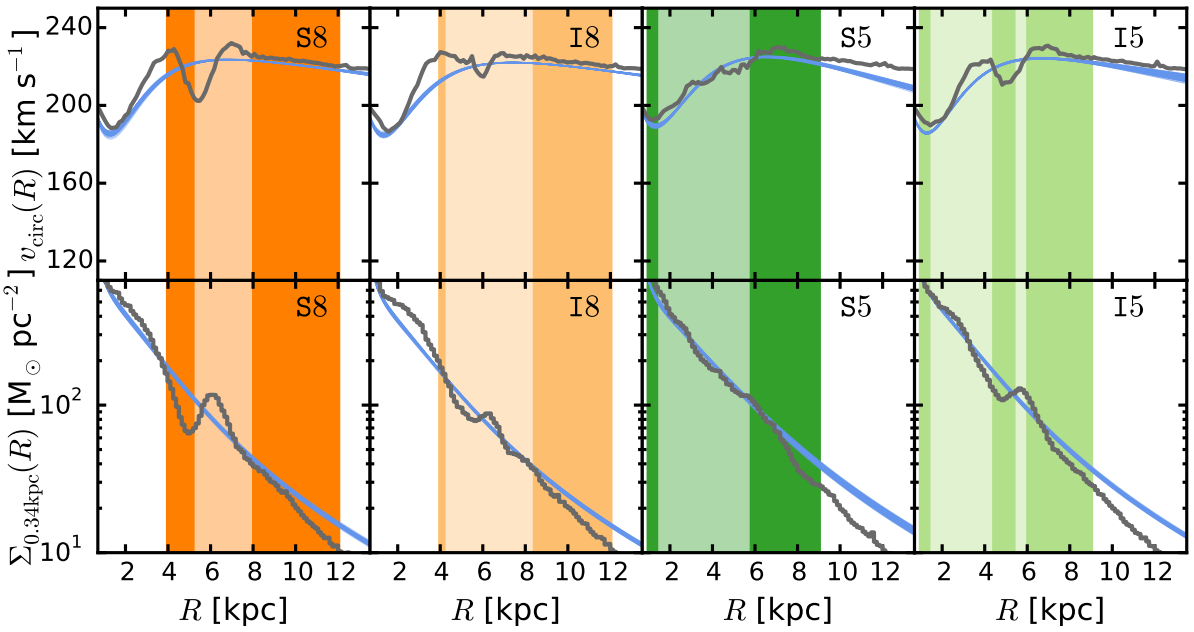


(b) $r_{\rm max} = 1$ kpc.

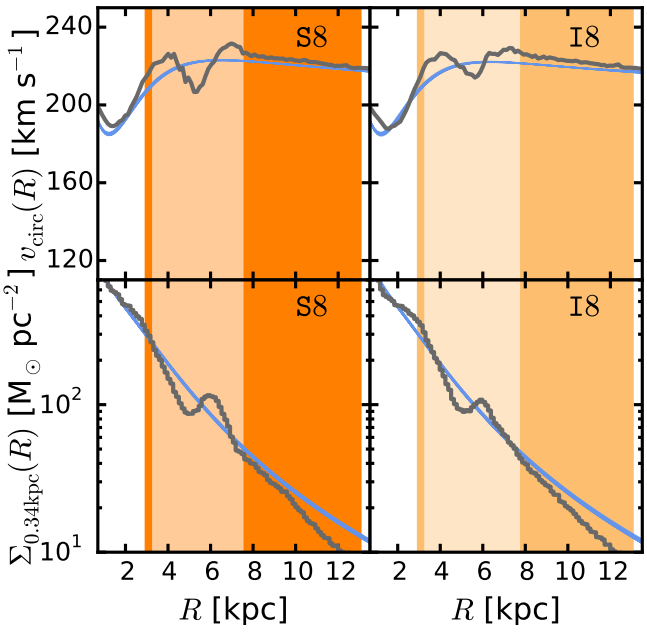
Figure 3.10: Same as Figure 3.9, but for all small survey volumes with $r_{\rm max} = 500$ pc and $r_{\rm max} = 1$ kpc.

the survey volume, $\Delta\phi = \phi_0 \pm \arcsin(r_{\rm max}/R_0)$, and within $|z| \leq 2z_{\rm s} = 0.34$ kpc, i.e., twice the scale height of the disk, which contains most of the disk mass. This is the matter distribution in which the stars are currently moving, and therefore the potential to which the modeling should be sensitive. In Figures 3.9-3.11, we also mark the survey volume and the radial bins of size $\Delta R = 200$ pc with the highest number of stars.

Even though the curves vary extremely between the individual data sets, it becomes very obvious that it is indeed the regions in which the majority of stars is located that drives the RoadMapping fit. Furthermore, whether this region is dominated by a spiral arm or an inter-arm region, and even if this region is only as small as $r_{\rm max} = 500$ pc, RoadMapping indeed constrains the local



(a) $r_{\rm max} = 4$ kpc.



(b) $r_{\rm max} = 5$ kpc.

Figure 3.11: Same as Figure 3.9, but for all big survey volumes with $r_{\rm max} = 4$ kpc and $r_{\rm max} = 5$ kpc.

potential where most of the stars of the data set are located. Also, the constraints are not only most accurate but also most precise in these regions.

Only in the two volumes with $r_{\rm max} = [0.5, 1]$ kpc at position `S8` does RoadMapping have some difficulties fitting the circular velocity curve; the model expects a flat or falling rotation curve and is presented with a steeply rising rotation curve due to the spiral arm dominating the region. However, given that RoadMapping recovers a good average surface density profile and the circular velocity at least at the center of the small volume, the fit is still quite successful.

In an application to real data in the MW, we would also have the possibility to impose some informative prior information on the potential shape (e.g., on the rotation curve), to avoid very

unrealistic results (see also discussion in Section 3.5.4).

3.4.2.3 Discussion of the Model Parameter Recovery

Figures 3.9-3.11 in the previous section have illustrated how well the potential is recovered by RoadMapping. Figure 3.12 compares the potential and qDF parameters found with RoadMapping to the parameters of the reference `DEHH-Pot` from Table 3.2. At first glance, there appear to be several discrepancies. In the following, we will discuss the deviations and explain why each set of parameters still corresponds to a good fit to the data.

Overall, the statistical random errors on the parameter recovery are very small for $N_* = 20,000$ and possible systematic errors dominate. There are only a few exceptions ($r_{\max} = 500$ pc, $r_{\max} = 1$ kpc at `I5`, $r_{\max} = 2$ kpc at `I8`), which we will discuss later.

We will first consider the parameters of the gravitational potential (left column in Figure 3.12). All volumes recover $v_{\rm circ}(R_\odot)$ within a few km s^{-1}; in the largest volumes—where the circular velocity curve is probed over several kiloparsecs—the estimate is the most accurate. The halo fraction $f_{\rm halo}$ of the radial force at the "solar" radius $R_\odot$ is very well recovered, especially for $r_{\max} \gtrsim 2$ kpc. The estimate that we get for the best-fit Miyamoto-Nagai disk scale height $b_{\rm disk}$ seems to be also approximately independent of the size of the volume. We can even recover the true halo scale length $a_{\rm halo}$; however, only for a volume as large as $r_{\max} = 5$ kpc. The models at $r_{\max} = 500$ pc appear to be too small to constrain the halo at all, and the MCMCs diverged completely for this parameter. Smaller volumes that underestimate $a_{\rm halo}$ get slightly larger estimates for the disk scale length $a_{\rm disk}$ and the overall radial density slope is then probably closer to the truth, even if the individual parameters are not. Outliers can often be explained by having a look at the data. The large disk scale length recovered from the $r_{\max} = 2$ kpc volume at `S5`, for example, mirrors the comparably flat matter distribution caused by two spiral arms close together and dominating the volume (see Figure 3.9(a) and the large $\sigma_{\Delta\rm Spiral}$ for this analysis in Figure 3.4(d)).

The right column of Figure 3.12 compares the recovered qDF parameters for the different survey volumes with the qDF parameters we obtained from fixing the potential model to the `DEHH-Pot` and fitting the qDF only in a $r_{\max} = 5$ kpc volume at `S8`. Even though the qDF parameters for small volumes are widely different for different positions within the galaxy, they all approach the values recovered with the `DEHH-Pot` for larger volumes. There seems, therefore, to be an overall best-fit qDF describing the average tracer distribution in the galaxy's disk. The only difference is in the $h_{\sigma,z}$ parameter, where the models fitting an `MNHH-Pot` recover a slightly larger value than the models using the known `DEHH-Pot`. The suspected reason is that the Miyamoto-Nagai disk flares at larger radii as compared to the double exponential-disk (see Figure 3.6), which leads to a less-steep radial decline in the vertical forces, and therefore mean vertical orbital energies $\langle E_z \rangle \sim \nu \times J_z$, and therefore to a slightly longer $h_{\sigma,z}$ scale length. In general, volumes centered on spiral arms have larger velocity dispersion parameters $\sigma_{R,0}$ and $\sigma_{z,0}$ as compared to volumes at the same radius R_0 but centered on an inter-arm region. Furthermore, the volumes at $R_0 = 5$ kpc with their stronger spiral arms have larger velocity dispersions than those at $R_0 = 8$ kpc—which is what we expect. Most volumes recover similar tracer scale lengths $h_R \sim 2.5 \pm 0.5$ kpc close to the known disk scale length $R_{\rm s}$. Only the volumes centered on the inter-arm region at $R_0 = 8$ kpc (position `I8`) recover much longer h_R. This might be related to the fact that volumes at `I8` are dominated by an especially extended inter-arm region. The volumes at `I5` with $r_{\max} = [0.5, 1]$ kpc were not able to constrain the tracer scale length at all because of the unfortunate position between the rising density wings of two strong spiral arms (see Figure 3.10).

There are a few survey volumes for which the recovered parameters show some peculiarities. The models from volumes with $r_{\max} = 1$ kpc at `I5` and $r_{\max} = 2$ kpc at `I8` reject the DM halo com-

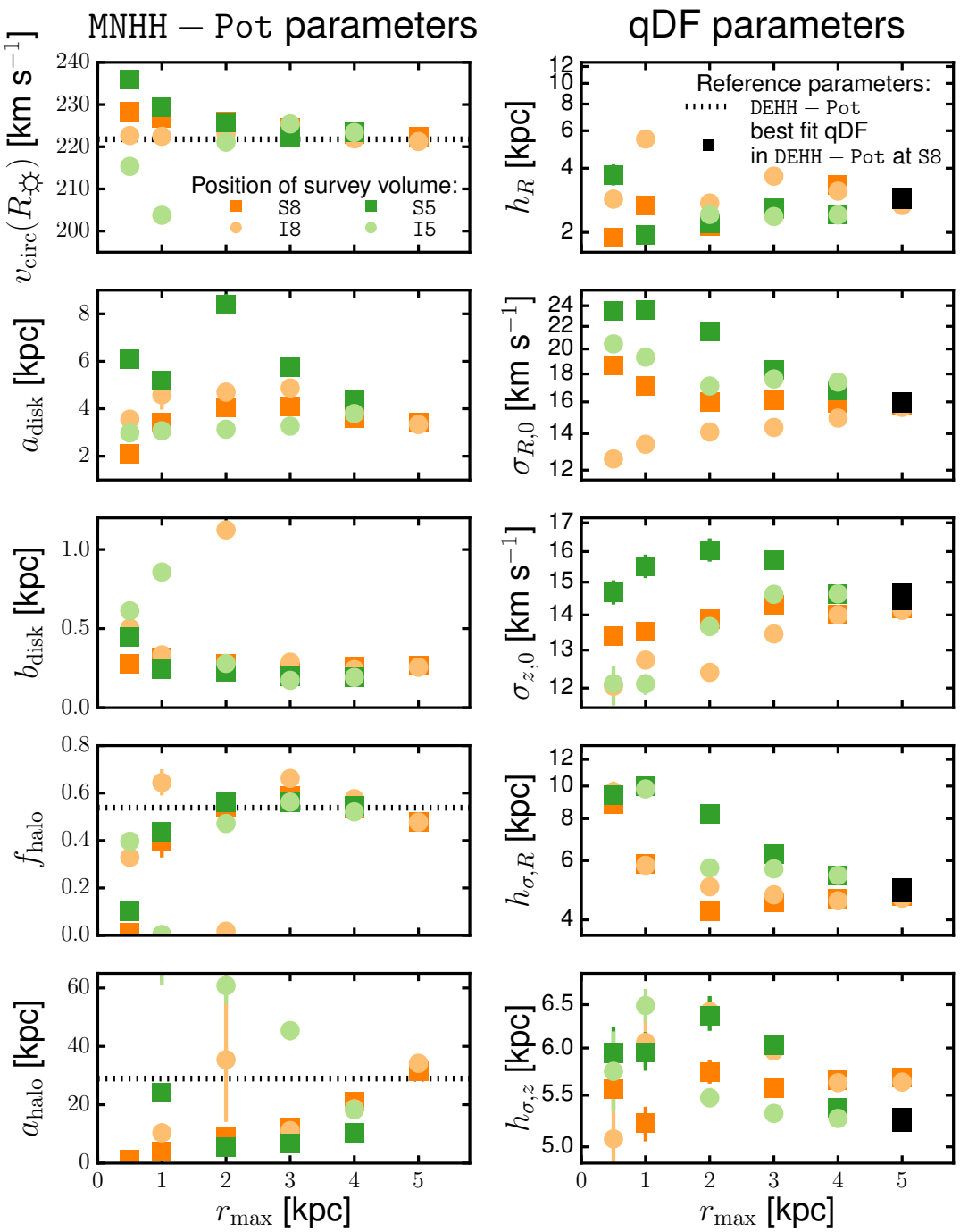


Figure 3.12: Overview of the model parameter estimates (`MNHH-Pot` parameters on the left, qDF parameters on the right) recovered with RoadMapping from 22 different data sets. All data sets were drawn from the same simulation snapshot, but from survey volumes at different positions in the galaxy (color-coded) and of different sizes ($r_{\rm max}$ as indicated on the x-axis). Note that all five qDF parameters are shown here on a logarithmic scale, because RoadMapping uses a logarithmically flat prior for them in the fit (see Equation (3.16)). The black dotted line shows the known model parameters from the reference potential `DEHH-Pot` in Table 3.2 (the Miyamoto-Nagai disk parameters $a_{\rm disk}$ and $b_{\rm disk}$ are related but not directly comparable to an exponential disk scale length and height). The black squares denote the qDF parameters we recovered by fixing the potential to the `DEHH-Pot`, centering a survey volume with $r_{\rm max} = 5$ kpc on the spiral arm at $R_0 = 8$ kpc (position `S8`), and fitting the qDF only. A survey volume with a radial coverage as large as $r_{\rm max} = 5$ kpc is required to properly recover all "true" model parameters. For smaller volumes, there seem to be quite large deviations between truth and model; that these recovered parameters still all correspond to successful RoadMapping fits to the data is discussed in Section 3.4.2.3.

pletely, i.e., $f_{\rm halo} = 0$. The corresponding halo scale lengths $a_{\rm halo}$ are therefore unconstrained,[19] while the corresponding disk scale heights $b_{\rm disk}$ are grossly overestimated to account for the missing contribution of the spherical halo. We have investigated the reason for this fitting result and found that for the way in which the spiral arms affect the circular velocity curve in these volumes, the recovered models with unusual radial profiles are indeed a better description for the data (see Figure 3.9(a) and 3.10(b)). Also, while most analyses average the vertical forces radially over the spiral arms (see Figures 3.7, lower right panel), for these analyses the averaging happens vertically, i.e., at approximately one scale height above the plane where the model's vertical forces are equally well recovered at all radii (in spiral arms and between), while at small and large $|z|$, the model is bad. RoadMapping therefore also found a good average fit model for the stars in these volumes.
Overall, we find that if the volume is large enough to average over several spiral arms and inter-arm regions, an unlucky positioning with respect to the spiral arms does not lead to strong biases in the parameter recovery. We stress again that for particularly large volumes, $r_{\rm max} = 5$ kpc, we were able to recover all model parameters, including the halo scale length $a_{\rm halo}$.

3.4.2.4 Recovering the Local Gravitational Forces

In the previous section, we found that the potential and qDF parameters recovered from different survey volumes can be quite different. While the differences can be explained qualitatively, it is not yet clear how good the corresponding potential constraints actually are in a quantitative sense. To test this, we calculate again $\Delta F_R(\boldsymbol{x}_{*,i})$ and $\Delta F_z(\boldsymbol{x}_{*,i})$ from Equations (3.23) and(3.24) at the position of each star $\boldsymbol{x}_{*,i}$ in each data set (analogous to Figure 3.7). From the corresponding histograms of number of stars versus ΔF, we derive the median and the $16th$ and $84th$ percentiles (1σ range) and show them in Figure 3.13. We chose this diagnostic because the forces at the positions of the stars are the quantities of the potential to which our modeling is sensitive.
The important key result from Figure 3.13 is that we get very close to recovering the true forces $\Delta F(\boldsymbol{x}_{*,i}) \lesssim 10\%$ at the positions of the majority of stars in the survey volume, no matter how large or small the survey volume is. On average, the force recovery is also unbiased[20] for the ensemble of stars.
Figure 3.13 also contains some subtle clues that suggest that the quality of the force recovery could be correlated with the position of the data set with respect to the spiral arms. We investigate this further by relating in Figure 3.14 the local force recovery, i.e., the distribution of $\Delta F_R(\boldsymbol{x}_{*,i})$ and $\Delta F_z(\boldsymbol{x}_{*,i})$ for each data set, to the relative spiral contrast within the respective survey volume, $\sigma_{\Delta\rm Spiral}$ (Equation (3.6); see also Figure 3.4(d)). Figure 3.14 shows that the average fraction of stars for which the recovery of the radial or vertical force is bad (i.e., larger than 10%) increases with increasing spiral contrast $\sigma_{\Delta\rm Spiral}$. This is as expected. Volumes in which the steep gradient in surface density around a strong spiral arm ($|\Delta_{{\rm Spiral},k}| > 0$) is not balanced by larger areas with less perturbations ($\Delta_{{\rm Spiral},k} \sim 0$) have (1) a large relative spiral contrast $\sigma_{\Delta\rm Spiral}$, and (2) a large relative number of stars affected by the non-axisymmetric kinematics of the spiral arms. For these individual stars, the axisymmetric RoadMapping model is less successful in recovering the correct forces. (However, as we saw in Figure 3.13, the ensemble average is unbiased even in these cases.)
Interestingly, and even though there is some scatter, the force recovery at a given $\sigma_{\Delta\rm Spiral}$ is on average very similar for the radial and vertical forces (compare the linear fits in Figure 3.14). This means that RoadMapping attempts to fit both the radial and vertical forces at the positions

[19] For these analyses and the $r_{\rm max} = 500$ pc analyses the fit could not constrain $a_{\rm halo}$ and the MCMC was diverging. We had to stop the MCMC after some time, so the $a_{\rm halo}$ might in truth be even less constrained than shown in Figure 3.12.

[20] An exception are the radial forces for small volumes strongly dominated by spiral arms (e.g., at `S5`). The small systematic bias in F_R is discussed in detail in Section 3.4.2.6.

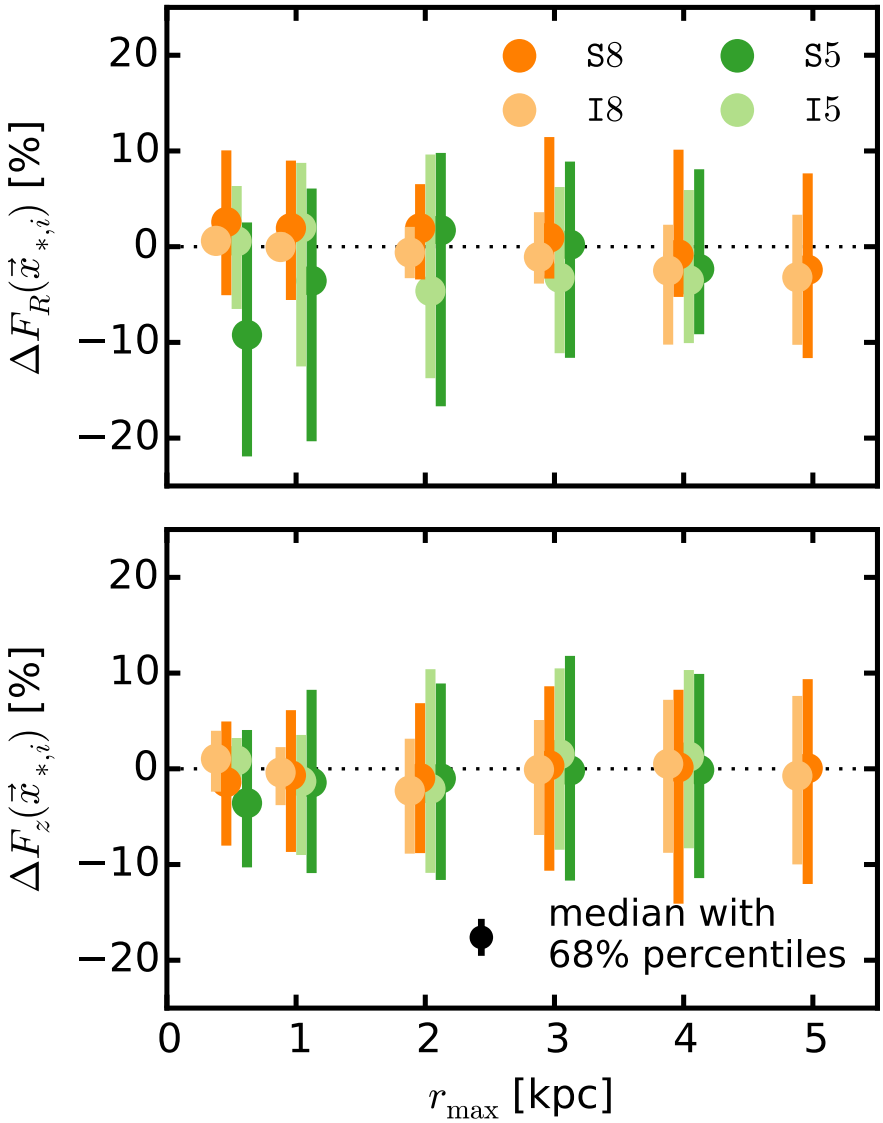


Figure 3.13: Accuracy of the radial (upper panel) and vertical (lower panel) gravitational forces recovered with RoadMapping from the suite of data sets introduced in Section 3.4.2.1 for the ensemble of stars in each data set. The x-axis denotes the radial size $r_{\max}$ of the survey volume belonging to each data set. (For presentation purposes, we added a small offset $\ll 1$ to $r_{\max}$ on the x-axis.) The y-axis shows the distribution of force residuals, $\Delta F_R(\boldsymbol{x}_{*,i})$ and $\Delta F_z(\boldsymbol{x}_{*,i})$ in Equations (3.23) and (3.24), at the positions of the stars $x_{*,i}$ that entered the analysis. In particular, we show here for each distribution of $\Delta F(\boldsymbol{x}_{*,i})$ the median as a dot with the $[16th,84th]$ percentile range as a bar. We find that the forces are very well recovered at the positions of the stars independent of the size of the volume.

of the stars, and is not particularly sensitive to just one of them.

As we saw in Figure 3.4(d), the spiral contrast $\sigma_{\Delta\text{Spiral}}$ increases for the different test volume positions approximately in this order: `I8` $\longrightarrow$ `S8` $\longrightarrow$ `I5` $\longrightarrow$ `S5`. From Figure 3.14, it follows that this is also the order in which the accuracy of the force recovery decreases. (We did not include this piece of additional information in Figure 3.14, but it can be seen in Figure 3.13, especially for the smaller volumes.)

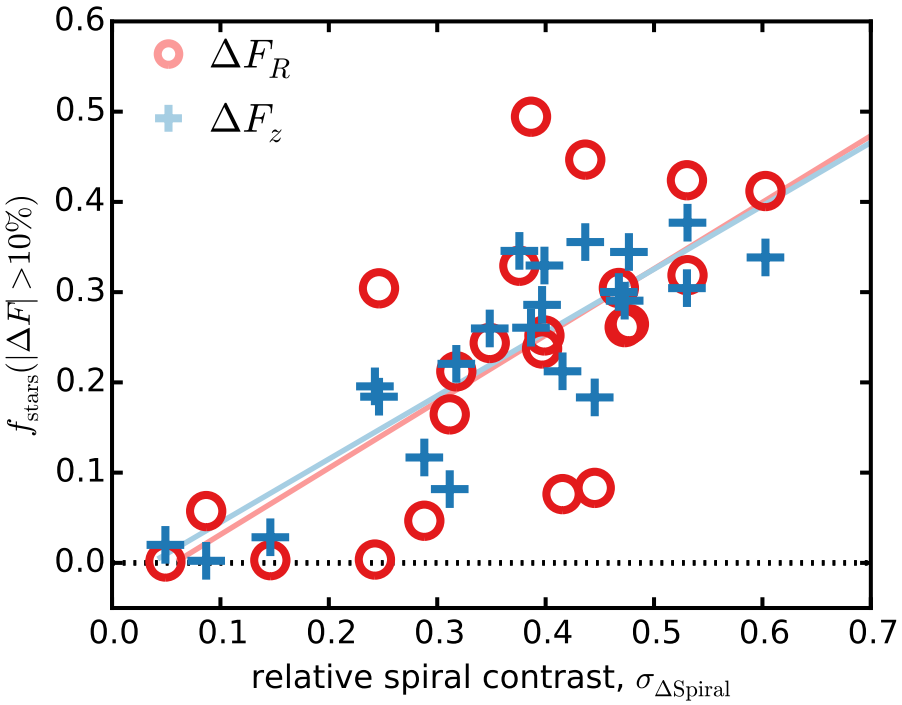


Figure 3.14: Influence of the spiral arm contrast on the recovery of the gravitational forces at the positions of the stars $x_{*,i}$ that entered the RoadMapping analysis. Each circle/cross pair corresponds to one of our 22 data sets. The relative spiral contrast on the x-axis is quantified as $\sigma_{\Delta\rm Spiral}$ calculated within each survey volume according to Equation (3.6) in Section 3.2.4. On the y-axis, the fraction of stars ($f_{\rm stars}$) in each data set is shown for which the radial (red circles) and vertical (blue crosses) force residual calculated from Equations (3.23)-(3.24) is larger than 10% (i.e., at 0 all stars have good force measurements, at 1 everything went wrong). The red and blue lines are linear fits to the radial and vertical force residual fraction, respectively, and are guides to the eye that show the clear and expected trend that in volumes with smaller spiral arm contrast, where comparably fewer stars are located in spiral arms, the axisymmetric best-fit model can recover the true gravitational forces also for more stars. On average, the radial and vertical forces are equally well recovered at a given spiral contrast.

3.4.2.5 Extrapolating the Gravitational Potential Model

It is also interesting to see how well the extrapolation of a recovered potential describes the overall gravitational potential of the galaxy. To investigate the extrapolability, we introduce another diagnostic that uses a cylindrical grid centered on the respective positions in Table 3.1, always having a radius of $r_{\rm max} = 5$ kpc and a height of $z = 1.5$ kpc both above and below the plane. In the (x, y) plane, the regular grid points have a distance of 0.25 kpc and in z they have a distance of 0.125 kpc to better sample the thin disk. (We throw out grid points close to the galactic center with $R < 0.125$ kpc.) We then evaluate at the position $\boldsymbol{x}_{g,j} \equiv (x_j, y_j, z_j)$ of each regular grid point the force residuals

$$\Delta F_R(\boldsymbol{x}_{g,j}) \equiv \frac{|F_{R,M}(R_j, z_j)| - |F_{R,T}(x_j, y_j, z_j)|}{|F_{R,\rm typ}(R_j)|} \tag{3.25}$$

$$\Delta F_z(\boldsymbol{x}_{g,j}) \equiv \frac{|F_{z,M}(R_j, z_j)| - |F_{z,T}(x_j, y_j, z_j)|}{|F_{z,\rm typ}(R_j)|}, \tag{3.26}$$

analogous to Equations (3.21)-(3.24). The two panels in Figure 3.15 show the [16th,84th] percentile range and the median of the grid points' distribution in $\Delta F_R(\boldsymbol{x}_{g,j})$ and $\Delta F_z(\boldsymbol{x}_{g,j})$.

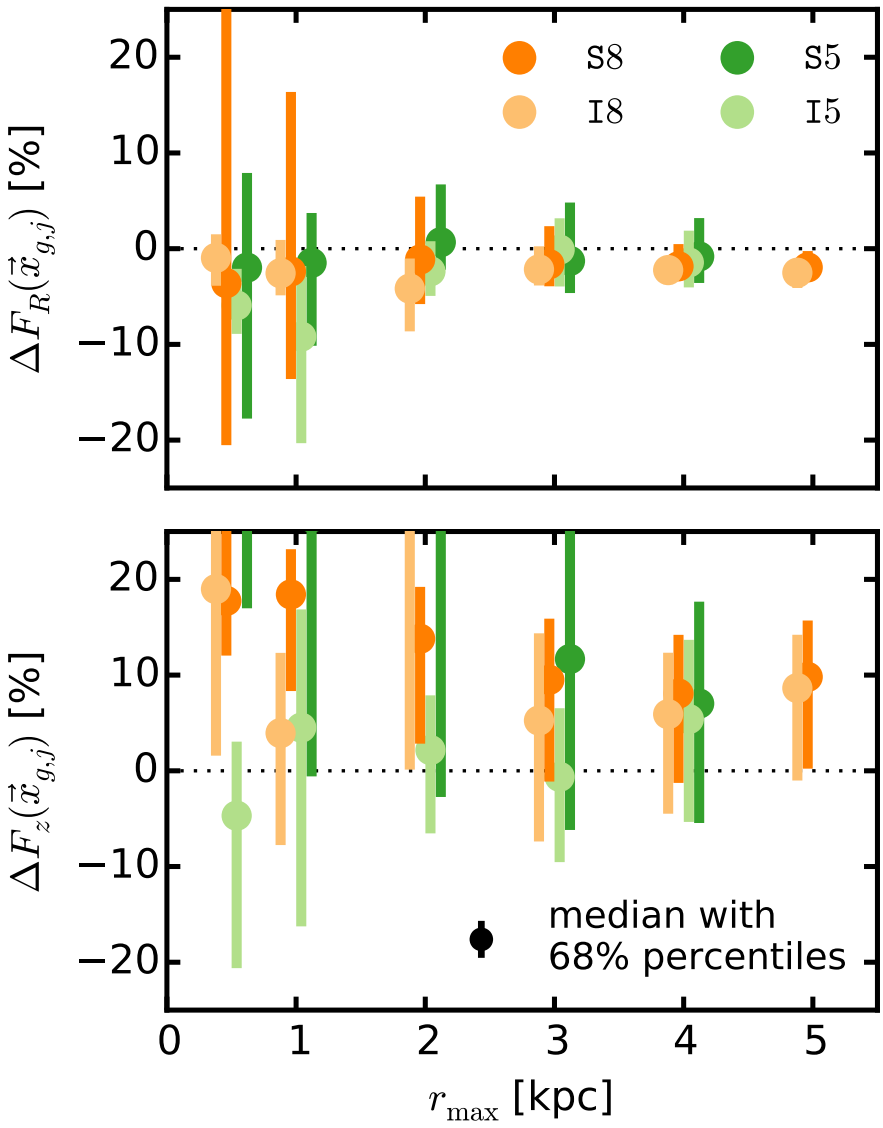


Figure 3.15: Extrapolability of a RoadMapping potential model. This figure is similar to Figure 3.13 and shows the radial (upper panel) and vertical (lower panel) gravitational force residuals. However, instead of calculating the residuals at the positions of the stars, we evaluated them here at regular grid points $\boldsymbol{x}_{g,j}$ in a large cylinder of $r_{\max} = 5$ kpc and height $|z| = 1.5$ kpc around each survey volumes' center: $\Delta F_R(\boldsymbol{x}_{g,j})$ and $\Delta F_z(\boldsymbol{x}_{g,j})$ in Equations (3.25) and (3.26). This demonstrates how well the model can be extrapolated out to 5 kpc, i.e., how close to the truth the corresponding volume-averaged model is. The x-axis shows the radial size $r_{\max}$ of the survey volume from which the model was derived. The dot and error bars denote the median and $[16th,84th]$ percentile range for each distribution of $\Delta F(\boldsymbol{x}_{g,j})$. The extrapolability works better for the radial forces than for the vertical forces. We account the systematic overestimation of the vertical forces to the spiral arms and the flaring of the disk model at large radii (see Section 3.4.2.6). We find that we need at least a survey volume of $r_{\max} = 3$ kpc to get a potential with a reasonable extrapolability.

First, we find that the radial forces are overall very well predicted, especially when derived from large survey volumes. There is, however, an overestimation of $\sim 5 - 20\%$ in the vertical forces (depending on volume size and position), which is induced by the spiral arms and is partly also due to a systematic error caused by the choice of potential model (see the explanation below in

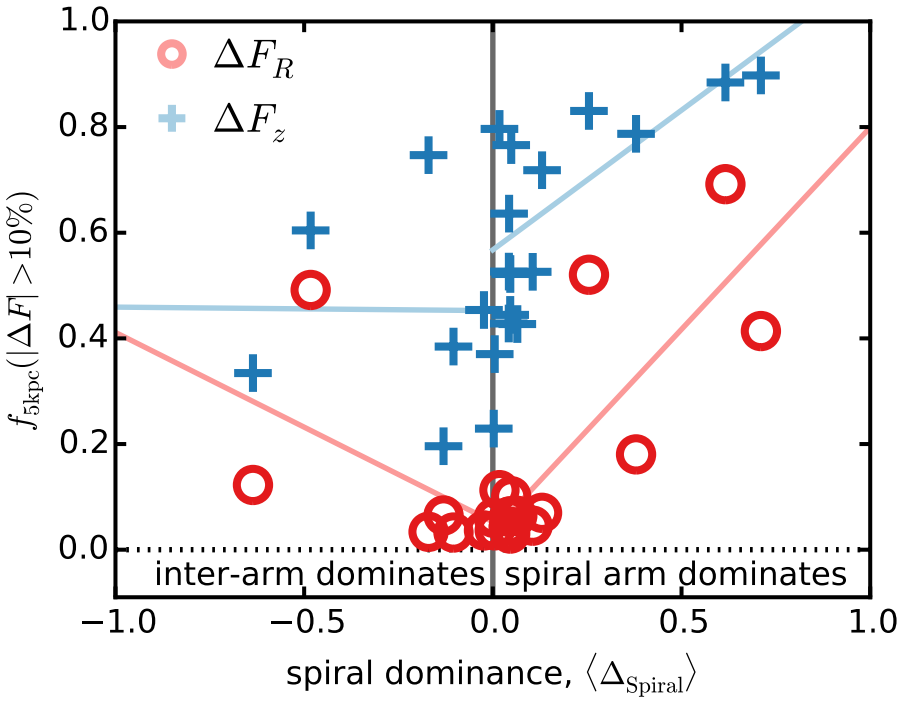


Figure 3.16: Influence of the spiral arm dominance on the extrapolability of the RoadMapping potential models. Each circle/cross pair corresponds to one of our 22 RoadMapping analyses. How much a spiral arm dominates within a given survey volume is quantified by $\langle \Delta_{\mathrm{Spiral}} \rangle$ from Equation (3.5) and shown on the x-axis. $\langle \Delta_{\mathrm{Spiral}} \rangle \sim 0$ means spiral arms and inter-arm regions dominate equally in the survey volume. The more negative $\langle \Delta_{\mathrm{Spiral}} \rangle$ is, the more dominated by an inter-arm region the survey volume is. The y-axis shows the volume fraction within which the radial (red circles) and vertical (blue crosses) gravitational force residuals are larger than 10%. In particular, we extrapolate each potential model within a cylindrical volume with $r_{\mathrm{max}} = 5$ kpc and $|z| \leq 1.5$ kpc, centered on the respective survey volume positions given in Table 3.1. $f_{5\mathrm{kpc}}(|\Delta F| > 10\%)$, therefore, quantifies how well a RoadMapping potential model can be extrapolated out to 5 kpc from the center of the survey volume; at $f_{5\mathrm{kpc}}(|\Delta F| > 10\%) \sim 0$, the extrapolation works best. The red and blue lines are linear fits and only serve as a guide to the eye. This figure demonstrates that the extrapolability of the model gets better the less a spiral arm dominates the data set. For the dominance of inter-arm regions, this trend is less pronounced. Overall, the radial force can be much better extrapolated than the vertical force, and models from data sets centered on inter-arm regions can be more reliably extrapolated than those centered on spiral arms.

Section 3.4.2.6).

Second, the constraints we get on the spatially averaged forces inside $r < 5$ kpc are almost as good when derived from a survey volume of $r_{\mathrm{max}} = 3$ kpc as compared to survey volumes of $r_{\mathrm{max}} = 4$ or 5 kpc. If we had to decide between a $r_{\mathrm{max}} = 3$ kpc volume with good data quality and a larger volume with worse data quality, we would lose nothing in terms of extrapolability when using the smaller volume (only the halo scale length might not be as well constrained, see Figure 3.12).

Third, there are further indications in Figure 3.15 that the position of the survey volume with respect to the spiral arms matters for the force recovery.

In Figure 3.16 we stress that even more by relating the extrapolability, i.e., the volume-averaged distribution of force residuals, $\Delta F_R(\boldsymbol{x}_{g,j})$ and $\Delta F_z(\boldsymbol{x}_{g,j})$, to the dominance of the spiral arm $\langle \Delta_{\mathrm{Spiral}} \rangle$ in the survey volume (Equation (3.5); see also Figure 3.4(c)). We derive the fraction of grid points $\boldsymbol{x}_{g,j}$ in the reference cylinder for each data set/RoadMapping model with forces that

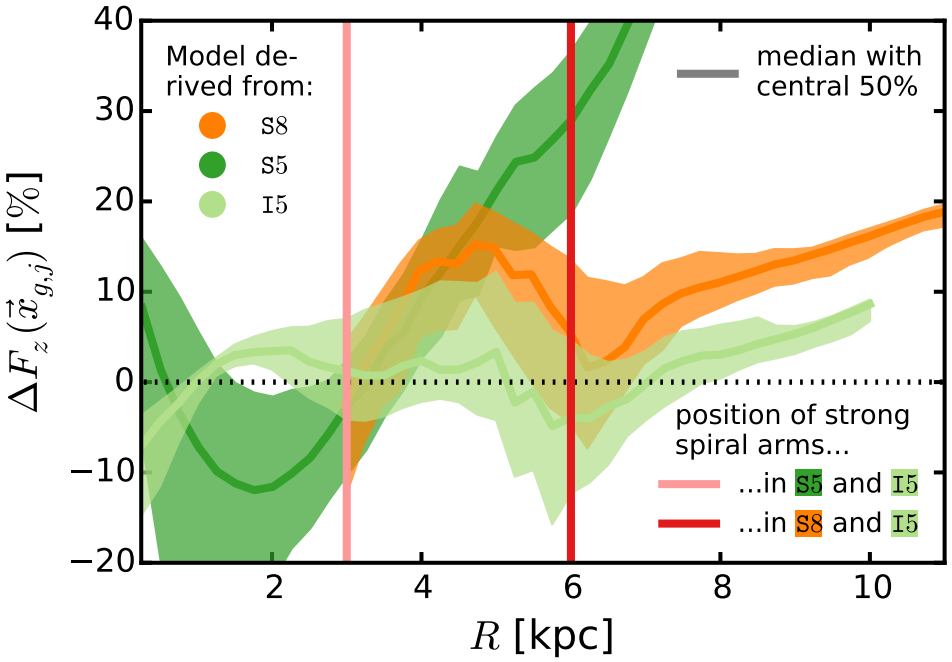


Figure 3.17: Examples for the extrapolability of a RoadMapping gravitational potential model as a function of galactocentric radius, R. In particular, we show the vertical force residuals $\Delta F_z(\boldsymbol{x}_{g,j})$ from Equation (3.26) for the RoadMapping potential models derived from three data sets with $r_{\max} = 2$ kpc at positions `S8`, `S5`, and `I5` (color-coded; see also Figure 3.1). The colored bands show the distribution (median with the central 50% percentiles) of ΔF_z's of all regular grid points $\boldsymbol{x}_{g,j}$ within a distance $r \leq 5$ kpc and $|z| \leq 1.5$ kpc from the survey volumes center at a given R. This figure demonstrates the origin of the bias in the vertical force prediction, which we found in Figure 3.15. Because the fit is driven by the excess of stars that feel stronger vertical forces in the spiral arms, e.g., at $R \sim 3$ kpc or $R \sim 6$ kpc, we get the vertical force right at these radii. We consequently overestimate it in the inter-arm regions. In addition, the assumed gravitational potential model family, `MNHH-Pot` with a Miyamoto-Nagai, disk, flares outside of $R \sim 8$ kpc as compared to the true exponential disk.

are misjudged by more than 10%, and plot it against $\langle \Delta_{\text{Spiral}} \rangle$. Positive or negative $\langle \Delta_{\text{Spiral}} \rangle$ quantify how much the spiral arms or the inter-arm regions dominate the corresponding survey volume, respectively. A low fraction of grid points with $|\Delta F| > 10\%$ means that the extrapolation of the potential model works well.

First, we note that there is a clear trend that the extrapolability gets worse if a spiral arm strongly dominated the survey volume from which the potential constraint was derived ($\langle \Delta_{\text{Spiral}} \rangle \gg 0$). The same trend can be seen for the dominance of inter-arm regions ($\langle \Delta_{\text{Spiral}} \rangle \ll 0$), but it is weaker and less clear.

Secondly, we note again that $F_z(\boldsymbol{x}_{g,j})$ is not predicted as well as $F_R(\boldsymbol{x}_{g,j})$. The reason for this is laid out in Section 3.4.2.6.

The main result of Figure 3.16 is, however, the following. The extrapolability of models derived from data sets drawn from survey volumes centered on inter-arm regions appears to be in general better than that of data sets centered on spiral arms. We suspect that the reason for this is that the stellar distribution between spiral arms is smoother, more extended, and closer to the overall axisymmetric average model, such that the potentials recovered from these volumes have real predictive power for a much larger volume.

3.4.2.6 Biases in the Potential Recovery caused by the Spiral Arms

What are the reasons for the biases that we observe in Figure 3.15?

The peak of the distribution in $\Delta F_R(\boldsymbol{x}_{g,j})$ (and also in $\Delta F_R(\boldsymbol{x}_{*,i})$ in Figure 3.13) is slightly ($\sim 1.5\%$) biased toward an underestimation of $|F_R|$ in our RoadMapping models. This bias already showed up in the circular velocity curve in Figure 3.6 and also in the reference potential model `DEHH-Pot`. We therefore suspect that this bias is caused by the spiral arms. The line of argument goes as follows. Spiral arms are very thin. If a spiral arm crosses the observation volume, both its leading side (at large radii) and its trailing side (at small radii) are also in the volume. Stars on the trailing side feel a lower gravitational pull toward the galaxy center than they would if there was no spiral arm. Because there are, in general, more stars at smaller radii, the RoadMapping fit is slightly biased to reproduce slightly weaker radial forces.[21]
The peak in the distribution of $\Delta F_z(\boldsymbol{x}_{g,j})$ is strongly biased toward an overestimation of $|F_z|$ by the RoadMapping model. We illustrate the reason for this in Figure 3.17, where we show how $\Delta F_z(\boldsymbol{x}_{g,j})$ varies as a function of R for three example data sets with $r_{\max} = 2$ kpc. One reason for this bias is a relic of the assumed `MNHH-Pot` disk model family, which flares outside of $R \sim 8$ kpc (see Section 3.3.3). The vertical forces in this region are therefore much higher in the model than in the true galaxy with its exponential disk. The main reason for this overestimation of F_z, however, comes directly from the spiral arms: there are much more stars in the spiral arms than in the inter-arm regions, and the stars in the spiral arm feel stronger vertical forces because of the higher surface mass density. The RoadMapping fit is driven by the majority of stars and the best-fit model therefore predicts, in general, higher vertical forces. As expected, the overestimation of F_z in Figure 3.15 is especially strong ($\sim 20\%$) for small survey volumes dominated by spiral arms, while small volumes dominated by an inter-arm region result in much better estimates for the spatially averaged $F_z(\boldsymbol{x}_{g,j})$ ($\sim 5\%$ bias). Large volumes lie somewhere in between (bias of $\sim 10\%$).
Why do the biases show up in different strengths in Figures 3.13 and 3.15?
The stellar number asymmetry in the trailing versus leading sides of spiral arms is much smaller than the stellar number asymmetry in the spiral arm versus the inter-arm region. The bias is therefore visible in the distribution of $\Delta F_R(\boldsymbol{x}_{*,i})$ (because the F_R recovery is biased only by a few stars, which leads to a bias that is visible for the majority of stars) and not in $\Delta F_z(\boldsymbol{x}_{*,i})$ (because the majority of stars bias the fit and we therefore also recover F_z for the majority of stars). The bias becomes particularly pronounced for $\Delta F_z(\boldsymbol{x}_{g,j})$ (because the inter-arm regions dominate when averaging spatially, which leads to a large average overestimation of F_z) and stays small for $\Delta F_R(\boldsymbol{x}_{g,j})$ (because trailing and leading sides of spiral arms are similarly important when averaging spatially; it therefore becomes visible that the bias is actually not that big).

3.5 Discussion and Summary

3.5.1 On the Informativeness of an Orbit Distribution Function

The qDF appears to be very informative. We did expect it to be at least a reasonable model for the overall symmetrized disk of the galaxy simulation, considering its initial setup as an axisymmetric, exponentially decreasing particle distribution that subsequently evolved as a mono-age population (see Sections 3.2.1 and 3.3.2). In Sections 3.4.1.1, 3.4.1.3, and 3.4.1.4, we demonstrated that the qDF is indeed a good average model for the tracer distribution in a large survey volume—even though the spiral arms did introduce considerable deviations.
We had, however, no indications beforehand of how well the axisymmetric qDF would perform in a small survey volume completely dominated by non-axisymmetric spiral arms. It would not have been surprising if RoadMapping had failed. However, in Section 3.4.2.2, it turned out that the

[21] In the special case that the survey volume coincidentally only contains part of a spiral arm—as was the case with the analyses for $r_{\max} = [0.5, 1]$ kpc at position `I5` (see Figure 3.10, right panels)—the fitting behaves differently anyway, as was already discussed in Section 3.4.2.3.

potential measurements were reliable even in most of the small volumes with $r_{\max} = [0.5, 1]$ kpc. Furthermore, the corresponding qDF parameters were tightly constrained and reasonable as well. We deduce that the qDF is indeed flexible and robust enough to work with data affected by non-axisymmetries.
That the corresponding potential constraints were reliable as well leads to the following conclusion: a potential model that does not fit the gravitational forces acting on the stars appears to lead to such an unrealistic orbit/action distribution, that a fit with even such a simple orbit DF as the qDF is impossible. This demonstrates once more how powerful the concept of an orbit DF is.

3.5.2 On the Restrictiveness of the Parametrized Potential Model

How much does the choice of potential model matter for the success of the modeling?
We used, on the one hand, a bulge and halo model that reproduces the true bulge and halo better than we can hope to use in reality for the MW. The fact that for all except the largest volumes the true halo scale length is not remotely recovered (and some small volumes even have $f_{\rm halo} = 0$), and that the contribution of the bulge to the overall potential is small (i.e., the bulge contribution to the total radial force at $R = 8$ kpc is only $\sim 9 - 10\%$) remedies this apparent advantage.
On the other hand, we used a disk model, the Miyamoto-Nagai disk, that we chose purely for its convenient parametric form and of which we know that it is not a good model for the simulation snapshot; especially not for the radial density profile at large radii (Smith et al. 2015). As we saw in most figures in this chapter, this leads to biases in predicting the potential at radii where we have only a few or no stars. However, because the spiral arms are such strong perturbations in the overall potential, a better disk model would not give much better results.
It appears that a potential model with a reasonable shape and flexibility (here, disk+bulge+halo structure with five free parameters) can do well enough in finding a good fit, both locally for small volumes and overall for large volumes.
This is in agreement with one of our key results of Chapter 2, Section 2.3.6, where we managed to successfully fit data from an MW-like (but axisymmetric) galaxy model with a bulgeless potential of a restrictive Stäckel form. This was illustrated in Figure 2.18. Considering that we used there the same number of stars, $N_* = 20,000$, the potential uncertainties were much greater than in the analogous figure of this chapter, Figure 3.6(a). We believe that this is how RoadMapping accounts for an inconvenient potential parametrization—by increasing the uncertainty of the model estimate—which is exactly as it should be.

3.5.3 *Gaia* Measurement Errors and Choosing the Survey Volume Size

Considering measurement uncertainties of distances and proper motions, we found in Chapter 2, Section 2.3.4, that for a survey volume with $r_{\max} = 3$ kpc, distance uncertainties of $< 10\%$ and proper motion uncertainties of less than 3 mas yr^{-1} RoadMapping still gives unbiased parameter results. Even if the proper motion errors are not perfectly known.
The measurement uncertainties of *Gaia* in proper motions (already in the first data release $\delta\mu \sim 1$ mas yr^{-1}; Lindegren et al. 2016) and in distances (at least for the final data release within 3 kpc and for bright stars; de Bruijne et al. 2014) lie below these limits.
In Chapter 2, we focused on recovering completely unbiased model parameters and found that RoadMapping is robust against moderate deviations of the model assumptions. In this chapter, we released the condition that the model parameters themselves had to be recovered accurately, but allowed RoadMapping to simply find an overall best fit for the data strongly affected by spiral arms—which was surprisingly successful in recovering the local potential even if the model parameters were not recovered.

We therefore presume that, in reality, we probably have an even larger margin of error than we found in Chapter 2, and before the measurement uncertainties noticeably muddle the constraints.
In addition, we found in this chapter that a volume of $r_{\max} = 3$ kpc should already be big enough to find an overall best-fit axisymmetric model for the Galaxy. At larger distances, dust starts affecting the measurements. Furthermore, inside of $R = 3 - 4$ kpc, the stellar motions become increasingly non-axisymmetric, possibly because of the Galactic bar (e.g., Reid et al. 2014; Bovy et al. 2015, see also Section 1.2.1).
Overall, we should therefore be very well-off by applying RoadMapping to the final *Gaia* data set within $r_{\max} = 3$ kpc only.
How well we can do with the first few *Gaia* data releases remains to be seen. The *Gaia* DR1 from September 2016 has parallax measurements of $\sim 16\%$ for red-clump giants at a distance of ~ 500 pc from the Sun (de Bruijne et al. 2014; Michalik et al. 2015a), which is not precise enough for RoadMapping. We could, however, use photometric distances, which should be precise enough for red-clump stars and even extend over a larger volume. The *Gaia* DR2 in April 2018 might, however, already cover ~ 2 kpc with good parallaxes for fainter stars as well.

3.5.4 Spiral Arms in the Solar Neighborhood

The Sun is located in one of the smaller spiral arms of the MW, the local Orion spur/arm (Morgan et al. 1953). Two of the MW's major spiral arms pass by the Sun within a few kiloparsecs: the Perseus arm is ~ 2 kpc from the Sun (toward the outer MW; Xu et al. 2006), and the Sagittarius arm at ~ 1 kpc (toward the Galactic center; Sato et al. 2010). It is, however, still under dispute which arms are actually major arms of the MW (Blaauw 1985; Xu et al. 2013; Zhang et al. 2013a).
How reliable the RoadMapping results from the *Gaia* DR1 will be depends on the strength of the local Orion arm, which will dominate the survey volume. RoadMapping had, for example, some difficulties recovering the circular velocity curve for small volumes (see Figure 3.10).
However, recent measurements of the MW's rotation curve (Bovy et al. 2012a; Reid et al. 2014) confirm again that it is flat. We could impose this condition as a prior constraint in RoadMapping (or fix the rotation curve slope as Bovy & Rix 2013 did in their RoadMapping analysis). Based on independent information on the location and strength of the MW spiral arms, we could also use the present approach to estimate systematic uncertainties.
For later *Gaia* data releases, where the tracers extend further into the Galaxy (~ 2 kpc from the Sun), the Sagittarius and Perseus arms could also play a role. In general, RoadMapping should do much better with larger volumes, and if several spiral arms and inter-arm regions within the survey volume average each other out. This averaging should work especially well if the MW is—like the simulation in this study—a four-armed spiral. The exact number of spiral arms, however, is still disputed (see Section 3.2.5). If *Gaia* should show that the MW is a two-armed grand-design spiral instead, having a very large survey volume would become even more important to achieve a good axisymmetric average RoadMapping model for the MW. Local measurements of the potential as in Figure 3.10 should, however, still be possible.

3.5.5 Absence of a Central Bar in the Simulation

The simulation we analyzed in this chapter does not have a prominent bar, and so we have not explicitly explored the impact of such a feature. Bars can play an important role in the dynamics when very small volumes near a resonance are considered (see, e.g., Dehnen 2000, and references in Section 3.2.5). When considering volumes of $\gtrsim 1$ kpc, we have no reason to believe that this should severely affect the robustness of such an analysis.

3.5.6 Interpreting RoadMapping Results

The two main findings of this chapters give a clear directive regarding how we should deal with any future result about the MW's potential derived with RoadMapping. (1) If the data spans a large volume, a significant proportion of the disk (at least $R \sim 5 - 11$ kpc), and averages over several spiral arms and inter-arm regions, we can trust and use the resulting model as an overall axisymmetric potential for the MW. For $R \sim 3 - 13$ kpc, we should even be able to make definite statements about the DM distribution. (2) If the data do not span such a large volume, we can still believe the local constraints. In particular, the surface density within one to two disk scale heights, the circular velocity within the survey volume, and the average gravitational forces where the majority of stars is located.
Fortunately, this is consistent with the procedure by Bovy & Rix (2013), who used RoadMapping to constrain the vertical force F_z for each MAP at only one radius that corresponds to a typical radius. It will be interesting to see whether RoadMapping potential constraints from *Gaia* data will agree with their findings.
In addition, one should always compare the distribution of the data and the recovered model in configuration space $(\boldsymbol{x}, \boldsymbol{v})$ and in action space, as we did in Section 3.4.1. This is not only a sanity check to confirm the goodness of the fit, but it might also reveal some substructure in the data that only becomes visible when comparing it to an axisymmetric smooth model.
The overall axisymmetric best-fit MW potential and disk DF from RoadMapping could subsequently be also used as zero-th order potential and DF for a perturbation-theory analysis of the response of the disk to spiral structure (Monari et al. 2016).

3.5.7 Summary

In this chapter, we investigated the robustness of RoadMapping when modeling a non-axisymmetric system. We explore this for the first time explicitly, by modeling a simulated spiral galaxy from D'Onghia et al. (2013), and by comparing the results to the true potential. This simulation has stronger spiral arms than we expect in the MW, and—except for the absence of a bar and thick and gas disk components—it has matter components similar to the MW, as discussed in Section 3.2.5. We find that RoadMapping-like action-based dynamical modeling is very robust against perturbations of spiral arms in this simulation, especially if the survey volume is large enough to encompass both spiral and inter-arm regions. In Section 3.4.1, we demonstrated this in detail for a single RoadMapping analysis of a data set with a spatial coverage of radius $r_{\max} = 4$ kpc around a position equivalent to that of the Sun.
In Section 3.4.2, we have investigated the role of survey volumes differing in size and position with respect to the spiral arms and inter-arm regions within the simulated galaxy. We find that the gravitational forces are mostly well recovered at the locations of the stars that entered the analysis. For survey volume sizes $r_{\max} \geq 3$ kpc, the recovered potential model already becomes a good average potential model for a large portion of the galaxy. For some positions of the survey volume center, e.g., in a smooth and not-too-depleted inter-arm region, smaller volumes can also give good overall constraints. If a small volume is dominated by a very strong spiral arm the constraints become less reliable, as expected. The correct DM halo scale length was, however, only recovered for a survey volume as large as $r_{\max} = 5$ kpc.
This overall robustness of RoadMapping is particularly notable because the breakdown of the assumption of axisymmetry implies a breakdown of several model assumptions simultaneously: (1) orbital actions are not fully conserved anymore, (2) the true potential is not spanned by the family of model potentials, (3) the qDF need not, or will not, describe the orbit distribution within spiral arms. However, the qDF seems to be informative enough to guide the fit to potential shapes that correctly measure the average surface density (within $\sim 2\times$ the disk scale height) and the circular velocity where most of the stars that entered the analysis are located—even for

small volumes with $r_{\max} = 500$ pc dominated by spiral arms.

3.5.8 Outlook

The results of this chapter imply that RoadMapping should be well-suited to making new measurements of the MW's gravitational potential with the upcoming *Gaia* data releases. It might even potentially work with the *Gaia* DR1 with its smaller coverage of the disk ($r_{\max} \sim 1$ kpc) because the local Orion arm, in which the Sun is located, is thought to be only a minor spiral arm in the MW and should not significantly disturb the modeling.
The next chapter is therefore dedicated to putting RoadMapping to test on data from *Gaia* DR1.

CHAPTER 4

RoadMapping Modeling of the Milky Way with Stars from *Gaia* and RAVE

4.1 Preface

Introduction. On September 14, 2016, the first *Gaia* data release provided the largest and most precise catalog of stellar proper motions to date. After our substantial improvements to the RoadMapping machinery in Chapter 2 and the elaborate tests of its robustness in both Chapter 2 on mock data, and in Chapter 3 on data from a spiral galaxy simulation, it was finally time for the first application of RoadMapping to real data in the MW disk since Bovy & Rix (2013). Our goal is to get new and precise measurements of the Galactic gravitational potential from action-based dynamical modeling.
The largest challenge of this application is (a) to piece together a data set from different surveys to get 6D phase-space coordinates and 2D chemical abundances of high-enough precision for a large amount of stars, and (b) to setup the SF of this stellar sample that mirrors the selection effects of the different surveys while being simple enough and having the correct functional form to be used in RoadMapping.

Structure of the chapter. In Section 4.2, we give an overview of the surveys that provided the data and explain how we put together a suitable data set for RoadMapping. We set up a SF for this data set as will be laid out in Section 4.3 and have a first look at the data properties in Section 4.4. In Section 4.5, we summarize the special modeling ingredients that we use in the RoadMapping analysis of the MW data. Section 4.6 presents all our results on the MW gravitational potential and the phase-space structure of data depending on abundance. And finally, Section 4.7 discusses several aspects of the modeling and the results, and provides an outlook and summary.

Attribution. My main accomplishments in this chapter are:

- Making the first full action-based dynamical modeling attempt of *Gaia* data, which gave constraints of high precision on the MW gravitational potential parameters and offered insight into the chemo-orbital structure of the Solar neighborhood.
- The idea how to create a sufficiently complex SF for RC stars in TGAS and RAVE that can be used in RoadMapping.

- piecing together a consistent data set with well-defined SF for RoadMapping with information from TGAS, *Gaia* DR1/secondary data set, RAVE, RAVE-on, APOGEE-RC and 2MASS.
- Major changes and improvements in the RoadMapping machinery, which now can use real data with covariant measurement errors, complex realistic SFs, different DFs, outlier models, and priors on the model parameters.
- Testing the code framework and different aspects of the modeling to confirm that the result is robust.
- Creating all content, tests, text, and figures in this chapter (with four exceptions for illustration purposes, stated in the respective figure captions).

In addition, my collaborators helped with the following:

- **Hans-Walter Rix (MPIA, Heidelberg)** had the idea to complement TGAS with the RAVE and RAVE-on catalogs, and to use RC stars to circumvent the distance precision problem. He also suggested the SF test in Section 4.7.1.
- **Jo Bovy (Uni Toronto)** contributed ideas to setting up the RoadMapping machinery for the *Gaia* data, helped with estimating the RC G-magnitude, and the total surface mass density from the Jeans equation. He also provided the code to estimate the distances for RC stars, and very helpful feedback on a draft of this chapter.
- **Georges Kordopatis (OCA, Nice)** provided the cross-matched catalog between TGAS, RAVE, and RAVE-on, including extinction estimates for each star, and helped with the RAVE quality flags.
- **Jennifer Wojno (AIP, Potsdam)** gave advice on the RAVE SF and provided the 2MASS stellar number counts.

I also thank **Ivan Minchev**, **Friedrich Anders (AIP, Potsdam)**, and **Simon White (MPA, Garching)** for helpful discussions on some aspects that found their way into this work.
This project was developed in part at the 2016 NYC Gaia Sprint, hosted by the Center for Computational Astrophysics at the Simons Foundation in New York City.
This work has made use of data from the European Space Agency (ESA) mission *Gaia* (`https://www.cosmos.esa.int/gaia`), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, `https://www.cosmos.esa.int/web/gaia/dpac/consortium`). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.

4.2 The Data

To perform a RoadMapping analysis of the MW, we require (i) 6D phase-space positions for a large ensemble of stars, including precise distances (see Sections 2.3.4 and 2.4.2), (ii) chemical abundance measurements to sort the stars into MAPs, and (iii) a well-defined SF. So far there is no survey that provides all of this. Luckily, with the first *Gaia* data release in 2016, we now have for the first time very precise proper motions for a large set of stars in the solar neighborhood. In combination with the spectroscopic survey RAVE, we managed to put together a suitable data set of RC standard candles for RoadMapping. In the following, we summarize the survey goals of *Gaia* and RAVE, introduce the data characteristics, and motivate our choices for the data selection.

4.2.1 Proper Motions from *Gaia*-TGAS

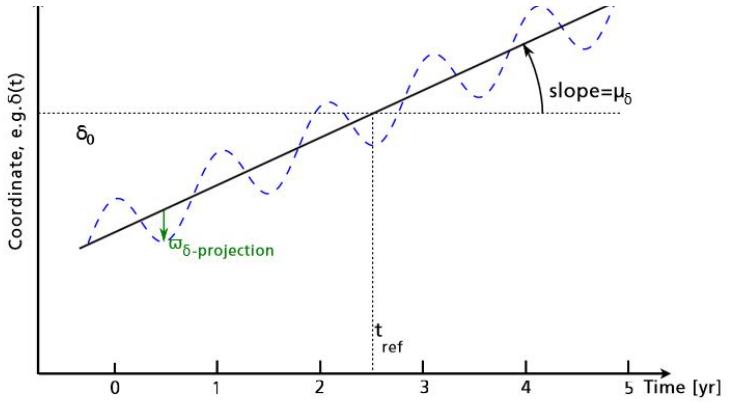
Survey overview: The *Gaia* mission. *Gaia* is an astrometric cornerstone satellite mission of the European Space Agency (ESA) (Perryman et al. 2001; de Bruijne 2012; Gaia Collaboration et al. 2016b) and the successor of Hipparcos (High Precision Parallax Collecting Satellite) (ESA 1997). The *Gaia* satellite was launched on December 19, 2013, and is now located at the Sun-Earth Lagrangian point L2. Its nominal operation will end in 2019. The *Gaia* telescope observes 1.1 billion stars of the MW while spinning and slowly precessing and continuously scanning the sky perpendicular to the spin axis. During the scan, the stars are transiting over the focal plane of the telescope, which contains different photometric and spectroscopic instruments (Gaia Collaboration et al. 2016b).

Ultimately, in ~ 2022, the astrometric instrument of *Gaia* will have recorded the position of each star about 70 times over the baseline of 5 years. This allows the calculation of (R.A., Dec.) positions at epoch J2015.0, proper motions ($\mu_{\mathrm{R.A.}*} \equiv \mu_{\mathrm{R.A.}} \times \cos(\mathrm{Dec.}), \mu_{\mathrm{Dec.}}$) and stellar parallaxes (ϖ) with high precision, $\sim 5 - 25$ μas (for $V < 15$ mag; ESA 2017b), for a large portion of the stellar content of the MW. The astrometric instrument also provides white-light magnitudes in the G-band ($330 - 1050$ nm) (Jordi et al. 2006, 2010). The first 5-parameter astrometric measurements together with G magnitudes were released on September 14, 2016, for stars with $2 - 3$ mag $\lesssim G < 20.7$ mag and typical position uncertainties of ~ 10 mas (Gaia Collaboration et al. 2016a; van Leeuwen et al. 2017; Lindegren et al. 2016).

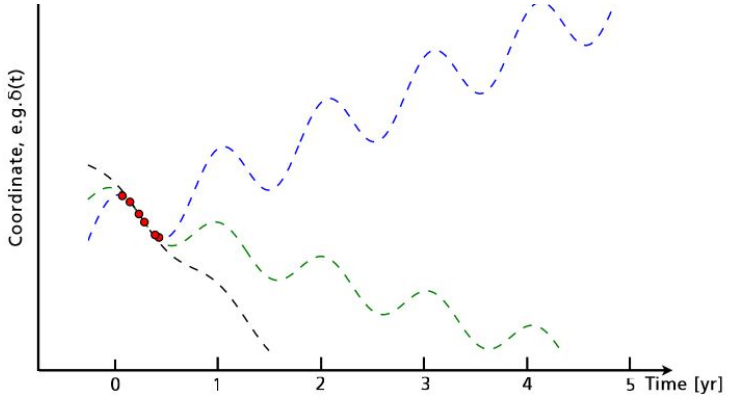
Gaia also has two spectro-photometric instruments on board to record integrated apparent magnitudes and low-resolution spectra in two band passes (blue: 330-680 nm, G_{BP}-band; red: 640-1000 nm, G_{RP} band) (Jordi et al. 2010). This will provide magnitudes and colors for all stars, and help to determine astrophysical parameters like effective temperature, surface gravity, age, mass, chemical composition, extinction etc. The first spectro-photometric results will be published in April 2018 (ESA 2017a).

The third instrument is the RVS (Katz et al. 2004) which takes spectra in the wavelength range $\lambda = 848 - 874$ nm with resolving power $R \sim 11,500$, to measure radial velocities of more than 100 million stars down to $G_{\mathrm{RVS}} \lesssim 11 - 14$ mag. These spectra will also contribute to the determination of stellar parameters. Median radial velocities will be made publicly available in April 2019; a full catalog of radial velocities and stellar parameters only in the subsequent data releases.

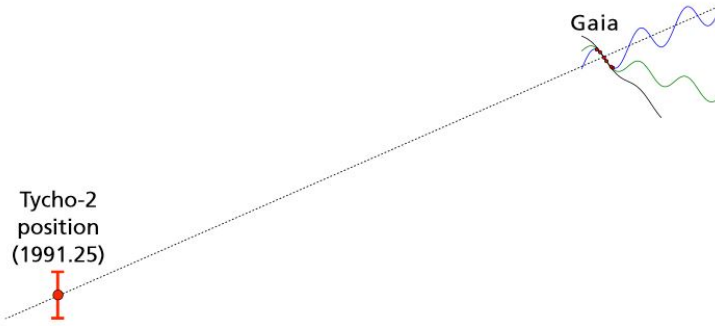
Catalog description: The Tycho-Gaia Astrometric Solution. The first *Gaia* DR from September 2016 was compiled from 14 months of *Gaia* observations and therefore suffered—in addition to lower precision due to less visits per star—from degeneracies between stellar parallax and proper motions, as illustrated in Figures 4.1(a) and 4.1(b). François Mignard proposed in 2009 the Hundred Thousand Proper Motions (HTPM) project (Mignard 2009; Michalik et al. 2014) and suggested to use the high-precision position measurements of Hipparcos from ~ 24 years ago to resolve the $\mu - \varpi$-degeneracy (see Figure 4.1(c)). The HTPM project evolved into the Tycho-Gaia Astrometric Solution (TGAS) Michalik et al. (2015a,b), which used, in addition to the 118,000 stars from the Hipparcos catalog (van Leeuwen 2007), also ~ 2.5 million stars from the *Tycho-2* catalog (Hog et al. 2000) with slightly lower precision in position measurements. *Gaia* DR1 now provides a reliable 5-parameter astrometric solution ($\mathrm{R.A.}, \mathrm{Dec.}, \mu_{\mathrm{R.A.}*}, \mu_{\mathrm{Dec.}}, \varpi$) for ~ 2 million *Tycho-2* stars ($G < 12$ mag) and $\sim 90,000$ Hipparcos stars with $G < 12$ mag, called the *primary astrometric data set* of *Gaia* DR1. Typical uncertainties are ~ 0.3 mas in the positions, ~ 1 mas yr^{-1} in the proper motions and ~ 0.3 mas (plus 0.3 mas in systematics) in the parallaxes (Gaia Collaboration et al. 2016a). In addition, there are strong correlations between the stellar coordinates—especially between the proper motions and the parallaxes—which can reach ± 1.



(a) Change of a star's position on the sky over the course of five years.



(b) Degeneracy between μ and ϖ for observations of less than a year.



(c) Solving the degeneracy using *Tycho-2* positions.

Figure 4.1: Illustration of the idea behind the Tycho-Gaia Astrometric Solution (TGAS) (Michalik et al. 2015a,b). Panel 4.1(a) shows how the apparent position (for example the declination Dec. = δ) of a star changes over the course of five years of *Gaia* observations. The mean slope is due to the intrinsic proper motion μ_δ of the star, while the apparent annual variations are due to the (projection of the) stellar parallax ϖ_δ. δ_0 marks the position of the star at a reference epoch $t_{\rm ref}$, for *Gaia* DR1 it is J2015.0. Panel 4.1(b) demonstrates that for observations of less than a year, there could be different combinations of $(\varpi_\delta, \mu_\delta)$ that plausibly explain the observed positions. Panel 4.1(c) finally illustrates how adding the single position of the *Tycho-2* catalog as a prior is sufficient to solve the degeneracy between proper motion and parallax. **Figure credit:** L. Lindegren & D. Michalik (Michalik 2016)

Phase-space coordinates for RoadMapping. The proper motions in TGAS have a precision similar to the smallest μ uncertainties that were previously possible to achieve with ground-based surveys (see Section 2.3.4). The TGAS proper motions and positions will therefore be very useful for our RoadMapping analysis.
A typical parallax error of 0.3 mas in TGAS would translate to a relative error of $\sim 15\%, 20\%, 30\%$ at distances of $\sim 0.5, 0.7, 1.0$ kpc from the Sun. This is way too large for RoadMapping to work (see Section 2.3.4). In Section 4.2.4 we will therefore use the more precise distances to RC standard candles as an alternative to the TGAS parallaxes.

4.2.2 Radial Velocities from the RAVE Survey

Motivation. *Gaia* will not provide radial velocity measurements before DR2 in April 2018. To have full 6D $(\boldsymbol{x}, \boldsymbol{v})$ coordinates for all TGAS stars, we need to combine the TGAS (R.A., Dec., ϖ, $\mu_{\mathrm{R.A.}*}$, $\mu_{\mathrm{Dec.}}$) with line-of-sight radial velocities v_{los} measured by a spectroscopic survey. In addition, a suitable spectroscopic survey would also need to provide chemical abundances, specifically [Fe/H] and [α/Fe], to be able to sort stars into MAPs. There are several spectroscopic surveys that overlap with TGAS and which are summarized in Table 1.2.
Tycho-2, on which the TGAS catalog is based on, is a survey of the solar neighbourhood with brightness limit $V \lesssim 11.5$ mag. The largest overlap with TGAS is therefore achieved by RAVE, which targeted brighter stars than the other spectroscopic surveys. There are 255,922 individual stars in both TGAS and RAVE.

Survey description. RAVE is a large-scale spectroscopic survey of the stars in the MW conducted with the 1.2-m UK Schmidt Telescope at the Australian Astronomical Observatory (AAO). Between 2003 and 2013, RAVE took spectra of 457,588 individual stars on the Southern hemisphere (Dec < 5 deg, see Figure 4.2).
The survey's main goal was to measure radial velocities. Hence, the spectral range of $\lambda = 8410 - 8795 \mathring{A}$ was chosen, which contains the Calcium II triplet. These absorption lines are strong also in low-metallicity stellar atmospheres and are therefore well-suited to measure radial velocities via Doppler shift even at low signal-to-noise (S/N > 10) and low spectral resolution ($R \approx 7,500$ for RAVE) to an accuracy better than 2 km s^{-1}. The *Gaia* RVS will cover a similar wavelength range, however with higher resolution (Katz et al. 2004). Until now, the RAVE collaboration has published five data releases: DR1 (Steinmetz et al. 2006), DR2 (Zwitter et al. 2008), DR3 (Siebert et al. 2011), DR4 (Kordopatis & RAVE Collaboration 2014), and DR5 (Kunder et al. 2017).
One of the secondary goals of RAVE was to spectroscopically determine stellar parameters $(T_{\mathrm{eff}}, \log g)$ and chemical abundances. The covered wavelength range does indeed contain some iron-peak element and α-element atomic transitions. RAVE DR5 (Kunder et al. 2017) performed a complete re-processing of all spectra, and a re-derivation and new calibration of stellar labels. However, the stellar labels suffer from large inaccuracies due to the small spectral range and low signal-to-noise, and the chemical abundances have high uncertainties of ~ 0.2 dex.
The RAVE collaboration aimed towards a simple target selection function and focused therefore on stars with 9 mag $\lesssim I \lesssim 12$ mag. The I-band ($\lambda = 8060 \pm 1490$ $\mathring{A}$) was chosen because it overlaps with the Ca II triplet. Below a latitude of $|b| < 25$ deg an additional weak color-cut of $(J - K_s) \geq 0.5$ was imposed to target especially red giant stars in the highly extincted regions in the plane of the disk. Wojno et al. (2017) showed that this selection is kinematically and chemically unbiased.
RAVE made use of a robotic fiber positioner and the large field of view (5.7 deg) of the 6dF multi-object spectrograph, so ≤ 150 spectra could be taken simultaneously. This required a target input catalog from which, for each pointing of the spectrograph, stars were randomly selected. For DR1-3, the input catalog was pieced together from the *Tycho-2* catalog (Hog et al. 2000)

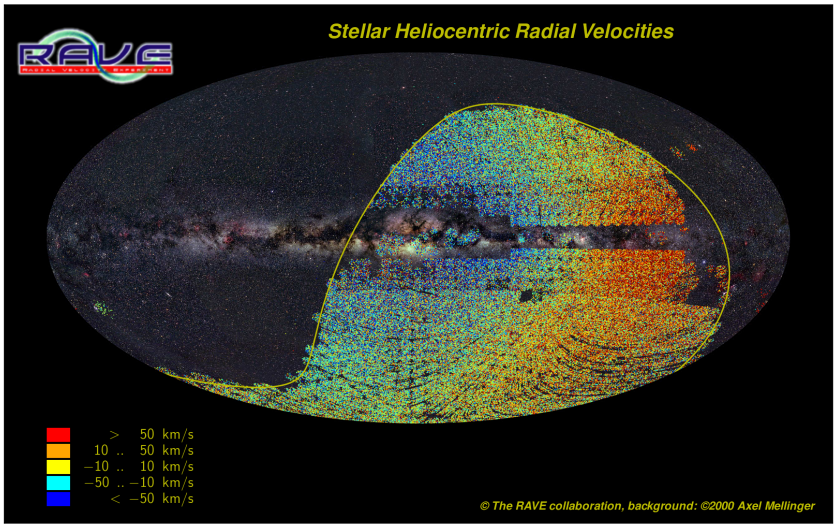


Figure 4.2: Heliocentric radial velocities of MW stars measured by the RAVE survey (DR4; Kordopatis & RAVE Collaboration 2014). The observed dipole in radial velocities is due to the Sun's peculiar motion with respect to the Galactic flow, $\Delta v \sim \sqrt{v_{\mathrm{X,GC},\odot}^2 + \left(\Omega_\odot \times R_\odot - v_{\mathrm{circ}}(R_\odot)\right)^2 + v_{\mathrm{Z,GC},\odot}^2} \sim 20$ km s^{-1} (Bland-Hawthorn & Gerhard 2016). This figure illustrates the footprint of the RAVE survey on the southern hemisphere, its location within the MW, and the incomplete sky coverage due to the pointings of the the circular fiber plate. **Figure credit:** The RAVE Collaboration (The RAVE database 2014); background image by Axel Mellinger (2000).

and the SuperCOSMOS Sky Survey (Hambly et al. 2001) (with interpolated I-band magnitudes). DR4-5 used an improved input catalog on the basis of DENIS DR3 (DENIS Consortium 2005) cross-matched with 2MASS (Skrutskie et al. 2006). Due to the random selection of stars from the input catalog by the fiber positioning algorithm, different number of pointings per field, selection from different magnitude bins with different exposure times, fiber collisions, and incomplete coverage of the sky due to circular fiber plates, the completeness of RAVE is a complicated function of the position on the sky and the I-band magnitude (see Figure 4.2). Wojno et al. (2017) estimated the completeness of RAVE therefore with respect to a parent sample, which was assumed to be complete. In particular, they used 2MASS as the parent sample.

Quality cuts. To only select stars with good quality measurements from RAVE, we pick stars for which the classification flags `c1_flag`, `c2_flag`, `c3_flag` are set to `'n'`, indicating a normal star. We further imposed some quality cuts recommended by Georges Kordopatis (OCA, Nice): The flag `algo_conv` had to be something other than `1`, which indicates that the algorithm for determining the stellar parameters converged; `delta_vlos` had to be smaller than 10 km s^{-1} to make sure that only well-determined radial velocities are used. These quality cuts might not be neccessary for a successful RoadMapping analysis: The former quality constraint is not needed if we do not use the RAVE stellar parameters, and the latter is not needed because the likelihood in RoadMapping properly takes into account the velocity uncertainties. Under the assumption that it does not depend on the 3D position of a star if its spectrum is too bad to determine good radial velocities and stellar parameters, this quality cuts will remove only random stars and not bias the modeling result.

4.2.3 Stellar Parameters and Chemical Abundances from RAVE-on

Motivation. For RoadMapping, we need [Fe/H] and [α/Fe] measurements to sort the stars into MAPs of bin widths Δ[Fe/H] = 0.1 dex and Δ[α/Fe] = 0.05 dex, analogous to Bovy &

Rix (2013); Bovy et al. (2012b,d,c). The RAVE survey provides iron and magnesium—an α element—measurements, derived from fitting spectral templates from physical models to the spectra. The abundances have uncertainties of 0.2 dex (and worse for S/N < 40 Kunder et al. (2017)). These uncertainties are too large to be used to sort the stars into MAPs if one strives to have low contamination (see also Section 2.4.3). Fortunately, Casey et al. (2017) provides the RAVE-on catalog of stellar parameters and chemical abundances for RAVE stars, which have higher precision and accuracy.

Catalog overview. Casey et al. (2017) performed a re-analysis of 520,781 RAVE DR5 spectra to derive stellar labels using a data-driven machine-learning algorithm called *The Cannon* (Ness et al. 2015, 2016). The idea behind *The Cannon* is to use a set of stars—the *training set*—that exists in both the survey A (here: RAVE DR5) of which one wishes to re-analyse the spectra, and in a different survey B, which provides reliable stellar labels for these stars. A model for the space spanned by the spectra in A is assumed, which models the normalized spectral flux at each wavelength-pixel as a second-order polynomial function of the stellar labels given by B. The coefficients of these models are fitted for in the *training step*. In the *test step*, this model with the best fit coefficients is used and independently fitted to all spectra in A, with the stellar labels now being the free parameters. In that way, stellar labels on the same scale as in survey B are assigned to the spectra in A. This approach seems to give reliable results also for spectra with low signal-to-noise, were physical models like the ones used in the RAVE stellar parameter pipeline (Kordopatis et al. 2011; Kordopatis & RAVE Collaboration 2014) have difficulties.
It was not possible for Casey et al. (2017) to find for the main-sequence stars a survey that provided at the same time good stellar labels, chemical abundances and a large enough overlap with RAVE to qualify as a training set. In the end, they used the K2/EPIC catalog (Huber et al. 2016) and fitted a three-parameter-only ($T_{\rm eff}, \log g, [{\rm Fe/H}]$) *Cannon* model to the spectra. The RAVE-on labels for main-sequence stars are therefore not very reliable.
For the RAVE giant stars, however, APOGEE from the SDSS DR13 (SDSS Collaboration et al. 2016) with its well-calibrated stellar labels ($T_{\rm eff}, \log g$) and chemical abundances (O, Mg, Si, Ca, Al, Fe, and Ni) for red giants, proved to be a good training set. Casey et al. (2017) used 536 giants with [Fe/H] = -1.79 to +0.26 dex, and supplemented the set with some well-studied low-metallicity giants. Their model for the giant stars was a second-order polynomial in 9 stellar labels (incl. 7 chemical abundances) on the APOGEE scale. Magnesium is the α-element for which the derived intrinsic uncertainties and the bias compared to stars in *Gaia*-ESO (Gilmore et al. 2012; Randich et al. 2013) turned out to be smallest. The RAVE-on labels which Casey et al. (2017) derived for the RAVE stars have intrinsic uncertainties of $\sigma_{T_{\rm eff}} \sim 80$ K, $\sigma_{\log g} \sim 0.15$ dex, $\sigma_{[{\rm Fe/H}]} \sim 0.07$ dex and $\sigma_{[{\rm Mg/H}]} \sim 0.08$ dex. This is comparable to the APOGEE uncertainties of $\sigma_{T_{\rm eff}} \sim 94$ K, $\sigma_{\log g} \sim 0.067$ dex, $\sigma_{[{\rm X/Fe}]} \sim 0.1$ dex (Holtzman et al. 2015) uncertainties for the chemical abundances. At [Fe/H]~ -4, or for hot stars, the RAVE-on labels are not reliable anymore, because the atomic transition lines become very weak.

Abundances for RoadMapping. RAVE-on stellar labels and abundances exist for 81,560 giant stars in the TGAS/RAVE cross-match. We will use [Fe/H] and the α-element magnesium, or specifically [Mg/Fe] = [Mg/H] - [Fe/H], to sort stars into MAPs for the RoadMapping analysis. The stellar labels ($T_{\rm eff}, \log g$) will be used in the following to select RC stars. We use only stars for which the RAVE-on quality flag `QC_flag==True`, which ensures that the model achieved a good χ^2 in the fit to the stellar spectrum and that the star is not very hot. There are several stars in our data set that have metallicities [Fe/H] > 0.26; because this is outside of the training set used in the *Cannon* analysis of RAVE-on, the extrapolated abundances of those stars should be taken with some grain of salt.
We consider stars to belong to the same MAPs if they fall in the same bin in the [Fe/H] vs.

[Mg/Fe] plane with bin width Δ[Fe/H] = 0.1 dex and Δ[Mg/Fe] = 0.05 dex, analogous to Bovy & Rix (2013).

4.2.4 Red Clump Stars as Standard Candles

Red clump stars. The red clump (RC) is a feature in the Hertzsprung-Russell diagram, located just blue-wards of the red giant branch (RGB). It corresponds to an overdensity of giant stars of almost constant color and luminosity. These stars are typically low-mass ($M \lesssim 1.8 M_\odot$; Carroll & Ostlie 2007, §13.2), relatively metal-rich and in the stable phase of helium core burning (CHeB). Older and more metal-poor stars settle on the bluer Horizontal Branch (HB), and more massive stars on the secondary RC during their CHeB phase. It is estimated that a third of all giant stars are RC stars (Girardi 2016).
After low-mass stars leave the main sequence, they ascend up the Red Giant Branch (RGB), their hydrogen-exhausted cores—not any longer supported by nuclear fusion and radiation pressure—collapse until electron-degeneracy pressure sets in, stabilizes the core and makes the equation of state independent of temperature. The hydrogen-burning shell dumps more helium on the core. This causes the core mass, density, and temperature to increase—the temperature even dramatically—, until a critical core mass of $M_{\mathrm{HeF}} \sim 0.47 M_\odot$ is reached and the triple-α process of helium burning starts explosively. This is called the helium core flash (HeF). Afterwards, the star settles in the stable phase of CHeB in the RC. Because the helium-core burning takes place at approximately the same core mass for all stars $M \lesssim 1.8 M_\odot$, the luminosity is approximately the same for all RC stars (Carroll & Ostlie 2007; Girardi 2016). More massive stars burn Helium under non-degenerate conditions and have therefore no sharply defined common luminosity in the Hertzsprung-Russell diagram—the secondary RC stars.
Their well-defined luminosity makes the low-mass RC stars ideal standard candles for distance estimation, as noted first by Cannon (1970).
Paczyński & Stanek (1998) were the first to actually estimate the absolute magnitude in the I-band for RC stars in the Hipparcos catalog (ESA 1997). Based on their Gaussian best fit $M_I^{\mathrm{RC}} = -0.279 \pm 0.088$, they measured the distance to the Galactic center to be $R_0 = 8.4 \pm 0.4$ kpc. Since then, RC distance indicators have proven to be useful in many studies, e.g., determination of the structure of the Galactic bar and bulge (Ness et al. 2013; Nataf et al. 2013; Wegg et al. 2015), the Galactic warp (Momany et al. 2006), the chemical evolution in the MW disk (Nidever et al. 2014; Bovy et al. 2016b), distances to the Magellanic Clouds (Girardi & Salaris 2001; Glatt et al. 2008), to only name a few. The accurate determination of the absolute RC magnitude in many passbands will continue to be an important research topic for years to come, due to the ever increasing accuracy of parallaxes in *Gaia*. While the mean magnitude (in the K band) of the RC used to be considered as not dependent on metallicity (Alves 2000; Pietrzyński et al. 2003), newer studies show a weak dependence on metallicity (van Helshoecht & Groenewegen 2007; Grocholski & Sarajedini 2002) and color (Salaris & Girardi 2002), which need to be taken into account when constructing models for the RC absolute magnitude for precise distance estimation.

Red clump selection and distance estimation algorithm according to Bovy et al. 2014. Bovy et al. (2014) constructed an algorithm to select RC stars from the APOGEE survey (Sloan Digital Sky Survey III's Apache Point Observatory Galactic Evolution Experiment) (Majewski et al. 2015), as part of the SDSS-III DR11/DR12 (Alam et al. 2015), based on $(J-K_s)$, [Fe/H], $\log g$, T_{eff}, and to assign distances to 5-10% precision. It is motivated by the agreement of the location of the RC in the $\log g$-vs.-T_{eff} plane of both the theoretical stellar isochrones from the Padova and Trieste Stellar Evolution Code (PARSEC) library (Bressan et al. 2012) and RC stars from the APOKASC catalog (APOGEE and the *Kepler* asteroseismic science consortium (KASC)) (Pinsonneault et al. 2014; Stello et al. 2013) using T_{eff} from APOGEE and asteroseismic $\log g$ from APOKASC.

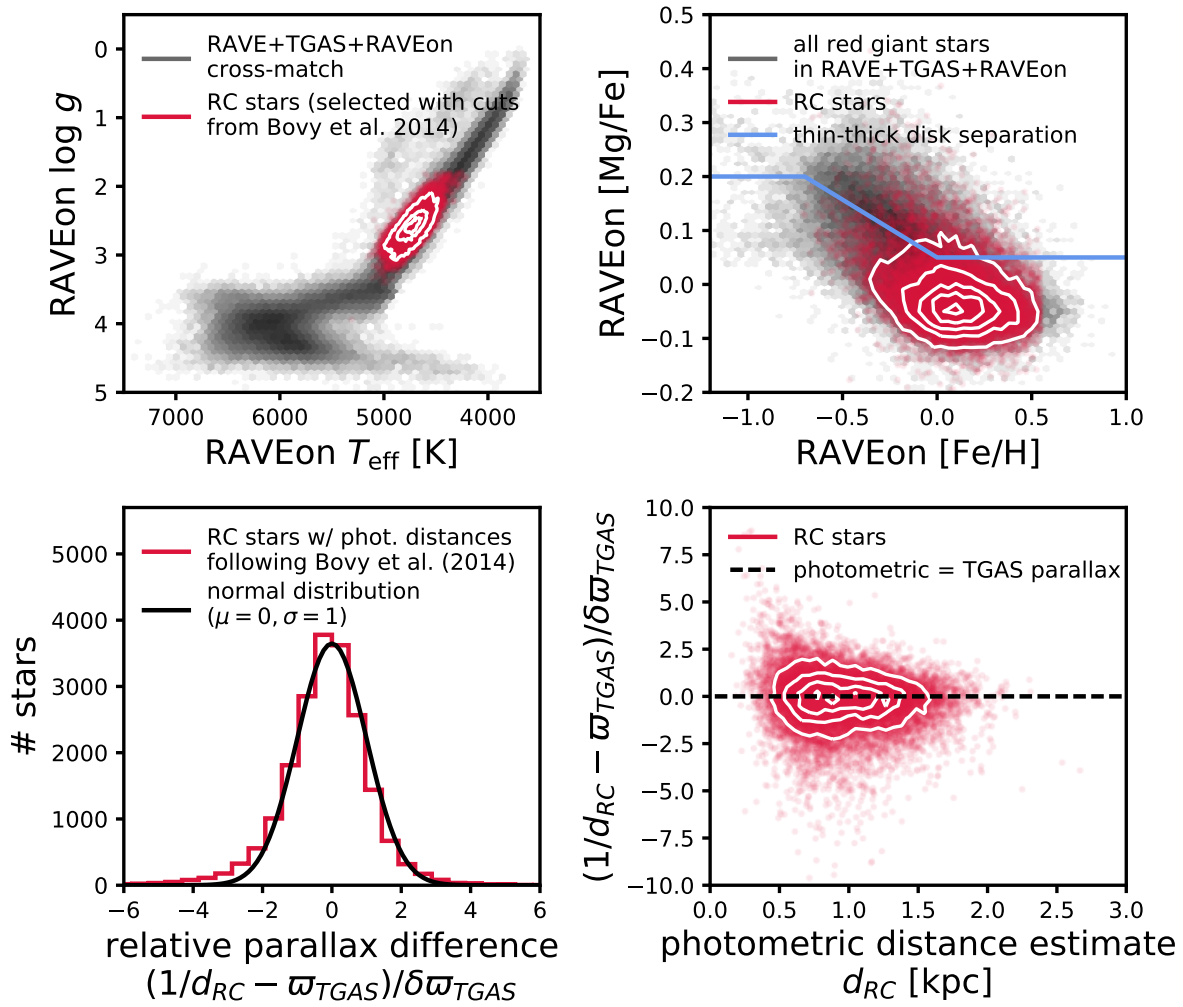


Figure 4.3: RC stars selected with the algorithm by Bovy et al. (2014) (summarized in Section 4.2.4) from the cross-match between the TGAS+RAVE+RAVE-on catalogs (provided by Georges Korodopatis, OCA, Nice). The upper left panel shows the spectroscopic Hertzsprung-Russell diagram based on the RAVE-on stellar labels, and the location of the selected RC stars. The upper right panel shows the RAVE-on [Fe/H] and [Mg/Fe] for the giant stars in RAVE-on only (because there are no [Mg/H] measurements for the main-sequence). We have overplotted the approximate classical separation between thin and thick disk based on their chemistry in blue (provided by Louise Howes (Lund Observatory), see also (Recio-Blanco et al. 2014; Kordopatis et al. 2015b)). The abundances of the RC stars suggest that our sample contains mostly metal-rich, α-young stars in the thin disk. The lower two panels compare the photometric distances $d_{\rm RC}$ estimated for the RC stars based on 2MASS photometry and the absolute magnitude model by Bovy et al. (2014) with the parallaxes measured by *Gaia* in the TGAS catalog. The histogram of the sample's relative parallax differences in the lower left panel is very close to a normal distribution, which means that, overall, the estimated photometric distances are indeed consistent and unbiased with respect to the TGAS parallaxes. The last panel shows that the sample is mostly located between $d_{\rm RC} \sim 0.4 - 2$ kpc and seems to suggest that this overall un-biased agreement does not depend on the distance of the star. There might be, however, a slight trend towards overestimation of the photometric distances.

To separate RC stars from RGB stars, Bovy et al. (2014) first applied a visual selection in $(\log g, T_{\rm eff})$ with a weak dependence on [Fe/H]; for completeness we repeat equations (2) and (3) from Bovy et al. (2014):

$$1.8 \leq \log g \leq 0.0018 \text{ dex K}^{-1} \left(T_{\rm eff} - T_{\rm eff}^{\rm ref}\left([\text{Fe/H}]\right)\right) + 2.5, \tag{4.1}$$

$$T_{\rm eff}^{\rm ref}\left([\text{Fe/H}]\right) = -382.5 \text{ K dex}[\text{Fe/H}] + 4607 \text{ K}. \tag{4.2}$$

To effectively cut most secondary RC stars, RGB and Asymptotic Giant Branch (AGB) stars, and to create a model for the absolute RC magnitude, Bovy et al. (2014) defined additional cuts in the Color-Magnitude Diagram (CMD)—absolute magnitude in the K_s band, M_{K_s}, versus de-reddened color, $(J-K_s)_0$, based on 2MASS NIR photometry (Skrutskie et al. 2006)—depending on metallicity Z. In particular, for each Z they constructed the stellar number density in the CMD—given assumptions for the IMF ($\rightarrow$ log-normal following Chabrier 2001) and star formation history ($\rightarrow$ constant)—and then applied cuts in $(J-K_s)_0$ which ensured that the spread in RC M_{K_s} for this stellar population is small and well-defined ($\sigma_{M_{K_s}} \lesssim 0.1 - 0.2$ mag), which is directly motivated by the physical origin of the primary RC with its constant luminosity. The cuts applied by Bovy et al. (2014) are:

$$\begin{aligned} Z &\geq 1.21\left((J-K_s)_0 - 0.05\right)^9 + 0.0011 && (4.3)\\ Z &\leq 2.58\left((J-K_s)_0 - 0.40\right)^3 + 0.0034 && (4.4)\\ 0.5 \leq (J-K_s)_0 &< 0.8 && (4.5)\\ Z &\leq 0.06, && (4.6) \end{aligned}$$

(their equations (6)-(8)). This provides a selection of RC stars with a purity of 93% and a model for $M_{K_s}\left[(J-K_s)_0, Z\right]$. The distances estimated from the distance modulus $\mu = K_{s,0} - M_{K_s}$ (with $K_{s,0}$ being the de-reddened K_s-band magnitude of the star) are expected to be biased only to within 2% and have random errors of 5% based on a calibration with Hipparcos RC stars. The algorithm is implemented in the `apogee` python package by Bovy (2016), which can be downloaded at `https://github.com/jobovy/apogee`.

RC stars and distances for RoadMapping. We use the cross-match based on *Tycho-2* IDs of the RAVE, TGAS and RAVE-on catalogs that was provided by Georges Kordopatis (OCA, Nice) at the *Gaia* Sprint Meeting at the Simons Foundation in New York City in October 2016. We first removed all duplicates from the data set, keeping for each star (identified via their RAVE ID) only the measurement with the highest signal-to-noise ratio. The cross-matched catalog contains 213,932 stars. We then selected all stars that satisfy the RAVE and RAVE-on quality conditions mentioned in Sections 4.2.2 and 4.2.3 and for which RAVE-on measurements for $\log g, T_{\rm eff}$, [Fe/H] and [Mg/H] exist. The latter condition removes the whole main-sequence, because [Mg/H] measurements exist only for giant stars in RAVE-on. We retain 81,560 giant stars.
The catalog by Georges Kordopatis also contains reddening estimates for all stars. He used the 2D dust maps by Schlegel et al. (1998), which provide the integrated reddening $E(B-V)$ along a given line-of-sight, to estimate the extinction of a star at the given 2D position without correction for distance. We show the reddening at the positions of the stars in the data set in the first panel in Figure 4.4. The extinction in the 2MASS J and K_s photometric bands was obtained using the following relations by McCall (2004):

$$\begin{aligned} A_J &= 0.819 \cdot E(B-V) && (4.7)\\ A_{K_s} &= 0.350 \cdot E(B-V). && (4.8) \end{aligned}$$

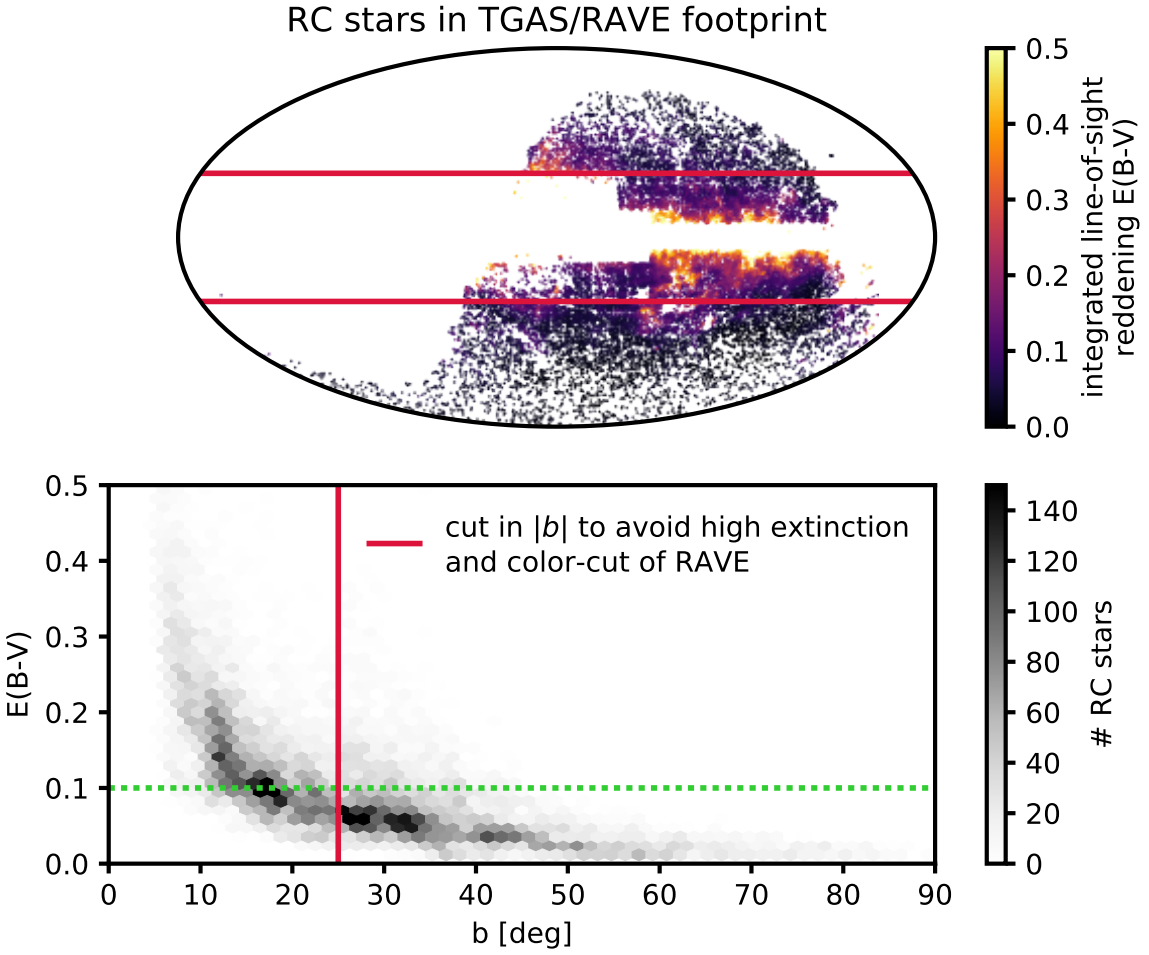


Figure 4.4: Reddening due to dust extinction at the positions of the RC stars in the TGAS/RAVE footprint. The upper panel shows the spatial distribution of RC stars, as selected following Section 4.2.4, and the 2D reddening estimates from Schlegel et al. (1998) as provided by Georges Kordopatis (OCA, Nice). The lower panel plots a 2D stellar number count histogram in reddening vs. Galactic latitude $|b|$. To achieve for most of the stars a reddening below $E(B-V) = 0.1$, the footprint should be restricted to $|b| > 25°$. In that way, the systematic errors in RC distances due to insufficient correction for reddening are small and the SF also does not have to incorporate a dust map.

We de-redden the 2MASS colors and magnitudes according to

$$(J-K_s)_0 = (J-K_s) - A_J + A_{K_s} \quad (4.9)$$

$$K_{s,0} = K_s - A_{K_s}. \quad (4.10)$$

Conveniently, the giant stars in RAVE-on have labels on the APOGEE scale, and we can directly apply the RC selection algorithm by Bovy et al. (2014) developed for APOGEE. We therefore use the de-reddened color and magnitude together with the RAVE-on $T_{\rm eff}$, $\log g$ and [Fe/H] (assuming these quantities had no uncertainties), select 19,836 likely RC stars from TGAS/RAVE/RAVE-on, and assign distances following the algorithm described in the previous paragraph. For the RoadMapping analysis, we assume for each RC star a distance uncertainty of 5%.

In Figure 4.3, we show our selection of RC stars in the Hertzsprung-Russell diagram and in the [Mg/Fe]-vs.-[Fe/H] plane of the RAVE-on stellar labels. The RC appears as a overdensity in the $\log g$-vs.-$T_{\rm eff}$ plane, but its location does strongly overlap with the RGB and there is no obvious

shift to higher temperatures that one usually expects. This might lead to some contamination of our sample by RGB stars. The comparison of the abundances of all RAVE-on giant stars with our RC sample shows that the algorithm by Bovy et al. (2014) mostly selected metal-rich stars in the thin disk. The lower panels in Figure 4.3 compare the estimated photometric distances, $d_{\rm RC}$, for the RC stars with their parallaxes from TGAS. Overall, they are fully consistent with each other and unbiased. There might be however a slight overestimation of the photometric distances. The stars are mostly located within $d_{\rm RC} \sim 0.4 - 2$ kpc from the Sun.

4.3 Selection Function

In dynamical modeling, taking into account the SF of the survey is crucial (Rix & Bovy 2013; see also Section 2.3.3). In this chapter, we attempt to model data from the *Gaia* DR1/RAVE cross-match; two surveys both having complex observation strategies (see Section 4.2). In the following, we present our strategy to set up a SF for RC stars in the TGAS/RAVE cross-match to use in the RoadMapping analysis.

4.3.1 Completeness of RAVE

The official RAVE SF. As described in Section 4.2.2, the overall survey SF of RAVE, with 9 mag $\leq I \leq$ 12 mag and a color cut below $|b| = 25$ deg, is relatively simple. It is also chemically and kinematically unbiased as shown by Wojno et al. (2017). The actual completeness of the survey is however a complicated function of position on the sky and the I-band magnitude. Wojno et al. (2017) chose a data-driven way to determine the completeness by calculating the ratio of stars in RAVE and the number of stars in 2MASS, which was used as complete parent sample. 2MASS (Skrutskie et al. 2006) is a NIR photometric survey, which provides magnitudes in the (J, H, K_s) photometric bands, covers the whole sky, and recorded point sources down to $J = 15.8$ mag and $K_s = 14.3$ mag. Above $|b| = 30$ deg, it has a completeness of 99%. Because it covers the full footprint of RAVE and is almost complete over RAVE's NIR magnitude range, it is a good choice for a parent sample. Wojno et al. (2017) used the 2MASS J and K_s bands to construct an approximate I-band magnitude:

$$I_{\rm 2MASS} = J + (J - K_s) + 0.2 \exp\left(\frac{(J - K_s) - 1.2}{0.2}\right) + 0.12. \tag{4.11}$$

They separated the sky into pixels of equal area, using the `HEALPIX` algorithm by Górski et al. (2005). An early version (November 2016) of their SF used 12,288 pixels (`HEALPIX` level 5, `NSIDE`=32, area$\simeq$ 3.36 deg), pixelated in the (l, b)-plane of the sky and using a `RING` ordering of the pixels. The stars in RAVE and 2MASS were then independently binned into these `HEALPIX` pixels and bins in $I_{\rm 2MASS}$ of 0.1 mag.

Our implementation of the RAVE SF. The first decision was to completely reject the Galactic bulge and disk region affected by the RAVE color-cut $|b| < 25$ dec from the RoadMapping modeling. In that way, the SF becomes considerably simpler, also because we avoid the highly-extincted regions ($E(B - V) > 0.1$) of the Galaxy, as demonstrated in Figure 4.4, and do not have to take a dust map into account.
The RAVE SF has a lot of spatial fine-structure (see Figure 4.2). In RoadMapping only the SF integrated over ϕ as function of (R, z) matters (cf. Equation (2.3)). Bovy et al. (2016a) showed, that SF variations on scales much smaller than the stellar density gradient do not matter for the density modeling. And in Section 2.3.3, we showed that there is a certain amount by which we can misjudge the SF before the RoadMapping results degrade. The very small bins in which Wojno et al. (2017) calculated the RAVE SF function (`HEALPIX NSIDE`=32, $\Delta I_{\rm 2MASS} = 0.1$ mag)

contains less than 30 RAVE stars per bin, which leads to a lot of statistical noise when calculating the completeness fraction $c_{\mathrm{RAVE}} = N_{\mathrm{RAVE}}/N_{\mathrm{2MASS}}$ per bin and pixel. All of this motivated us to use larger pixels and magnitude bins than Wojno et al. (2017). We only used the 2MASS counts from an early version of Wojno et al. (2017), re-binned them, and and also binned the RAVE DR5 stars (Kunder et al. 2017)[22] (after removing duplicates) into pixels in (l, b) with `HEALPIX NSIDE` $= 16$, and magnitude bins $\Delta I_{\mathrm{2mass}} = 0.5$ mag in the range 9 mag $< I_{\mathrm{2MASS}} \leq 12$ mag. The completeness fraction is

$$c_{\mathrm{RAVE}}(\texttt{HEALPIX\_ID}, I_{\mathrm{2MASS}}) = \frac{N_{\mathrm{RAVE}}(\texttt{HEALPIX\_ID}, I_{\mathrm{2MASS}})}{N_{\mathrm{2MASS}}(\texttt{HEALPIX\_ID}, I_{\mathrm{2MASS}})}. \tag{4.12}$$

We did not consider any quality flags before binning the RAVE stars, because the SF uses only positions and 2MASS photometry, and no spectroscopy at all. This approach is acceptable as long as the assumption holds that if a spectrum is good or bad does not depend on the 3D position of the star. We assume this to be indeed the case, given the fact that in RAVE the exposure time was adapted for pointings with fainter stars to achieve a higher signal-to-noise. The first row of panels in Figure 4.7 illustrates the RAVE SF we are using.

4.3.2 Completeness of TGAS

Idea and Motivation. As of September 2016 to February 2017, when we prepared and ran the RoadMapping analysis for TGAS data, there was no published version of the TGAS SF available yet. So we decided to set up a SF analogous to the RAVE SF, and to use the full *Gaia* DR1 (secondary data set) as a parent sample for TGAS.
To qualify as a parent sample, we need to gauge if *Gaia* is complete over a given magnitude range. We use the histogram in G-magnitude in Figure 4.5(a) together with a simple argument to do that. We assume that the stars were uniformly distributed and spatially well mixed in magnitude, which might be, at least for the direct solar neighborhood, an assumption that is not too far from the truth. Then the number of stars at a given distance would be $N(d) \propto d^2$, from which follows $N(G \mid M_G) \propto \left(10^{0.2(G-M_G)+1}\right)^2$. Then the number of stars per G-bin would be $N(G) = \int N(G \mid M_G) \cdot N(M_G)\, \mathrm{d}M_G \propto 10^{0.4G}$, where $N(M_G)$ is the total number of stars with absolute magnitude M_G, spatially uniformly distributed. Therefore

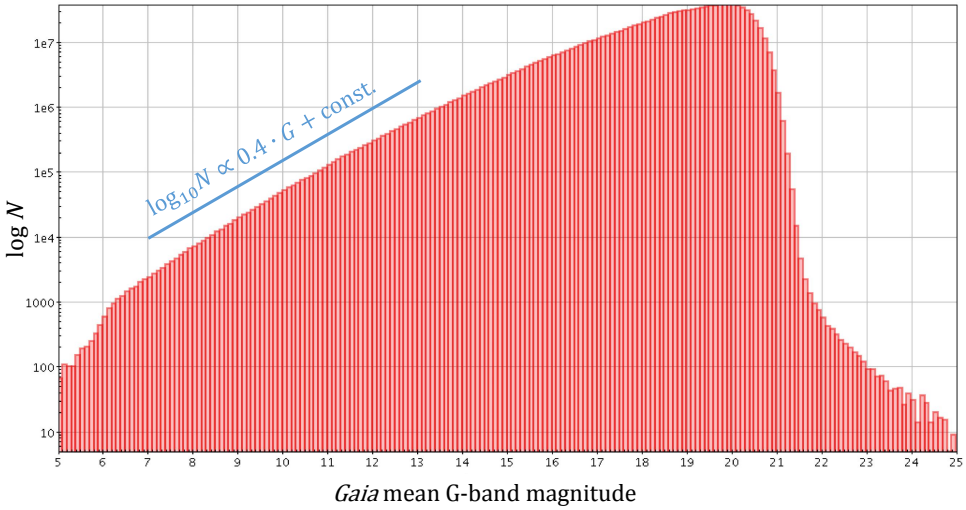
$$\log_{10} N(G) = 0.4 \cdot G + \text{const.}, \tag{4.13}$$

if the data set was complete. Indeed, in the region 7 mag $\leq G \lesssim 13$ mag the logarithmic histogram for the secondary data set in Figure 4.5(a) follows approximately a straight line with slope ~ 0.4, while for brighter and fainter stars the number counts drop. From this simple argument we deduce that in this magnitude range the secondary data set in *Gaia* DR1 is expected to be approximately complete and could be used as a parent sample for the SF of TGAS. The primary data set of *Gaia* DR1, TGAS in Figure 4.5(b), is also mostly located between 7 mag $< G <$ 13 mag. We use

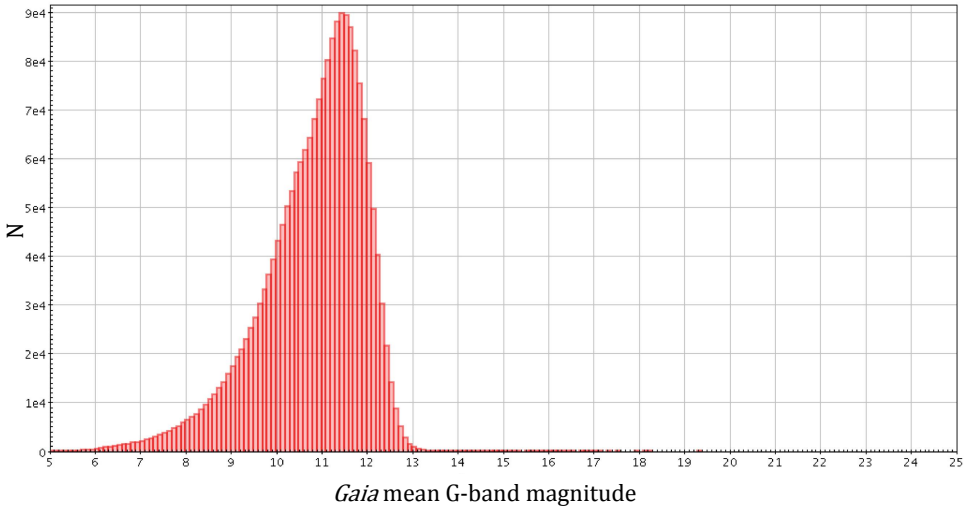
$$c_{\mathrm{TGAS}}(l, b, G) = \frac{N_{\mathrm{TGAS}}(l, b, G)}{N_{\mathrm{Gaia}}(l, b, G)} \tag{4.14}$$

as an estimate for the completeness of TGAS, with $N_{\mathrm{TGAS}}(l, b, G)$ being the number of stars per `HEALPIX` pixel at (l, b) and per G-magnitude bin in TGAS, and $N_{\mathrm{Gaia}}(l, b, G)$ the number of stars in the secondary data set of *Gaia* DR1.
The assumption that the secondary data set of *Gaia* DR1 is complete in 7 mag $\leq G \lesssim 13$ mag is

[22] The RAVE data is online available on the RAVE webpage, `https://www.rave-survey.org/downloads/`.



(a) Distribution of G-magnitudes in *Gaia* DR1/secondary catalog.



(b) Distribution of G-magnitudes in *Gaia* DR1/primary catalog/TGAS

Figure 4.5: Number of stars per *Gaia* G-band magnitude in the secondary data set in *Gaia* DR1 (upper panel) and TGAS (lower panel). In the upper panel we overplotted the simple completeness estimate in Equation (4.13) in blue, demonstrating that *Gaia* is approximately compelte in the range $G \in [7, 13]$ mag. **Figure credit:** The *Gaia* Archive (ESA 2017c).

quite simplified, as can be seen in Figure 4.6(a). The number counts of *Gaia* stars still show the imprints of the *Gaia* scanning law (Gaia Collaboration et al. 2016b, §5.2), which comes from the fact that stars that had fewer than five focal plane transits in the 14 months of operation are not included in *Gaia* DR1 (Arenou et al. 2017). This is a clear caveat. Also, the limiting magnitude varies with position. Given the following arguments and assumptions, we can still use *Gaia* as parent sample for 7 mag $< G <$ 13 mag:

- The limiting magnitude of *Gaia* DR1 is $G > 13$ mag everywhere (Arenou et al. 2017, their Figure 4b).
- The completeness due to the scanning law of *Gaia* is worst at high latitudes in the North-West and South-East of the MW. The RAVE footprint excludes the South-East region of the MW anyway, and also the North-West high-latitude region is only partly covered by RAVE (cf. Figure 4.2). The scanning-law might therefore not bias the SF too much.

- Most of the *Tycho-2* and therefore TGAS stars have $V \sim 11$ mag (the $\sim 99\%$ completeness limit of *Tycho-2* Høg et al. 2000), which corresponds to a distance of $d \sim 1.1$ kpc for RC stars ($M_I \sim -0.22, (V-I) \sim 1$; Girardi 2016). We will see later in Section 4.4.1 that most of the stars in our modeling are inside this distance from the Sun anyway. Some deviations in the low-completeness outer-regions of the SF should not matter too much for the modeling, as we found in Section 2.3.3.

Alternatives for a parent sample would have been:

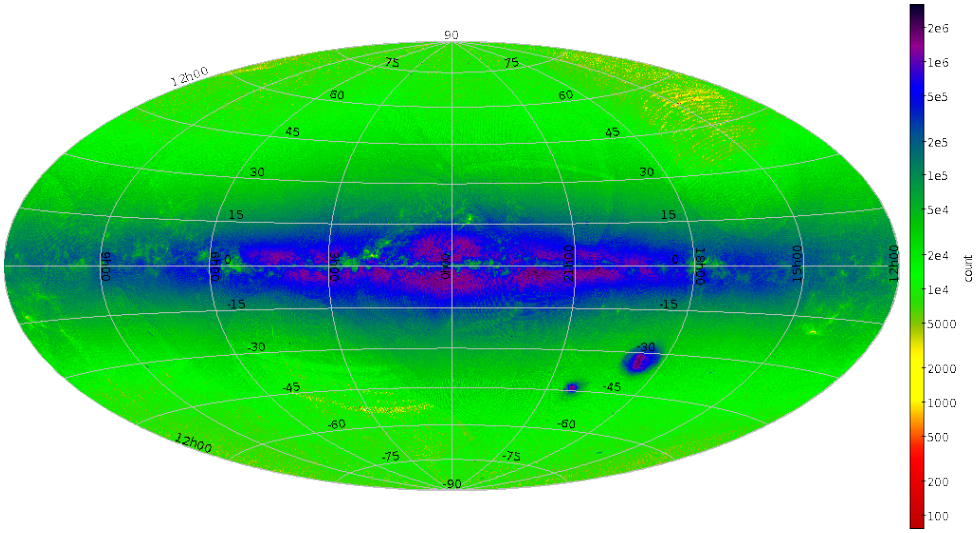
- *Tycho-2*. Advantage: Well-understood completeness at $V \lesssim 11$ mag. Disadvantage: The SF of TGAS would be more brightness-limited than even RAVE.

- 2MASS. Advantage: Complete down to faint magnitudes. Disadvantage: Completeness defined in terms of NIR magnitudes. *Tycho-2* and *Gaia*, however, strive to be complete in visible/white-light magnitudes. A SF using 2MASS would therefore have to be defined as a function of color in addition to magnitude and position (cf. Bovy 2017).

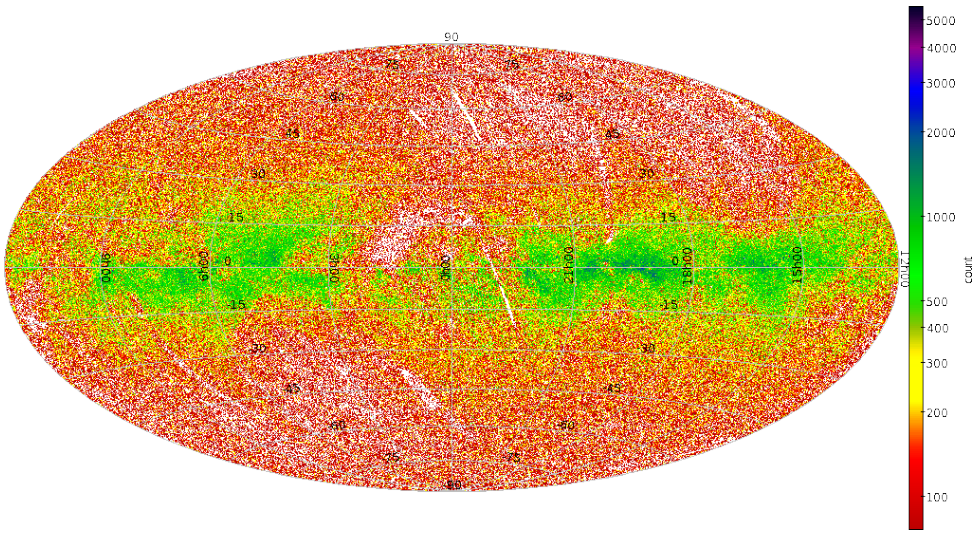
Realisation. The goal is to bin all TGAS and *Gaia* DR1 stars into bins of $\Delta G = 0.5$ mag and `HEALPIX` pixels of `NSIDE` = 16 / level = 4 in (l, b) (ring-like pixelation scheme), analogous to the RAVE SF in Section 4.3.1. The *Gaia* `source_id` can be directly used to generate the `HEALPIX_ID`. However, *Gaia* uses `HEALPIX` pixels in (R.A.,Dec.) and a nested pixelation scheme. We proceeded by using the *Gaia* Archive (ESA 2017c) to bin the whole catalog into `HEALPIX` pixels of level = 10 and $\Delta G = 0.1$ mag with the following Astronomical Data Query Language (ADQL) command:

```
SELECT
     (FLOOR(phot_g_mean_mag*10)/10+0.1) AS mag_bin,
     (source_id / 549755813888) AS healpx10,
     COUNT(*) AS num FROM gaiadr1.Gaia_source
WHERE
     (b > 25 OR b < -25)
     AND (dec < 5)
     AND (phot_g_mean_mag >= 7)
     AND (phot_g_mean_mag <= 14)
GROUP BY
     healpx10, mag_bin
ORDER BY
     healpx10, mag_bin
```

and analogous for TGAS, where `gaiadr1.Gaia_source` was replaced by `gaiadr1.tgas_source`. `(FLOOR(phot_g_mean_mag*10)/10+0.1)` returns the magnitude bin G_i for a star with $G_i \leq G < G_i + 0.1$. `source_id / 549755813888` returns the `HEALPIX_ID` of level 10 (Gaia@AIP 2016). To reduce the size of the query, we excluded regions $|b| < 25$ deg (plane of the disk) and Dec. > 5 deg (northern hemisphere), that are not part of our RAVE SF or the RAVE survey, respectively.
The second step is then to sum the counts of all level 10 / nest-pixelation `HEALPIX` pixels in (R.A.,Dec.) (as provided by the *Gaia* Archive) that fall into the same level 4 / ring-pixelation pixel in (l, b), to achieve the same pixelation scheme as our RAVE SF. We further bin the counts into bins of $\Delta G = 0.5$ mag.



(a) Density of stars in *Gaia* DR1/secondary catalog.



(b) Density of stars in *Gaia* DR1/primary catalog/TGAS.

Figure 4.6: 2D stellar number counts of stars in *Gaia* DR1 on the plane of the sky in (l, b) coordinates. The upper panel shows the number density of all 1.1 billion stars of the secondary data set in *Gaia* DR1, and the lower panel all 2 million stars in TGAS, the primary data set of *Gaia* DR1. **Figure credit:** The *Gaia* Archive (ESA 2017c).

Proceeding in the same way for the TGAS and the full *Gaia* DR1 catalog, we arrive at the completeness fraction

$$c_{\mathrm{TGAS}}(\texttt{HEALPIX\_ID}, G) = \frac{N_{\mathrm{TGAS}}(\texttt{HEALPIX\_ID}, G)}{N_{\mathrm{Gaia}}(\texttt{HEALPIX\_ID}, G)}, \tag{4.15}$$

which is illustrated in the second row of Figure 4.7.

4.3.3 Selection Function and Likelihood Normalisation for Red Clump Stars in RoadMapping

The full SF for a MAP of RC giants. The SF describes the probability that a star, given it exists, ended up in the final data set. In creating a MAP of RC stars, we applied several

selections. If a star is a RC star depends on $\log g, T_{\rm eff}, \text{color}, \text{metallicity}$ as described in Section 4.2.4, i.e.,

$$c_{\rm RC}(\log g, T_{\rm eff}, \text{color}, [\text{Fe/H}]) = \begin{cases} 1 & \text{if RC star according to Bovy et al. (2014)} \\ 0 & \text{otherwise} \end{cases}. \tag{4.16}$$

If the star ended up in a given MAP depends on the chemical abundances[23]

$$\begin{aligned} & c_{\rm MAP}([\text{Fe/H}], [\text{Mg/Fe}] \mid [\text{Fe/H}]_0, [\text{Mg/Fe}]_0) && (4.17)\\ & \approx \delta([\text{Fe/H}] - [\text{Fe/H}]_0) \times \delta([\text{Mg/Fe}] - [\text{Mg/Fe}]_0). && (4.18) \end{aligned}$$

The probability of a star to end up in RAVE depends on its $I_{\rm 2MASS}$ magnitude and sky position according to Equation (4.12) (which incorporates also the RAVE footprint on the southern hemisphere). The probability that the star is part of TGAS is given by Equation (4.15) and depends on the G magnitude in addition to the sky position. Here we assume that the selection whether a star is in TGAS or in RAVE is completely independent of each other. This is not fully true, as the input catalog for RAVE DR1-3 was partly based on *Tycho-2* (see Section 4.2.2). As DR3 only contains $\sim 15\%$ of the stars in DR5, we might still consider RAVE and TGAS as independent.
Overall, the SF for one of the RC MAPs we are using is

$$\begin{aligned} & \text{SF}_{\rm data}(l, b, I_{\rm 2MASS}, G, \log g, T_{\rm eff}, \text{color}, [\text{Fe/H}], [\text{Mg/Fe}] \mid [\text{Fe/H}]_0, [\text{Mg/Fe}]_0) && (4.19)\\ \propto\ & c_{\rm RC}(\log g, T_{\rm eff}, \text{color}, [\text{Fe/H}]) && (4.20)\\ & \times c_{\rm MAP}([\text{Fe/H}], [\text{Mg/Fe}] \mid [\text{Fe/H}]_0, [\text{Mg/Fe}]_0) && (4.21)\\ & \times c_{\rm RAVE}(l, b, I_{\rm 2MASS}) && (4.22)\\ & \times c_{\rm TGAS}(l, b, G). && (4.23) \end{aligned}$$

Modeling assumptions for RC MAPs in RoadMapping. In RoadMapping we only need the SF for the normalization of the likelihood in Equation (2.3), where the product of the DF and the data SF is integrated over the whole space of observables.
The RoadMapping modeling assumes that the distribution of all stars in the Galactic disk is fully described by a DF

$$\text{DF}_{\rm disk}(\boldsymbol{x}, \boldsymbol{v}, [\text{Fe/H}], [\text{Mg/Fe}] \mid p_M) \tag{4.24}$$

with model parameters p_M. If we consider a single MAP, then

$$\begin{aligned} & \int \text{d}[\text{Fe/H}]\, \text{d}[\text{Mg/Fe}] && (4.25)\\ & \qquad \text{DF}_{\rm disk}(\boldsymbol{x}, \boldsymbol{v}, [\text{Fe/H}], [\text{Mg/Fe}] \mid p_M) && (4.26)\\ & \qquad \times c_{\rm MAP}([\text{Fe/H}], [\text{Mg/Fe}] \mid [\text{Fe/H}]_0, [\text{Mg/Fe}]_0) && (4.27)\\ \propto\ & \text{DF}_{\rm disk}(\boldsymbol{x}, \boldsymbol{v}, [\text{Fe/H}]_0, [\text{Mg/Fe}]_0 \mid p_M) && (4.28)\\ =\ & \text{qDF}(\boldsymbol{x}, \boldsymbol{v} \mid p_M), && (4.29) \end{aligned}$$

where the latter step incorporates the basic assumption of RoadMapping, that a MAP is described by the qDF.
We require an additional assumption, that assumes that RC stars in a given MAP are distributed in phase-space following the same physical DF as all stars in this MAP. This assumption might

[23] In practice, we define a MAP as bin with finite width in the chemical abundance space rather than a delta-function, but as we, in the long-term, strive to use very small abundance bins, we use the delta-function already here to simplify the argumentation.

not be too far off from the reality, when one considers a MAP as being a proxy for a mono-age population and assumes that its IMF did not vary spatially. (The IMF is however allowed to vary with chemical abundance, and the star-formation efficiency can vary spatially.) Then all stars including the RC giants were subject to the same radial migration and heating processes, such that their current-time DF should not differ between giants and dwarfs.

The RC MAP SF in the likelihood normalisation. The above assumptions reduce the likelihood normalisation considerably to

$$\begin{aligned}
M_{\rm tot}(p_M) &\propto \int \mathrm{d}\boldsymbol{x}\ \mathrm{d}\boldsymbol{v}\ \mathrm{d}\log g\ \mathrm{d}T_{\rm eff}\ \mathrm{dcolor}\ \mathrm{d[Fe/H]}\ \mathrm{d[Mg/Fe]} \\
&\qquad \mathrm{DF}_{\rm disk}(\boldsymbol{x},\boldsymbol{v},\mathrm{[Fe/H]},\mathrm{[Mg/Fe]}) \times \mathrm{SF}_{\rm data}([...]) \\
&\propto \int \mathrm{d}\boldsymbol{x}\ \mathrm{d}\boldsymbol{v}\ \mathrm{d}\log g\ \mathrm{d}T_{\rm eff}\,\mathrm{dcolor} \\
&\qquad \mathrm{qDF}(\boldsymbol{x},\boldsymbol{v}) \times c_{\rm RC}([...,\mathrm{[Fe/H]}_0]) \times c_{\rm RAVE}(l,b,I_{\rm 2M}) \times c_{\rm TGAS}(l,b,G) \\
&\propto\!\!\!\!\sim \int \mathrm{d}\boldsymbol{x}\ \mathrm{d}\boldsymbol{v} \\
&\qquad \mathrm{qDF}(\boldsymbol{x},\boldsymbol{v}) \times c_{\rm RAVE}(l,b,I_{\rm 2M}=\mu+M_{I_{\rm 2M},\rm RC}) \times c_{\rm TGAS}(l,b,G=\mu+M_{G,\rm RC}) \\
&\equiv \int \mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{v}\ \mathrm{qDF}(\boldsymbol{x},\boldsymbol{v}) \times \mathrm{SF}_{\rm TGAS/RAVE}(l,b,\mu) \qquad (4.30)
\end{aligned}$$

where we assumed in the second-to-last step, that RC stars are standard candles with approximately constant luminosity in both $I_{\rm 2MASS}$ and G band, and have therefore also the same color $(I_{\rm 2MASS}-G)_{\rm RC}$, and $\mu = I_{\rm 2MASS} - M_{I_{\rm 2M},\rm RC} = G - M_{G,\rm RC}$ is the distance modulus. Model-independent constant pre-factors of the likelihood normalisation do not matter for the RoadMapping modeling and are ignored here. The SF is now reduced to be a function of position $\boldsymbol{x}$ only. We illustrate $\mathrm{SF}_{\rm TGAS/RAVE}(l,b,d) = c_{\rm TGAS}(l,b,d=10^{\mu/5+1}) \times c_{\rm RAVE}(l,b,d=10^{\mu/5+1})$ in the third row of Figure 4.7.

We further simplify

$$\begin{aligned}
M_{\rm tot} &\propto \int \mathrm{d}\boldsymbol{x} \int \mathrm{d}\boldsymbol{v}\ \mathrm{qDF}(\boldsymbol{x},\boldsymbol{v}\mid p_M) \cdot \mathrm{SF}_{\rm TGAS/RAVE}(\boldsymbol{x}) && (4.31)\\
&= \int \mathrm{d}\boldsymbol{x}\ \rho_{\rm qDF}(\boldsymbol{x}\mid p_M) \cdot \mathrm{SF}_{\rm TGAS/RAVE}(\boldsymbol{x}) && (4.32)\\
&= \int \mathrm{d}R\,\mathrm{d}z\ \rho_{\rm qDF}(R,z\mid p_M) \cdot R \cdot \int \mathrm{d}\phi\ \mathrm{SF}_{\rm TGAS/RAVE}(\boldsymbol{x}) && (4.33)\\
&= \int \mathrm{d}R\,\mathrm{d}z\ \rho_{\rm qDF}(R,z\mid p_M) \cdot \mathscr{S}(R,z) && (4.34)
\end{aligned}$$

for an axisymmetric DF generating the tracer density $\rho_{\rm DM}$ that is a function of R and z only. We define

$$\mathscr{S}(R,z) \equiv R \cdot \int \mathrm{d}\phi\ \mathrm{SF}(\boldsymbol{x}) \qquad (4.35)$$

to be the *effective azimuthal length of the SF* at (R,z). Even if the SF is a complicated function of $\boldsymbol{x}$, we can at least pre-calculate the effective azimuthal length of the SF and reduce the normalisation to be only a two-dimensional integral.

Implementation details in RoadMapping. As absolute magnitude for the RC stars in the I-band we use

$$M_{I,\rm RC} \approx -0.22 \text{ mag}. \qquad (4.36)$$

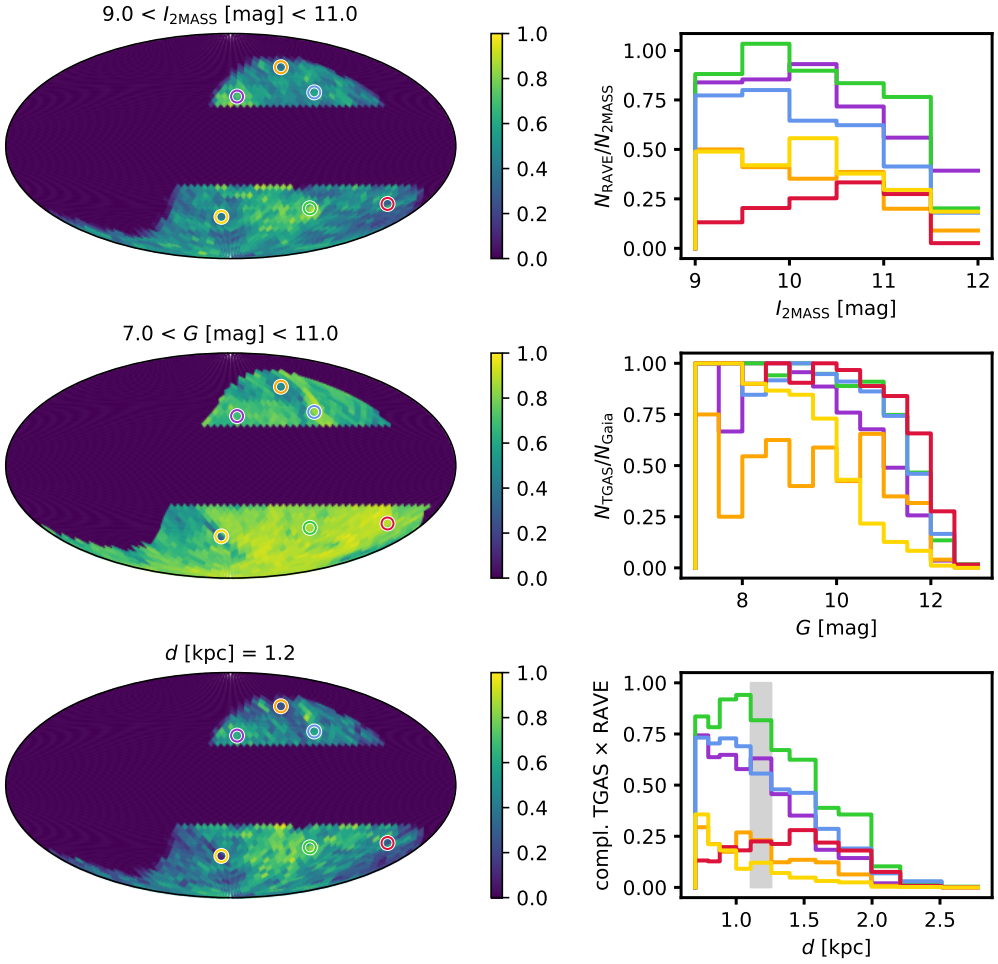


Figure 4.7: Completeness of the RAVE and TGAS catalogs, as estimated in `HEALPIX` pixels and magnitude bins by comparing to the number of stars in a parent sample, which is assumed to be complete in the considered magnitude range. In particular, we used 2MASS as parent sample for RAVE, and the full *Gaia* DR1 for TGAS. The first row of panels shows the RAVE SF we are using for the RoadMapping modeling. It is a combination of the RAVE footprint (Wojno et al. 2017; §2.3), our cut at $|b| = 25$ deg to avoid highly extincted regions in the Galactic plane (see Figure 4.4) and the RAVE color-cut in the disk, and the completeness fraction calculated following Equation (4.12). We restrict the SF to 9 mag $< I_{2MASS} \leq 12$ mag, the main magnitude range of the RAVE survey. The completeness function is calculated in `HEALPIX`els with `NSIDE` $= 16$ and in bins $\Delta I_{2MASS} = 0.5$. The second row of panels illustrates the TGAS completeness fraction following Equation (4.15), in the magnitude range 7 mag $\leq G \leq 13$ mag with bin size $\Delta G = 0.5$ mag. We show only the region of the RAVE footprint that is relevant for the modeling, even though the TGAS completeness could also be estimated outside of it. There are some regions of lower completeness north and south of the central bulge region, which directly comes from the *Gaia* scanning law. In these regions, stars were observed less often and do therefore not allow a reliable estimation of the astrometric parameters (see also the strip-like underdensities in Figure 4.6(b)). The lower row of panels assumes that the completeness fraction is the same for all stars as well as RC stars only, and uses the known absolute magnitude of RC stars to transform completeness in magnitude bins into completeness in distance bins (see Section 4.3.3). This allows to multiply the RAVE and TGAS SFs directly to achieve a joint completeness for the cross-match of these two catalogs. It is clear that our choice of spatial and magnitude bins to calculate the completeness are a good compromise between conserving some fine-structure, while also having a mostly smooth SF, which should be easier to integrate over the Galactocentric ϕ-coordinate.

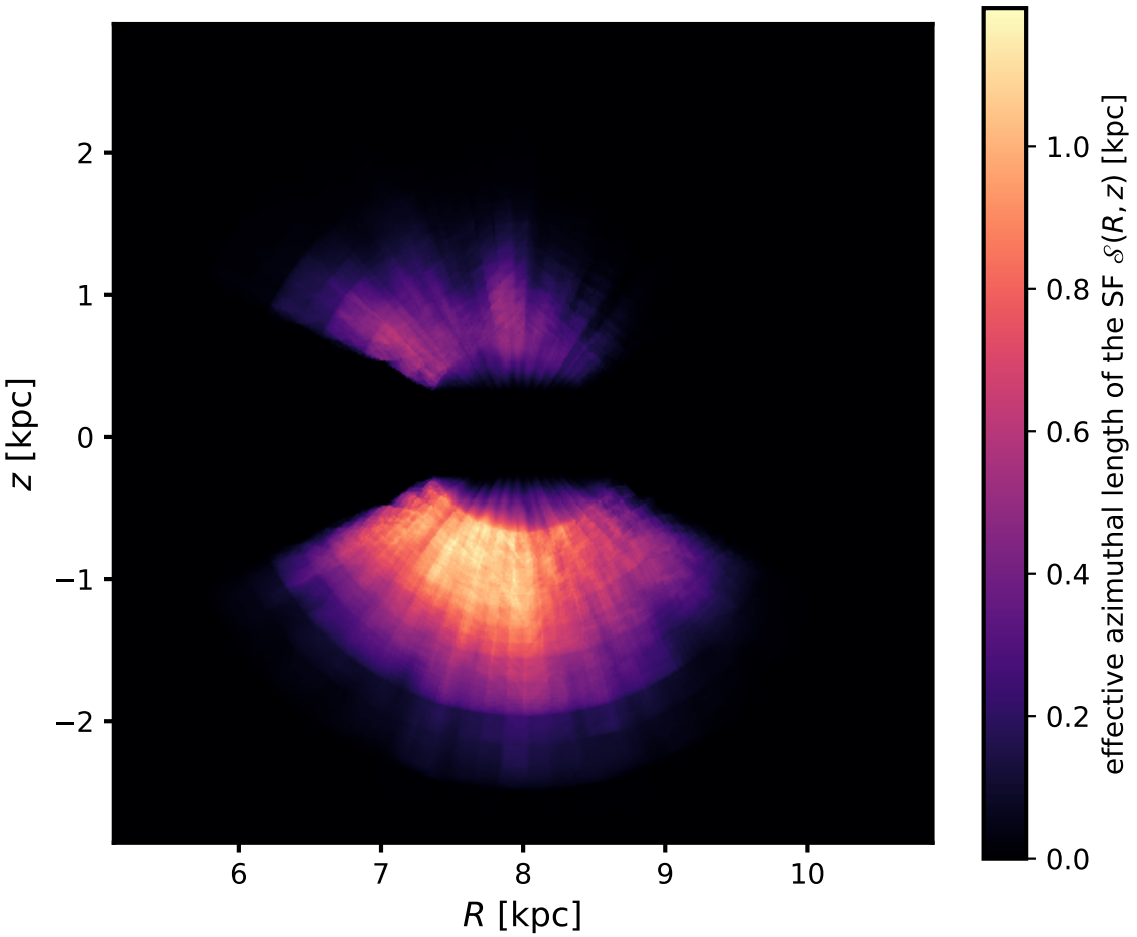


Figure 4.8: Effective azimuthal length $\mathscr{S}(R, z)$ of the TGAS/RAVE SF for RC stars, which is illustrated in the lower panels of Figure 4.7. $\mathscr{S}$ is calculated by integrating the SF(l, b, d) at a given Galacocentric (R, z) over ϕ following Equation (4.35). This is the pre-calculated SF that is used in our RoadMapping analyses. It illustrates that within the MW, the RAVE footprint restricts the survey volume to lay mostly below the plane, that we exluded the Galactic disk, and that the completeness in Figure 4.7 is a function of logarithmic distance bins, has a fixed minimum radius from the Sun corresponding to $I_{\rm 2mass} = 9$ mag, and decreases further out.

estimated by Groenewegen (2008) from Hipparcos parallaxes. For the *Gaia* G-band there was no official estimate for the absolute RC magnitude published by the end of 2016, so we estimated

$$\begin{aligned} (V-I)_{\rm RC} &= (M_{V,\rm RC} - M_{I,\rm RC}) \sim 1 \text{ mag} \\ \Rightarrow M_{V,\rm RC} &\sim 0.78 \text{ mag} \\ \Rightarrow M_{G,\rm RC} &\sim 0.5 \text{ mag} \end{aligned}$$

where we used the transformation by Jordi et al. (2010) between Johnson V and Johnson-Cousins $(V-I)$ and the *Gaia* G-band magnitude

$$G = V - 0.0257 - 0.0924 \cdot (V-I) - 0.1623 \cdot (V-I)^2 + 0.0090 \cot(V-I)^3. \tag{4.37}$$

In RoadMapping, we implemented the SF as a geometrical spherical shell around a given solar position (see Section 4.5.4) with a minimum and maximum radius. As RAVE with its $9 \leq I \leq 12$

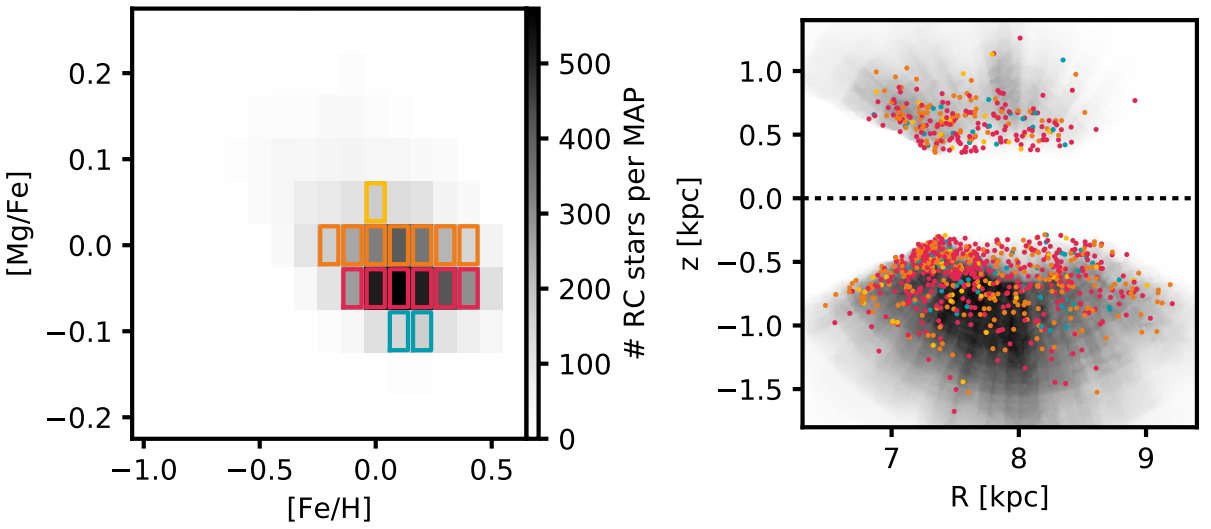


Figure 4.9: Distribution of RC stars within the TGAS/RAVE SF in the [Mg/Fe]-vs.-[Fe/H]-plane and Galactocentric (R, z). In the left panel we mark the 16 MAPs that we are going to analyse with RoadMapping (color-coded by [Mg/Fe]). They have between 151 and 573 stars. The right panel shows the spatial distribution of the stars, again color-coded by [Mg/Fe]. We show in grey-scale the effective azimuthal length of the SF from Figure 4.8.

magnitude range is spatially more restricted than the TGAS $7 \leq G \leq 13$, we use as outer radius and inner radius

$$d_{\max} = 10^{(12 \text{ mag} - M_{I,RC})/5+1} \text{ pc} \sim 2.8 \text{ kpc} \quad (4.38)$$

$$d_{\min} = 10^{(9 \text{ mag} - M_{I,RC})/5+1} \text{ pc} \sim 0.7 \text{ kpc}. \quad (4.39)$$

We also remove the stars outside of this shell from the data set, using the general assumption of RoadMapping that the 3D positions of stars are always perfectly known (even if there are distance uncertainties).
Each position within this shell has a completeness value given by the $\text{SF}_{\text{TGAS/RAVE}}(l, b, d)$. We setup a 400x400 pixel grid in (R, z) with side length $2 \times d_{\max}$, and integrate $\text{SF}_{\text{TGAS/RAVE}}$ over the shell along the Galactocentric ϕ direction to get the effective azimuth $\mathscr{S}(R, z)$ of the SF in Equation (4.35). Figure 4.8 shows this pre-calculated SF that is used in all the RoadMapping analyses.

4.4 First Glance at the Data

In Section 4.2, we set up a data set of RC stars with high-precision 6D phase-space coordinates; in particular (R.A.,decl.,$\mu_{\text{R.A.}*}$,$\mu_{\text{decl.}}$) from TGAS, v_{los} from RAVE, d_{RC} following Bovy et al. (2014). In addition, we will use [Fe/H] and [Mg/Fe] from RAVE-on. We restrict the data to the spatial region covered by the SF described in Section 4.3. In the following, we will have a first look at the physical characteristics of the data.

4.4.1 Size and Location of the Data

In the RoadMapping analysis, we want to investigate stellar MAPs in [Fe/H] and [Mg/Fe] independently. The left panel in Figure 4.9 shows the number of RC stars that fall into the same abundance bin with size Δ[Fe/H] = 0.1 dex and Δ[Mg/Fe] = 0.05 dex. There are in total 16 MAPs with between 151 and 573 RC stars. Bovy & Rix (2013) used MAPs with 100 to $\sim$ 800

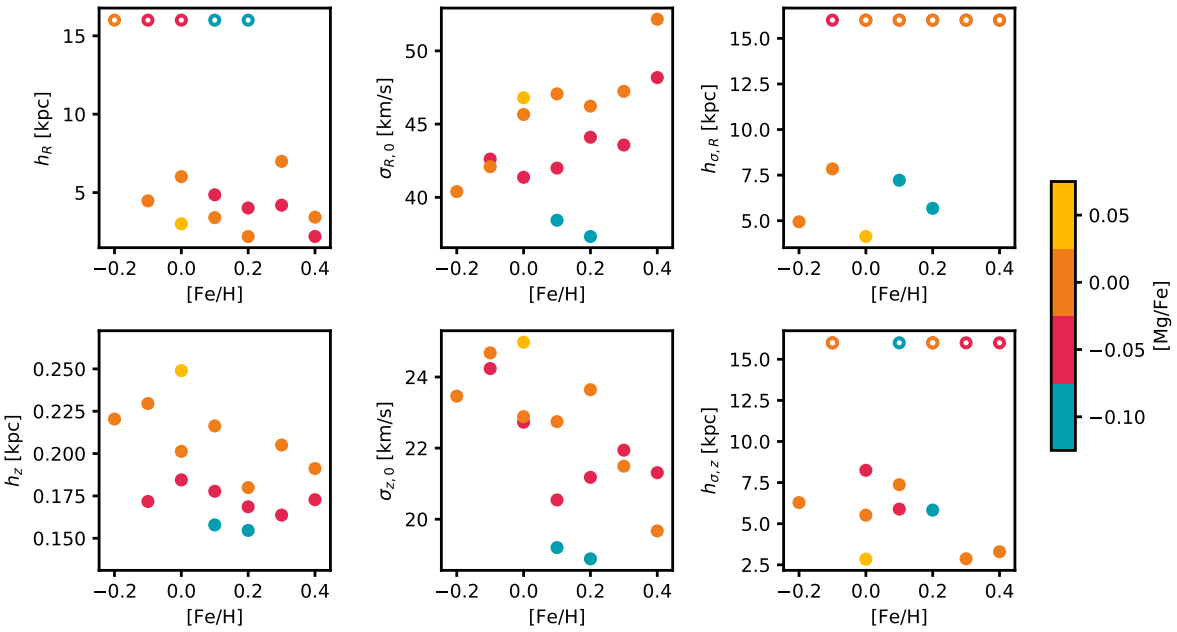


Figure 4.10: Non-dynamical fits of exponential tracer density and velocity dispersion profiles to the RC stars of the different MAPs shown in Figure 4.9. The model parameters are the radial tracer scale length and vertical scale height h_R and h_z, the radial and vertical velocity dispersion at the Sun $\sigma_{R,0}$ and $\sigma_{z,0}$, and the radial velocity dispersion scale lengths $h_{\sigma,R}$ and $h_{\sigma,z}$, defined in the model and likelihoods in Equations (4.42)-(4.44). If the scale lengths could not be fitted and/or were pegged at the upper fit limit of 16 kpc, we marked the corresponding MAP with an open circle. The parameters shown in this figure are used as initial values for the RoadMapping fit.

G-dwarf stars. Because we plan to fit more model parameters than Bovy & Rix (2013), we use only MAPs that have more than 150 stars. We mark the MAPs in the left panel of Figure 4.9 that we are going to analyse with RoadMapping.

The right panel of Figure 4.9 shows the spatial distribution in Galactocentric (R, z) coordinates. The coordinate transformation from heliocentric observables to Galactocentric cylindrical coordinates was performed using functions from the `galpy.bovy_coords` `python` module (Bovy 2015) and the phase-space coordinates of the Sun given in Section 4.5.4. We overplot the SF from Figure 4.8. While the SF has the largest $\mathscr{S}$ and therefore the largest effective survey volume at $R \sim 7.5$ kpc and $z \sim -1$ kpc, most stars are located close to the edge of the SF at $z \sim -0.5$ kpc and small radii, consistent with a disk tracer density distribution that decreases with $|z|$. The spatial coverage of our data sample is relatively small; we're only probing $R \in [7, 9]$ kpc and $|z| \in [0.3, 1]$ kpc. The spatial coverage of the SEGUE G-dwarfs in Bovy et al. (2012d); Bovy & Rix (2013) with $R \in [6, 12]$ kpc and $|z| \in [0.4, 4]$ kpc was, in comparison, much larger.

4.4.2 Spatial and Kinematic Structure of the Data

Before applying the RoadMapping analysis to the MAP data, we perform a preparatory examination on the spatial and kinematic structure of the data. For that, we fit the quasi-isothermal velocity dispersion profile

$$\sigma_\lambda(R) = \sigma_{\lambda,0} \cdot \exp\left(-\frac{R - R_\odot}{h_{\sigma,\lambda}}\right) \qquad \text{with} \quad \lambda \in [R, z] \tag{4.40}$$

and exponential tracer density

$$\rho(R,z) \propto \exp\left(-\frac{R}{h_R} - \frac{|z|}{h_z}\right) \tag{4.41}$$

to the Galactocentric coordinates (R, z, v_R, v_z) of the stars within each MAP by separately maximizing the likelihoods

$$\mathscr{L}(\{R_i, z_i\} \mid h_R, h_z) = \prod_i \frac{\rho(R_i, z_i)}{\int \mathrm{d}\boldsymbol{x}\ \rho(R,z) \times \mathrm{SF}(\boldsymbol{x})} \tag{4.42}$$

$$\mathscr{L}(\{R_i, v_{R,i}\} \mid \sigma_{R,0}, h_{\sigma,R}) = \prod_i \mathscr{N}[\mu = 0, \sigma = \sigma_R(R_i)](v_{R,i}) \tag{4.43}$$

$$\mathscr{L}(\{R_i, v_{z,i}\} \mid \sigma_{z,0}, h_{\sigma,z}) = \prod_i \mathscr{N}[\mu = 0, \sigma = \sigma_z(R_i)](v_{z,i}) \tag{4.44}$$

for the model parameters using the `python` function `scipy.optimize.minimize` and the SF from Section 4.3. For the scale lengths $(h_R, h_{\sigma,R}, h_{\sigma,z})$, we set an upper limit of 16 kpc for the fit. Figure 4.10 shows the best-fit parameters for all the MAPs.

In concordance with the studies by Bovy et al. (2012b,d,c, 2016b), the data exhibits with decreasing [Mg/Fe] an increasing radial tracer scale length, and decreasing vertical scale height as well as velocity dispersion. h_R and h_z both decrease with increasing [Fe/H]. There is a clear trend, that the vertical velocity dispersion decreases with [Fe/H], while the radial velocity dispersion increases with [Fe/H]. The latter is surprising as the radial velocity dispersion in the low-α regime, is naively expected to decrease with metallicity. We will look at this in more detail in Sections 4.6.3 and 4.7.4. The velocity dispersion scale lengths do not show any clear trends, and for some of the MAPs they could not even be properly fitted. We attribute the fact, that we could not constrain the scale lengths very well, to (i) the fact that the distributions in the thin disk are expected to be quite flat anyway and (ii) the small radial extent of the survey volume, over which the velocity dispersion does not vary much and (from inspecting the dispersion in radial bins) shows a large scatter.

4.4.3 Estimates for the Orbital Location

After examining the current physical positions and the SF corrected population distributions of the MAPs in the previous sections, we now estimate which region of the Galaxy is probed by the *orbits* of the stars. In particular, we gauge the mean guiding-center radius $R_g(L_z)$ (Equation 1.9) and the mean maximum height $z_{\max}$ (see, for example, Panel (e) in Figure 1.14) of all orbits in a default MW potential. For simplicity and because a crude estimate is sufficient, we use `galpy`'s `MWPotential2014` and the `Orbit.zmax` and `Potential.rl` methods. Figure 4.11 shows the corresponding mean values for all MAPs. We see very nicely the expected trend (see Section 1.2.3) that more metal-poor and more α-rich stars live in general further out in the Galaxy. The mean R_g's cover for the different MAPs only a range $R_g \in [7.0, 7.8]$ kpc, which is very close to the current radial peak location of the stars in Figures 4.9 and 4.13. This was to be expected, given that we consider kinematically cold thin disk MAPs for which the asymmetric drift is small, i.e., orbits are in their majority near-circular anyway. The distribution of mean $z_{\max}$ with abundance simply mirrors the distribution of h_z in Figure 4.10. The SF restricts the stars to be at $|z| \gtrsim 0.4$ kpc, the orbits reach up to $z_{\max} \sim 0.7 - 0.9$. So $z \sim 0.6$ is most likely the region that we probe best with our stars.

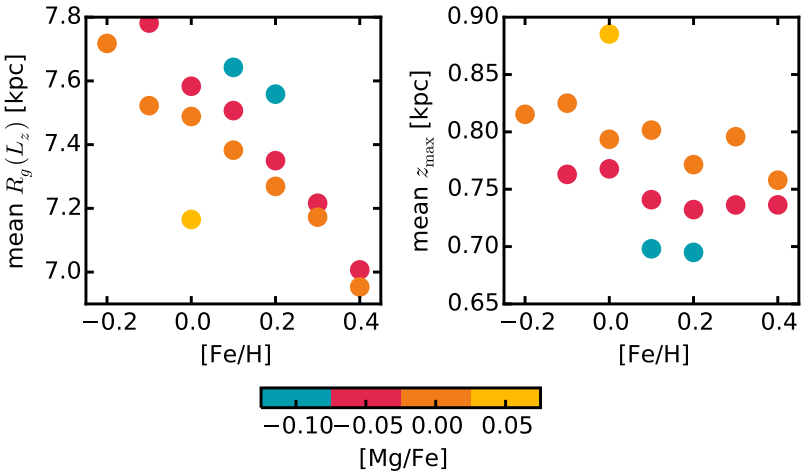


Figure 4.11: Mean guiding-center radius $R_g(L_z)$ (Equation 1.9) and mean maximum height of the orbits $z_{\rm max}$ for the stars in the MAPs considered in this work. The orbital parameters were estimated in the default MW potential `MWPotential2014` by Bovy (2015). This gives a feeling for the region within the Galaxy which is probed by the orbits of the stars.

4.5 Further Modeling Ingredients for the RoadMapping Analysis

In this section, we summarize ingredients used in the RoadMapping analysis that are specific for the analysis of TGAS/RAVE data in this chapter and differ from what we used in the RoadMapping analyses of the previous chapters.

4.5.1 Potential Model

Given the relatively low number of stars per MAP (between $\sim 150-580$ stars; cf. Bovy & Rix (2013) who had $\sim 100-800$ stars per MAP) and its small spatial coverage (see Figure 4.9), we use a simple gravitational potential model with four free parameters to fit to the TGAS/RAVE data,

$$p_\Phi = \{v_{\rm circ}(R_\odot), R_{\rm s,disk}, z_{\rm s,disk}, f_{\rm halo}(R_\odot)\} \tag{4.45}$$

Circular velocity at the Sun. As potential parameter that scales the total mass of the MW and for which we fit, we use, as usual, the total circular velocity at the Sun $v_{\rm circ}(R_\odot)$.

Disk. We use only one disk component to describe the contributions of thin and thick disk stars, gas disk and a potential DM disk. In particular, the *Three Miyamoto-Nagai (3MN) disk approximation to an exponential disk* available in `galpy` (`galpy.potential.MN3ExponentialDiskPotential`), which implements the approximation of an exponential disk

$$\rho(R,z) \propto \exp\left(-\frac{R}{R_{\rm s,disk}}\right) \mathrm{sech}^2\left(-\frac{|z|}{z_{\rm s,disk}}\right) \tag{4.46}$$

by three Miyamoto-Nagai disks (with its flaring radial disk profile; see Equation (3.17)) from Smith et al. (2015). This has two advantages: (1) This disk model has a realistic exponential density profile (Freeman 1970) and a vertical density structure that is widely used to describe galactic disks. It does not flare at larger radii, where a single Miyamoto-Nagai disk predicts too much mass in the disk as compared to an exponential. (2) This 3MN disk potential has a closed form for Φ and allows fast and analytic calculations of gravitational forces. This speeds up the

estimation of actions extensively (compare to Sections 2.2.3 and 3.3.3), which is an advantage to directly using the density profile in Equation (4.46). The free parameters of the disk model are the radial scale length $R_{\rm s,disk}$ and disk scale height $z_{\rm s,disk}$.

Halo. For the halo, we use the classical Navarro, Frenk, and White (NFW) DM halo profile (Navarro et al. 1997)

$$\rho(r) \propto \frac{1}{\frac{r}{a_{\rm halo}}\left(1 + \frac{r}{a_{\rm halo}}\right)^2} \tag{4.47}$$

to incorporate the spherical matter contributions of DM and the potentially negligible mass of the stellar halo. From our study in Chapter 3, specifically Figure 3.12, we know that to get the halo's scale length right (in the realistic case that we cannot 100% trust the disk model to be the perfect model for the disk and given that there might be non-axisymmetries in the data), the data needs to have a large radial extent. As we don't have large radial coverage, we cannot hope to constrain the scale length. So far, there exist also no other reliable measurements for the halo scale length $a_{\rm halo}$ and the concentration $c = r_{\rm vir}/a_{\rm halo}$ of the MW's DM halo. Bland-Hawthorn & Gerhard (2016) estimated that with a concentration of $c \approx 10$, the MW might have a NFW scale length of $a_{\rm halo} = 25$ kpc. Bovy (2015)'s `MWPotential2014`, which is a fit of a simple model to dynamical constraints, contains a NFW halo with scale length $a_{\rm halo} = 16$ kpc at a concentration of $c = 15.3$. Both estimates correspond to a virial radius of the order of ~ 250 kpc (compare also Table 8 in Bland-Hawthorn & Gerhard 2016). McMillan (2011) found as the best-fit NFW halo scale length to several kinematic constraints $a_{\rm halo} = 18 \pm 4.3$ kpc (their Table 2). For simplicity, we set the halo scale length in our NFW fixed to $a_{\rm halo} = 18$ kpc. The halo parameter that we are fitting for in RoadMapping, is the relative contribution of the halo to the radial force at the Sun,

$$f_{\rm halo}(R_\odot) = \left.\frac{F_{\rm R,halo}}{F_{\rm R,disk} + F_{\rm R,halo}}\right|_{R=R_\odot, z=0} = \frac{v^2_{\rm circ,halo}(R_\odot)}{v^2_{\rm circ,disk}(R_\odot) + v^2_{\rm circ,halo}(R_\odot)}. \tag{4.48}$$

This parameter is motivated by the definition of the rotational support in studies that investigate if the disk is maximal (Sackett 1997; see also Section 1.1.5). Together with $v_{\rm circ}(R_\odot)$, it determines the total mass of the halo and its concentration.

Bulge. The data in the solar neighborhood will not constrain the bulge, because its contribution to the total gravitational force is quite small. We therefore keep the bulge in the modeling fixed. We use a Hernquist bulge (see Equation (3.1)) with scale length $a_{\rm bulge} = 600$ pc and $M_{\rm bulge,tot} = 10^{10}$ M$_\odot$. Bovy & Rix (2013) used a similar bulge model; their total mass of 4×10^9 M$_\odot$ was, however, a bit lower. They chose this mass to achieve a similar rotation curve with their Hernquist bulge as compared to when a realistic cut-off power-law bulge model is used (e.g. McMillan 2011). The effect on the modeling result will not depend much on the exact choice of the bulge model, so we set the mass in the initial test to 10^{10} M$_\odot$ as a compromise between Bovy & Rix (2013) and the total stellar bulge mass of $1.4 - 1.7$ M$_\odot$ recorded in Portail et al. (2015); Bland-Hawthorn & Gerhard (2016).

All fixed and free model parameters are summarized in Table 4.1.

4.5.2 Distribution Function for the Disk and Outlier Model for the Halo

Motivating the use of the qDF for the stellar disk. To model the MW stars, we will continue using the qDF in the form of Binney & McMillan (2011) given in Equations (1.58)-(1.65) that proved to be successful in Chapter 3 and Bovy & Rix (2013). Work by Bovy et al. (2016b) suggests, however, that instead of using a purely exponential radial density profile, we should

rather use a more general form of the qDF for our low-[Mg/Fe] sample of MW stars that also allows for a break-radius and inverse scale lengths. Given that our data only has a small radial extent around the solar radius (which means that the determination of radial scale lengths will be difficult anyway, see Figure 4.10) and is mostly metal-rich (and therefore has the break-radius at smaller radii; Bovy et al. 2016b), we think that the qDF in its original form is a good starting point for an initial analysis of the Gaia data.

Outliers and a model for the stellar halo. There are several aspects why we have good reason to believe that our data is contaminated. First, the RC selection following Bovy et al. (2014) is suspected to have a $\sim 5\%$ contamination of RGB stars. For these stars, the distance estimations will be slightly wrong. Second, the chemical abundances from RAVE-on are expected to have uncertainties of ~ 0.07 dex (see Casey et al. 2017 and Section 4.2.3), while our MAP bins have $\Delta[\mathrm{Fe/H}] = 0.1$ dex and $\Delta[\mathrm{Mg/Fe}] = 0.05$ dex. We therefore expect a non-negligible contamination by stars from neighboring MAPs, especially from different [Mg/Fe], which will for the low-[Mg/Fe] MAPs cause a rise in velocity dispersion. Also, above [Fe/H] = 0.26 the *Cannon* abundances in RAVE-on are extrapolations only, so the contamination in this regime might be even larger. Third, not all stars in the solar neighborhood are actually disk stars. We also expect a fraction of $\sim 0.5\%$ of stars from the MW's stellar halo (Jurić et al. 2008), of which also some are metal-rich (e.g., Zoccali et al. 2003; Bonaca et al. 2017).
The proper way to capture outlier stars in our modeling, is to use the qDF together with an outlier model. We decided to only use a very simple outlier model that describes a stellar halo component, and leave the other possible outliers un-modelled. Our outlier model has a constant spatial density, is non-rotating[24] and has a high isotropic velocity dispersion that decreases outwards[25] (see Section 1.2.1). In particular, we use an outlier DF

$$\begin{aligned}
\mathrm{DF}_{\mathrm{out}}(\boldsymbol{x},\boldsymbol{v}) &= p_{\mathrm{halo},\rho}(\boldsymbol{x}) \times p_{\mathrm{halo,vel}}(v_R \mid R) \times p_{\mathrm{halo,vel}}(v_T \mid R) \times p_{\mathrm{halo,vel}}(v_z \mid R) && (4.49)\\
\text{with } p_{\mathrm{halo},\rho}(\boldsymbol{x}) &= \text{const.} && (4.50)\\
p_{\mathrm{halo,vel}}(v_\lambda \mid R) &= \frac{1}{\sqrt{2\pi}\sigma_{\mathrm{out}}(R)} \exp\left(-\frac{v_\lambda^2}{2\sigma_{\mathrm{out}}(R)^2}\right) \quad \text{with} \lambda \in [R,T,z] && (4.51)\\
\text{and } \sigma_{\mathrm{out}}(R) &= \sigma_{\mathrm{out}} \times \exp\left(-\frac{R-R_\odot}{h_{\mathrm{out}}}\right) && (4.52)\\
\sigma_{\mathrm{out}} &= 100 \text{ km s}^{-1} && (4.53)\\
h_{\mathrm{out}} &= 6 \text{ kpc}, && (4.54)
\end{aligned}$$

analogous to Bovy & Rix (2013). We introduce the outlier fraction p_{out}, which describes the probability that a star is not a disk star, but rather a halo star or otherwise an outlier, or in other words, the fraction of halo-outlier stars in the sample. In addition to the qDF parameters

$$p_{\mathrm{DF}} \equiv \left\{\ln h_R^{\mathrm{qDF}}, \ln \sigma_{R,0}^{\mathrm{qDF}}, \ln \sigma_{z,0}^{\mathrm{qDF}}, \ln h_{\sigma,R}^{\mathrm{qDF}}, \ln h_{\sigma,z}^{\mathrm{qDF}}\right\} \tag{4.55}$$

we will also fit for the outlier fraction

$$p_{\mathrm{out}} \tag{4.56}$$

in the RoadMapping modeling.

We summarize all model parameters again in Table 4.1.

[24] However, Bonaca et al. (2017) found that the in-situ formed, metal-rich component of the stellar halo might actually be rotating.

[25] For the solar neighborhood, the Galactocentric spherical radius r is approximately equal to the cylindrical radius R.

Table 4.1. Summary of the fixed and free model parameters for the RoadMapping analysis of the TGAS/RAVE data.

Category	Component	Parameter		Value	Fitted?	Prior
Φ	**Rotation curve**	circular velocity at the Sun	$v_{\rm circ}(R_\odot)$	-	free	flat
		slope of $v_{\rm circ}$ at the Sun	$\frac{{\rm d}\ln v_{\rm circ}}{{\rm d}\ln R}$	-	derived	Equation (4.67)
	Exponential disk	radial scale length	$R_{\rm s,disk}$	-	free	flat
	(Equation 4.46)	vertical scale height	$z_{\rm s,disk}$	-	free	flat
	(Smith et al. 2015)					
	NFW halo	halo scale length	$a_{\rm halo}$	18 kpc	fixed	-
	(Equation 4.47)	halo fraction at the Sun	$f_{\rm halo}(R_\odot)$	-	free	flat
	(Navarro et al. 1997)					
	Hernquist bulge	bulge scale length	$a_{\rm halo}$	600 pc	fixed	-
	(Equation 3.1)	total bulge mass	$M_{\rm bulge,tot}$	10^{10} M$_\odot$	fixed	-
	(Hernquist 1990)					
DF	**quasi-isothermal**	radial tracer scale length	$h_R^{\rm qDF}$	-	free	log. flat in [0.5,20] kpc
	distribution function	radial velocity disp. at Sun	$\sigma_{\rm R,0}$	-	free	log. flat
	(qDF)	vertical velocity disp. at Sun	$\sigma_{\rm z,0}$	-	free	log. flat
	(Equations (1.58)-(1.65))	scale length of $\sigma_R(R)$	$h_{\sigma,R}$	-	free	log. flat in [0.5,20] kpc
	(Binney & McMillan 2011)	scale length of $\sigma_z(R)$	$h_{\sigma,R}$	-	free	log. flat in [0.5,20] kpc
outlier	**simple stellar**	outlier fraction	$p_{\rm out}$	-	free	flat in [0,1]
model	**halo model**	velocity disp. at Sun	$\sigma_{\rm out,0}$	100 $\frac{\rm km}{\rm s}$	fixed	-
	(Equations 4.49-4.51)	scale length of $\sigma_{\rm out}(R)$	$h_{\rm out}$	6 kpc	fixed	-

4.5.3 Likelihood and Priors

The joint likelihood. The likelihood of the ith star to be drawn from the model, which consists of a disk model, specified by $\mathrm{qDF}(\boldsymbol{x},\boldsymbol{v} \mid p_M)$ with the current model parameters p_M, and a halo outlier model $\mathrm{DF}_{\mathrm{out}}(\boldsymbol{x},\boldsymbol{v})$ with p_{out} giving the fraction of halo/outlier stars, is (following Equation (17) in Hogg et al. 2010, and Equation (26) in Bovy & Rix 2013),

$$\mathscr{L}_i = (1-p_{\mathrm{out}})\mathscr{L}_{i,\mathrm{disk}} + p_{\mathrm{out}}\mathscr{L}_{i,\mathrm{out}}, \tag{4.57}$$

where

$$\mathscr{L}_{i,\mathrm{disk}} = \frac{1}{(r_o\cdot v_o)^3}\frac{\mathrm{qDF}(\boldsymbol{x}_i,\boldsymbol{v}_i \mid p_M)}{\int \mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{v}\ \mathrm{qDF}(\boldsymbol{x},\boldsymbol{v}\mid p_M)\times \mathrm{SF}(\boldsymbol{x})} \tag{4.58}$$

as in Equation (2.3) and

$$\begin{aligned}\mathscr{L}_{i,\mathrm{out}} &= \frac{1}{(r_o\cdot v_o)^3}\frac{\mathrm{DF}_{\mathrm{out}}(\boldsymbol{x}_i,\boldsymbol{v}_i)}{\int \mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{v}\ \mathrm{DF}_{\mathrm{out}}(\boldsymbol{x},\boldsymbol{v})\times \mathrm{SF}(\boldsymbol{x})} &(4.59)\\ &= \frac{1}{(r_o\cdot v_o)^3}\frac{\mathrm{DF}_{\mathrm{out}}(\boldsymbol{x}_i,\boldsymbol{v}_i)}{p_{\mathrm{halo},\rho}\times\int \mathrm{d}\boldsymbol{x}\ \mathrm{SF}(\boldsymbol{x})}. &(4.60)\end{aligned}$$

In the last step we used Equations (4.49)-(4.51). Both distribution functions, qDF and $\mathrm{DF}_{\mathrm{out}}$, need to be properly normalized for the likelihood by integrating the spatial density over the SF. The prefactor $(r_o\cdot v_o)^{-3}$ is needed for the likelihood to have the correct units, i.e., to satisfy

$$\int \mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{v}\ \mathscr{L}(\boldsymbol{x},\boldsymbol{v}\mid p_M) = 1 \tag{4.61}$$

and therefore

$$[\mathscr{L}(\boldsymbol{x},\boldsymbol{v}\mid p_M)] = [\mathrm{length}]^{-3}\times[\mathrm{velocity}]^{-3}. \tag{4.62}$$

As the likelihood is evaluated in `galpy`'s natural units, where

$$\begin{aligned}[\mathrm{length}] &= R_\odot &(4.63)\\ [\mathrm{velocity}] &= v_{\mathrm{circ}}(R_\odot) &(4.64)\end{aligned}$$

and $v_{\mathrm{circ}}(R_\odot)$ is one of the free model parameters, the units of the likelihood are important and we divide the likelihood by $(r_o\cdot v_o)^3$ with

$$\begin{aligned}r_o &= R_\odot/8\ \mathrm{kpc} &(4.65)\\ v_o &= v_{\mathrm{circ}}(R_\odot)/220\ \mathrm{km\ s^{-1}}. &(4.66)\end{aligned}$$

(We took care of the units also in Chapter 2 and 3, but we mention it only here for completeness.) Because we now use a proper outlier model in the analysis, we are not applying the robust likelihood strategy from Chapter 3 and Equation (3.12) anymore.

Priors. In Chapter 3, specifically Figure 3.10, we found that it could be difficult to get the slope of the rotation curve correct for survey volumes with radial coverage of 2 kpc or less, like our data. To help the modeling, we therefore use a prior on the slope of the rotation curve. Bovy et al. (2012a) have measured the MW's circular velocity curve from dynamical modeling of APOGEE data with a radial coverage of 4 kpc $< R <$ 14 kpc and precise line-of-sight velocities. They assumed an axisymmetric rotation model and corrected for the asymmetric drift of the kinematically warm stellar tracers. In their Figure 5, they show the posterior PDF of the slope of their best-fit rotation curve, which preferred flatness. Bovy & Rix (2013) described this PDF

with an approximation of the form

$$p(x) = \begin{cases} 0 & \text{if } x > 0.04, \\ W \exp(-W) & \text{otherwise.} \end{cases} \tag{4.67}$$

$$\text{where } \frac{\mathrm{d} \ln v_{\mathrm{circ}}(R_\odot)}{\mathrm{d} \ln R} \equiv x \text{ is the slope of the rotation curve} \tag{4.68}$$

$$\text{and with } W \equiv \left(1 - \frac{x}{0.04}\right) \tag{4.69}$$

(their Equation 41). We use this as a prior in our RoadMapping modeling as well.
We saw in Figure 4.10 that the relatively small radial coverage of the data will make it potentially difficult to determine radial trends in the velocities. Also, as discussed in Section 4.5.2, the tracer scale length of our current form of the qDF might prefer $\ln h_R^{\mathrm{qDF}} \to \infty$ (Bovy et al. 2016b). Some initial tests showed that it helps the RoadMapping modeling to achieve convergence of the MCMC, if we impose a flat prior on all three qDF scale lengths in the range

$$h_R^{\mathrm{qDF}}, h_{\sigma,R}^{\mathrm{qDF}}, h_{\sigma,z}^{\mathrm{qDF}} \in [0.5, 20] \text{ kpc}. \tag{4.70}$$

4.5.4 Additional Details

Position of the Sun in the Galaxy. One of the major assumptions in the modeling is, that we assume the position and velocity within the Galactocentric rest-frame to be fixed and known. We set it to be

$$R_\odot = X_{\mathrm{GC},\odot} = 8 \text{ kpc} \quad \text{(Ghez et al. 2008)} \tag{4.71}$$

$$Y_{\mathrm{GC},\odot} = 0 \quad \text{(by definition)} \tag{4.72}$$

$$Z_{\mathrm{GC},\odot} = 25 \text{ pc} \quad \text{(Jurić et al. 2008)} \tag{4.73}$$

$$v_{X,\mathrm{GC},\odot} = -11.1 \text{ km s}^{-1} \quad \text{(Schönrich et al. 2010)} \tag{4.74}$$

$$v_{Y,\mathrm{GC},\odot} = \Omega_{\mathrm{GC},\odot} \times R_\odot \tag{4.75}$$

$$v_{Z,\mathrm{GC},\odot} = 7.25 \text{ km s}^{-1} \quad \text{(Schönrich et al. 2010)} \tag{4.76}$$

with $\Omega_{\mathrm{GC},\odot} = 30.24 \text{km s}^{-1}\text{kpc}^{-1}$ (Bland-Hawthorn & Gerhard 2016) being derived from the proper motion of Sagittarius A* in the Galactic center. The Solar phase-space coordinates enter RoadMapping (i) when transforming the heliocentric observables into the Galactocentric data used in the modeling, and (ii) setting up the SF within the Galaxy. Especially fixing $R_\odot$ will lead to systematics, because it sets both the length and velocity scale (via $v_{Y,\mathrm{GC},\odot} \sim v_{\mathrm{circ}}(R_\odot) + V_\odot$) of the data and modeling. The current overall best estimate from several studies by Bland-Hawthorn & Gerhard (2016) is $R_\odot = 8.2 \pm 0.1$ kpc, a bit larger than the default $R_\odot = 8$ kpc we are using. One task for the future will be to re-run the RoadMapping modeling with a slightly different $R_\odot$ to estimate the systematic error introduced by its uncertainty.

Measurement uncertainties. As the TGAS uncertainties in positions are basically negligible, we set for simplicity $\sigma_{\mathrm{R.A.}} = \sigma_{\mathrm{Dec.}} = 0$ in the RoadMapping analysis. We do not expect this to bias the modeling result at all. In RoadMapping we do take into account the 3D velocity error ellipse, spanned by the correlated TGAS $(\mu_{\mathrm{R.A.}*}, \mu_{\mathrm{dec.}})$ together with the 5% uncertainty in the photometric distance d_{RC} (see Section 4.2.4), and the independent uncertainty in the v_{los} measurement by RAVE. We sample this error ellipse with 1000 MC samples, which should be sufficient, given that this is more samples than in all tests in Chapter 2 with at the same time smaller proper motion uncertainties. Because we have a maximum of 600 stars per MAP, the computational expense of this high sampling of the error ellipse is small.

Numerical precision of the likelihood normalisation. The SF in Figure 4.8 has a lot of small scale structure as compared to the spherical SFs with constant completeness that we used in Chapters 2 and 3. It is therefore worth to confirm explicitly, that our adopted numerical precision ($N_x = 20, N_v = 28, n_\sigma = 5.5$) for calculating the likelihood normalisation in Section 2.2.8 is high enough. We proceeded analogously to Figure 2.4 and were able to show that, for a potential and qDF similar to our best-fit in Figure 4.12 in the next section, and for 600 stars, the adopted precision is more than sufficient.
The integral over (R, z) in Equation (4.30) is, as always, performed with a 40th-order Gauss-Legendre integration. We run some tests with RoadMapping using the SF in Figure 4.8 together with mock data created analogously to Section 2.2.6. In these tests, we properly recovered the six free parameters of the mock data model, $v_{\rm circ}(R_\odot)$ and the qDF parameters, indicating that the implementation of the SF and the normalisation integral over (R, z) work indeed correctly.

Nested-grid search and fiducial qDF. In Section 2.2.10 we have discussed the importance of adapting the fiducial qDF, required for the stability of the normalisation calculation, during the nested-grid search of the parameter space. To save computation time, we only run 15 iterations of the nested-grid search, before starting the MCMC (see Section 2.2.10). We use the estimates in Figure 4.10 to initiate the first iteration of the grid-search. The velocity dispersion scale lengths were however hard to measure from the data, so we set the initial $h_{\sigma,R} = 8$ kpc and $h_{\sigma,z} = 7$ kpc.

4.6 Results from RoadMapping Modeling of TGAS/RAVE Data

We use the RC MAP data from TGAS/RAVE in Section 4.2 and model it with RoadMapping in the SF from Section 4.3 using the potential model and modeling ingredients from Section 4.5. This section summarizes the results.

4.6.1 First Considerations about the Best-fit Model

The RoadMapping modeling of each of the the MAPs took between 4 and 8 days on 32 cores.[26] Figure 4.12 summarizes the best-fit model parameters for all 16 MAPs.

The qDF parameters. The right column in Figure 4.12 shows the best-fit qDFs. The qDF parameters are the parameters of scaling profiles (see Section 1.5.2) and cannot be directly compared to the values from the non-dynamical fit in Figure 4.10. We still see the same trends with abundance in $h_R^{\rm qDF}$, $\sigma_{z,0}^{\rm qDF}$, and $\sigma_{R,0}^{\rm qDF}$ as in Figure 4.10. Given that we fit the qDF parameters in log-space and show 1σ error bars in Figure 4.12, the scale lengths are overall not very well constrained and some are even pegged at the upper limit of the prior (black dotted lines). This was to be expected, given the relatively small radial extent of the data, $R \in [7, 9]$ kpc, and the low number of stars per MAP.

The potential parameters. In the potential parameters (left column of Figure 4.12), the first thing to notice is, that the different MAPs all give very similar results on the parameters (we will discuss this later also in Figure 4.14). The MAP with the highest number of stars is the one at [Fe/H] $= 0.1$, [Mg/Fe] $= -0.05$ (red). All other MAPs with a lower number of stars are scattered around this MAP's potential estimate. Secondly, there are no clear trends with abundance. Did we expect any? In Chapter 3 we found that the modeling is sensitive to the

[26] The duration of the run-time depended on whether the MCMC converged during the first run. Some data sets had to run longer, and some required flat priors on the qDF scale lengths to achieve convergence (see Section 4.5.3).

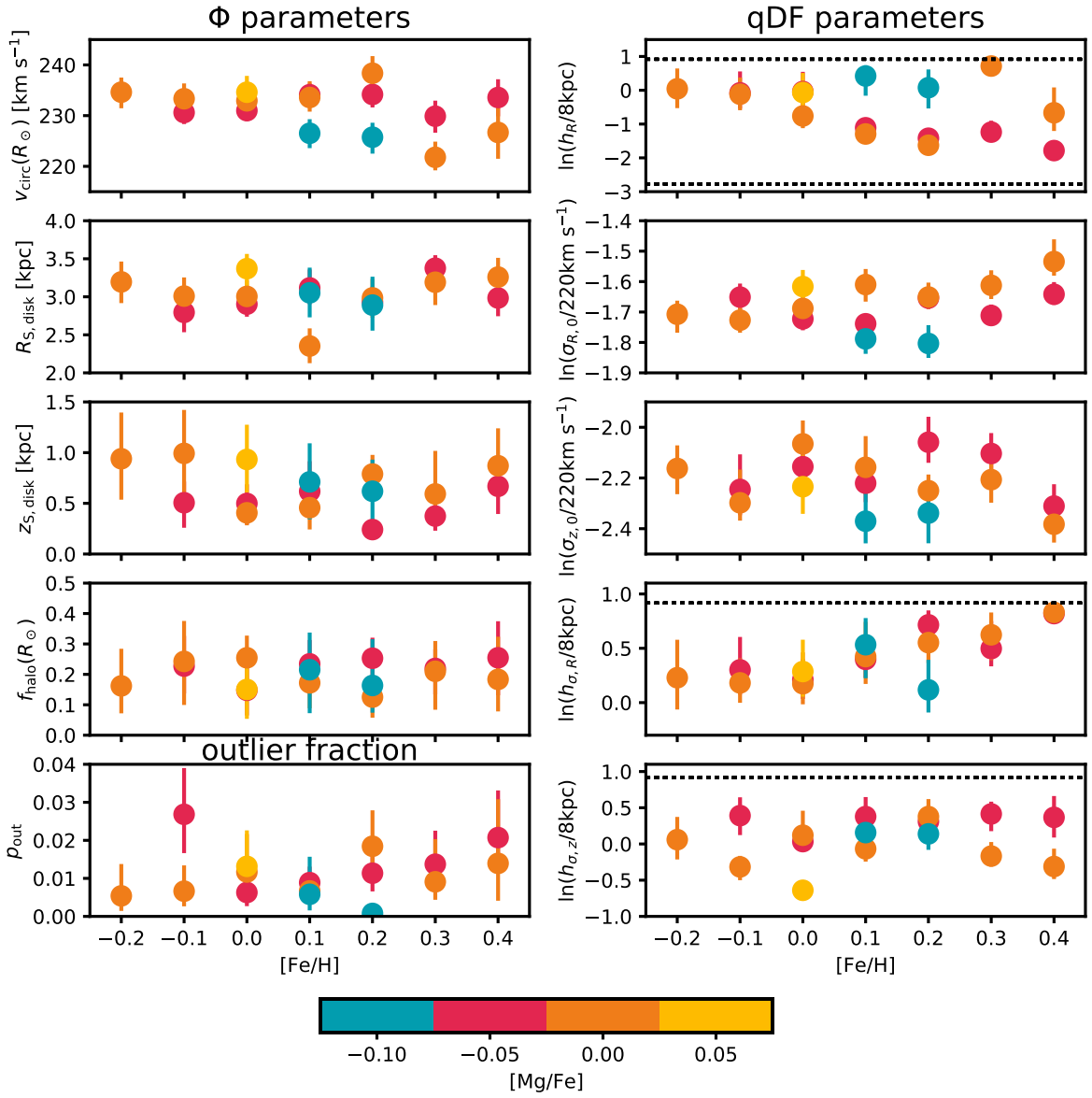


Figure 4.12: Overview of the best-fit model parameters for the MW gravitational potential Φ, the qDF, and the outlier model (summarized in Table 4.1) derived with RoadMapping for the 16 independent MAPs of TGAS/RAVE RC stars shown in Figure 4.9. The x-axis denotes the [Fe/H], and the color the [Mg/Fe] of the corresponding MAP. We show the median together with the [16th,84th] percentiles of the marginalized posterior PDF for each model parameter. The horizontal black dotted lines mark the flat prior in [0.5,20] kpc that we imposed on the qDF scale lengths $h_R, h_{\sigma,R}, h_{\sigma,z}$ (see Section 4.5.3). We see similar trends with abundance in the qDF parameters that we already found in Figure 4.10. The potential parameters derived from different MAPs are all more or less consistent with each other, and there are no clear trends with abundance. The outlier fraction is with 1-3% very low and consistent with the expected low number of halo stars that we tried to capture with the outlier model (see Section 4.5.2).

gravitational forces at the peak location of the data. MAPs of different chemical abundances have, according to their different scale lengths and heights, different mean orbital (R, z) and constrain the potential in different regions (see also Figure 2.18 in Chapter 2). However, the chemical and spatial range probed by our data is so small (see Figure 4.11) that this effect is modest at best, and we expect randomness due to low-number statistics to dominate. The only exception are maybe the $v_{\rm circ}(R_\odot)$ measurements of the [Mg/Fe] = −0.1 MAPs (blue), which is

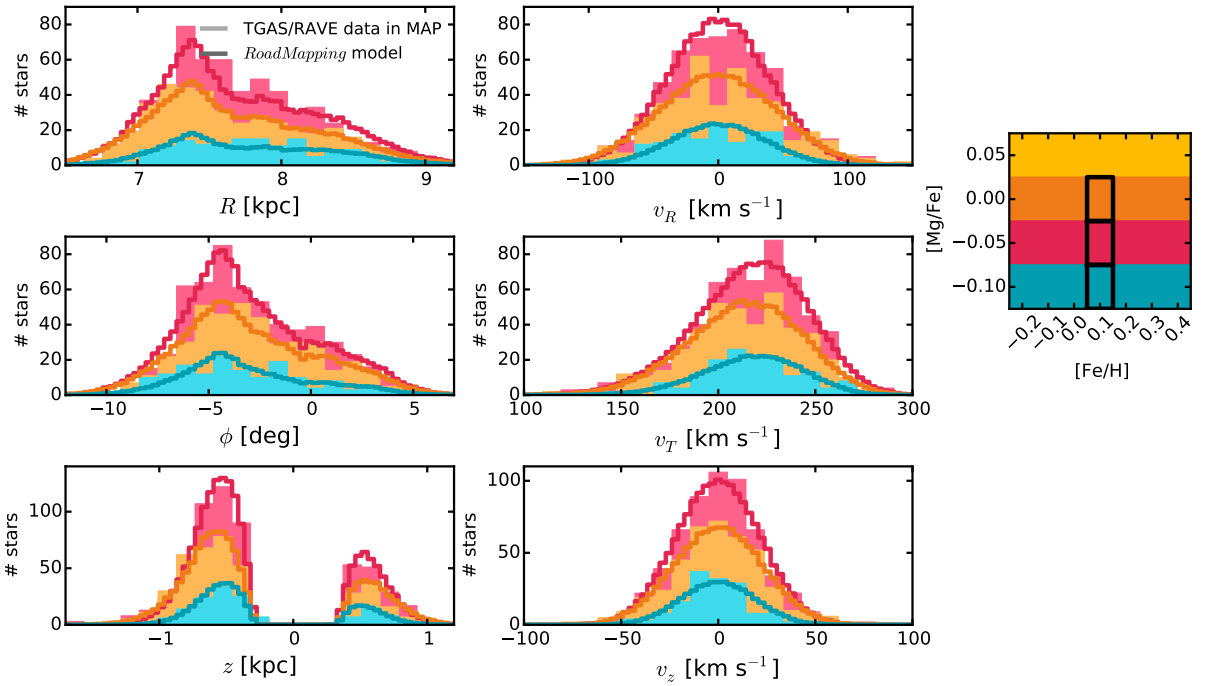


Figure 4.13: Comparison of the one-dimensional distributions in Galactocentric position-velocity space of three example MAPs of the TGAS/RAVE data (bright histograms) with the best-fit model from RoadMapping (dark line). The (R, z, v_R, v_T, v_z) as shown here entered RoadMapping as data. The model distribution was generated by MC sampling of the best-fit qDF in the best-fit potential (with the peak parameters from Figure 4.12), by shaping the distribution according to the SF in Section 4.3.3 (analogous to Section 2.2.6), and scaling it to the number of stars in the original MAP. All MAPs show good agreement between data and model; here we only present the MAPs with [Fe/H] = 0.1 and [Mg/Fe] = $[-0.1, -0.05, 0.0]$ (as color-coded in the right-panel). The [Mg/Fe] = 0.0 model exhibits a slightly better fit to the radial profile than in the [Mg/Fe] = −0.05 MAP; all other MAPs (not shown here) are somewhere in between.

lower by a few km s^{-1}. We attribute this to these data sets having considerably lower h_z (Figure 4.10) and mean $z_{\rm max}$ (Figure 4.11). It appears, that we measure a lower $v_{\rm circ}(R_\odot)$ closer to the plane of the disk. This will need further investigation with larger or more extended data sets in the future.

The outlier fraction. The outlier model mimicking a halo population (see Section 4.5.2) prefers fractions of only 1-3% for all MAPs. This is encouraging because it means that the qDF is overall a good model for the data. Also, we only expected a maximum of 5% halo stars in the Solar neighborhood. We note, that the RAVE-on abundance measurements for stars with [Fe/H] ≥ 0.3 are less reliable, because this is outside of the *Cannon* training set, and we expect more contamination. This could explain the slight trend towards higher outlier fractions with metallicity.

Comparing the best-fit model to the data. Do these model parameters actually correspond to physical tracer distributions that are a good match for the data? To test this, we show in Figure 4.13 for three selected MAPs the one-dimensional histograms for the Galactocentric cylindrical data coordinates that entered the RoadMapping analysis. We compare this to the corresponding model implied by the peak best-fit Φ and qDF in Figure 4.12 found with RoadMapping and shaped by the SF. Overall, the fit was indeed successful given the low number of stars.
The ϕ coordinate was not used in the axisymmetrc RoadMapping modeling. The comparison

between data and model is therefore an independent test of the SF and how axisymmetric the data really is. Given that there is a good agreement in the azimuthal distribution, we deduce that the data, which is located away from the Galactic plane by construction, is not strongly affected by non-axisymmetries. In addition, this means that the SF is powerful in describing the data distribution.
The recovery of the radial tracer density profile of the example MAP at [Mg/Fe] = 0 (orange) is very good, the one at [Mg/Fe] = −0.05 (red) slightly less so[27], but still acceptable. We have inspected the fits of all MAPs and confirmed that all MAPs are equally well fitted. The MAP at [Mg/Fe] = −0.1 has an almost flat radial profile, which cannot be properly recovered with the qDF. This is also mirrored in Figure 4.12 by the $h_R^{\rm qDF}$ parameter pegged at the prior's upper limit of 20 kpc. This confirms the findings by Bovy et al. (2016b), who found that low-α, Solar metallicity populations have flat density profiles.
The overall velocity distribution is well-recovered, even though the data velocities are affected by uncertainties, while the model velocities were derived from a uncertainty-deconvolved likelihood fit. This demonstrates how small the velocity uncertainties have become in the age of Gaia.
We have visually inspected also the two-dimensional residuals between data and model, but found no indications that there are obvious or systematic trends in misfitting the data.

4.6.2 Our Best-fit Estimate for the Galactic Gravitational Potential

On combining potential estimates of different MAPs. Bovy & Rix (2013) combined their independent potential estimates from different MAPs by using only the $\Sigma_{\rm tot,1.1kpc}$ measurements at one best radius for each MAP. In particular, they determined the radius at which the correlation between the surface density profile and disk scale length implied by the PDF of a given MAP was minimal. They presumed that this was the radius at which $\Sigma_{\rm tot,1.1kpc}(R)$ could be constrained most robustly (see also Sections 1.5.3 and 4.7.2). This was necessary, because their potential model had only two free parameters and they were only modeling the vertical motions of the data. The individual potential constraints were therefore lacking generality. Given that their MAPs spanned a wide range in abundance and scale radii, and had a large radial extent, they could cover also a large range of radii with these $\Sigma_{\rm tot,1.1kpc}$ and derived the beautiful surface density profile shown in Figure 1.20.
Our modeling has different strengths and weaknesses. We have four free potential parameters, which means that our model spans a much larger space of possible potentials. We also model the full phase-space distribution of stars having very precise measurements. That means, we expect also to get good constraints on the radial, not only the vertical forces. On the other hand, as compared to Bovy & Rix (2013), our data set is tiny considering the abundance and spatial extent. Considering our results in Chapter 3, we are optimistic to get at least local constraints on the overall gravitational potential.
In our situation, the proper way to combine the different MAP measurements is to multiply all their PDFs: The qDF parameters of different MAPs are independent of each other, so we treat them as nuissance parameters and marginalize over them. The potential parameters of the different analyses, however, describe the same potential. And indeed, in Figure 4.12 we saw already that we got very similar potential constraints for all MAPs. This allows us to directly multiply the potential PDFs. The number of stars per MAP is very different; but according to the central limit theorem this translates directly into the width of the PDF. MAPs with less stars will therefore also have less influence on the composite PDF.

Our joint constraint on the MW potential. The potential PDFs are only available in the form of unnormalized MCMC samples, so we fitted a four-dimensional multivariate normal distribution

[27] This could also be due to the absolute scaling of the best-fit model.

Table 4.2. Summary of the RoadMapping best-fit potential parameters, and comparison with estimates from the literature.

Potential parameter		Best RoadMapping fit	Estimate from the literature
Circular velocity at the Sun	$v_{\rm circ}(R_\odot)$	231.4 ± 0.7 km s^{-1}	238 ± 15 km s^{-1} (Bland-Hawthorn & Gerhard 2016), §6.4.2
Disk scale length	$R_{\rm s,disk}$	3.01 ± 0.05 kpc	2.6 ± 0.5 kpc$^{(a)}$ (Bland-Hawthorn & Gerhard 2016), §5.1.2
Disk scale height	$z_{\rm s,disk}$	490 ± 50 pc	300^{+150}_{-80} pc$^{(a)}$ (Bland-Hawthorn & Gerhard 2016), §5.1.1
Halo fraction at the Sun	$f_{\rm halo}(R_\odot)$	0.22 ± 0.02	$0.3^{+0.1}_{-0.09}$ (Bovy & Rix 2013), Fig. 21
Total surface density at 1.1 kpc	$\Sigma_{\rm tot,1.1kpc}(R_\odot)$	98 ± 3 M$_\odot$pc^{-2}	70 ± 5 M$_\odot$pc^{-2} (Bland-Hawthorn & Gerhard 2016), §5.4.2
Rotational support of the disk	$\left.\frac{v^2_{\rm circ,disk}}{v^2_{\rm circ,tot}}\right\vert_{R=2.2R_{\rm s,disk}}$	$72\% \pm 2\%$	$69\% \pm 6\%^{(b)}$ (Bovy & Rix 2013)
Slope of the rotation curve	$\mathrm{d}(\ln v_{\rm circ})/\mathrm{d}(\ln R)\vert_{R_\odot}$	-0.07 ± 0.01	-0.06 ± 0.05 (Bovy & Rix 2013)
Local DM density	$\rho_{\rm DM,\odot}$	0.0049 ± 0.0005 M$_\odot$ pc^{-3}	0.0065 ± 0.0023 M$_\odot$ pc^{-3} (Zhang et al. 2013b)

$^{(a)}$Thin disk only.

$^{(b)}$At 2.2 times a scale length of 2.15 kpc.

Note. — The four quantities at the bottom of the table are derived quantities that we did not explicitly fit for.

to each of the potential PDFs. As can be seen in Figure 4.14, there are covariances between the potential parameters (especially between $z_{\rm s,disk}$ and $f_{\rm halo}(R_\odot)$, which together determine the local disk-to-halo mass budget and the vertical mass distribution), but no strong degeneracies of non-Gaussian shape. The covariant normal distribution that we get from multiplying all normalized PDF Gaussians is shown in black in Figure 4.14. This potential constraint derived from using 5,005 stars has very high precision and the best-fit parameters are summarized in Table 4.2.

We also compare it to estimates from the literature, both in Figure 4.14 and Table 4.2. $v_{\rm circ}(R_\odot)$ is fully consistent with previous measurements and has even much higher precision. $R_{\rm s,disk}$ and $f_{\rm halo}(R_\odot)$ are also nicely consistent with the estimates from the literature (even though they only lie just inside of the uncertainty boundary). The peak estimate for our disk scale height $z_{\rm s,disk}$ is still close, but ~ 50 pc outside of the previously measured range of thin disk scale heights. We note that we used only one disk component in our potential model, so the scale height we got is a combination of contributions from thin, thick, and gas disks, which might not be well described by only one $\mathrm{sech}^2(-|z|/z_{\rm s,disk})$ vertical profile (Equation 4.46). Because the thin disk is the most massive disk component in the Solar neighborhood (Bland-Hawthorn & Gerhard 2016, §5.1.3, thick to thin disk surface density ratio at $R_\odot$: $12\% \pm 4\%$), we chose to compare it here to our composite disk estimate. Overall, we get realistic values for the MW potential parameters with high precision.

Rotation curve and disk maximality. The first panel in Figure 4.15 shows the circular velocity curve implied by the composite over all best-fit in Table 4.2. The circular velocity curve is very tightly constrained and has an almost flat slope of

$$\left.\frac{\mathrm{d}(\ln v_{\rm circ})}{\mathrm{d}(\ln R)}\right|_{R=R_\odot} = -0.07 \pm 0.01 \tag{4.77}$$

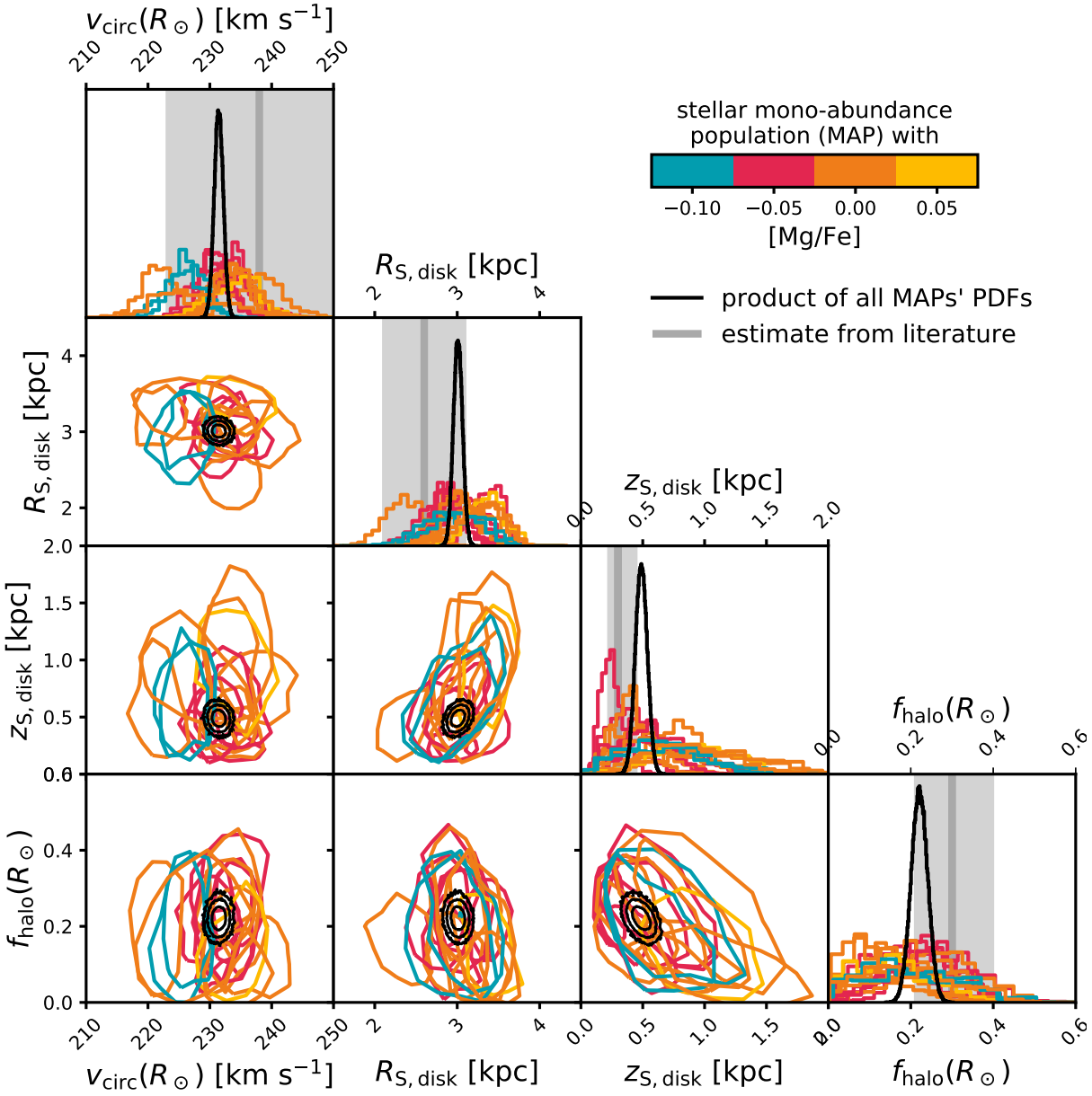


Figure 4.14: Corner plot showing the posterior PDF for the model parameters of the MW's gravitational potential for all 16 MAPs, color-coded by their [Mg/Fe]-abundance. The colored contours enclose 68% of the probability for each independent potential constraint and they are consistent with each other for all MAPs. We have fitted a multivariate Gaussian to each of the PDFs (projected to the potential parameters only, treating the qDF parameters and outlier fraction as nuissance parameters for each MAP) and multiplied all of them. The composite constraint on the potential from all MAPs is shown in black (1-, 2- and 3-sigma contours). Estimates from the Literature are overplotted in the panels with the one-dimensional PDF projections of the potential parameters. The composite best-fit and the literature estimates are summarized in Table 4.2. Except of maybe the disk scale height which we overestimate by a bit, our estimates have all high precision and are consistent with previous estimates for the MW. A flip book version of this figure can be found at the end of this thesis in the Bibliography.

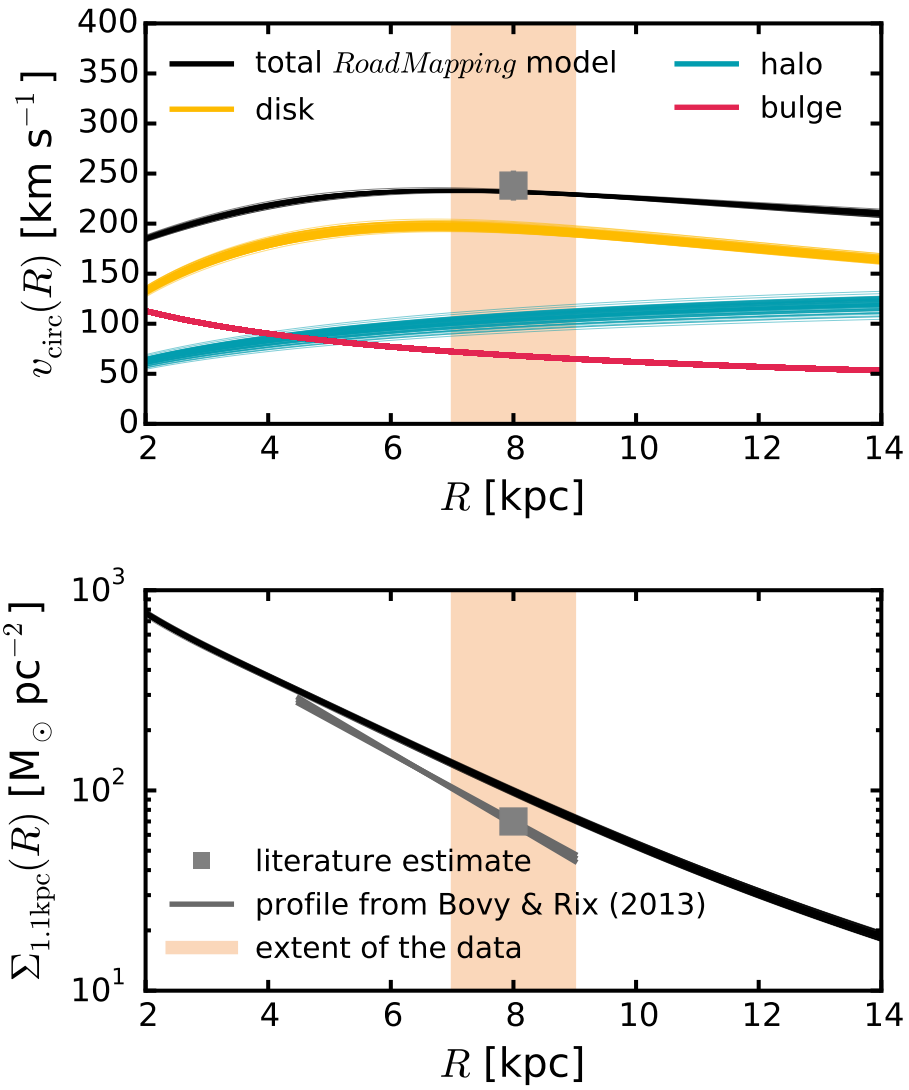


Figure 4.15: Circular velocity curve and radial surface density profile at $z = 1.1$ kpc for the composite best-fit model from RoadMapping and the TGAS/RAVE RC data in Table 4.2. From the PDF marked black in Figure 4.14, we drew 100 different MC potential samples that are individually plotted here to indicate the width of the constraint implied by the PDF. The overall rotation curve is very tightly constrained, flat, as proposed by the prior, and agrees well with the estimate from literature. We kept the bulge fixed in the analysis, as well as the halo scale length. The model prefers a maximum disk with rotational support of $\sim 72\%$ at 2.2 disk scale lengths, in agreement with previous measurements. The surface density profile is with $\Sigma_{\rm tot,1.1kpc} \sim 98$ M$_\odot$pc^{-2} more massive than previous estimates, which prefer $\Sigma_{\rm tot,1.1kpc} \sim 70$ M$_\odot$pc^{-2} since Kuijken & Gilmore (1991).

at the Sun, as proposed by the prior in Equation (4.67). The rotation curve is slightly declining with radius, consistent to what Bovy & Rix (2013) found. We also show the decomposition of the rotation curve into its contributions of the disk, halo, and (fixed) bulge. Our model has a very massive disk with rotational support at $2.2 \times R_{\rm s,disk} \sim 6.6$ kpc (where the disk rotation curve peaks) of $(v_{\rm circ,disk}/v_{\rm circ,tot})^2|_{R=2.2R_{\rm s,disk}} = 72\% \pm 2\%$ or, as fraction of the circular velocity,

$$(v_{\rm circ,disk}/v_{\rm circ,tot})|_{R=2.2R_{\rm s,disk}} = 85\% \pm 1\%. \tag{4.78}$$

The definition for a maximum disk following Sackett (1997) is $(v_{\rm circ,disk}/v_{\rm circ,tot})|_{R=2.2R_{\rm s,disk}} = 85\% \pm 10\%$. Our measurements confirm therefore that the MW has a maximum disk (Sackett 1997; Bovy & Rix 2013), but has a slightly higher rotational support than some estimates in the literature (Gerhard 1999; Piffl et al. 2014).

Surface density profile. The second panel in Figure 4.15 shows our best-fit surface density profile at $|z| = 1.1$ kpc and compares it to estimates from literature: the current best estimate by Bland-Hawthorn & Gerhard (2016) and the profile which Bovy & Rix (2013) measured (see Figure 1.20). It is obvious that our total surface mass profile is significantly more massive than the estimates from literature: ~ 98 $M_\odot$ pc^{-2} at the Sun inside 1.1 kpc, versus 70 $M_\odot$ pc^{-2} in the literature (Bland-Hawthorn & Gerhard 2016; see also Section 4.7.3). What is the reason for this discrepancy? We will discuss this in more detail in Section 4.7.1.

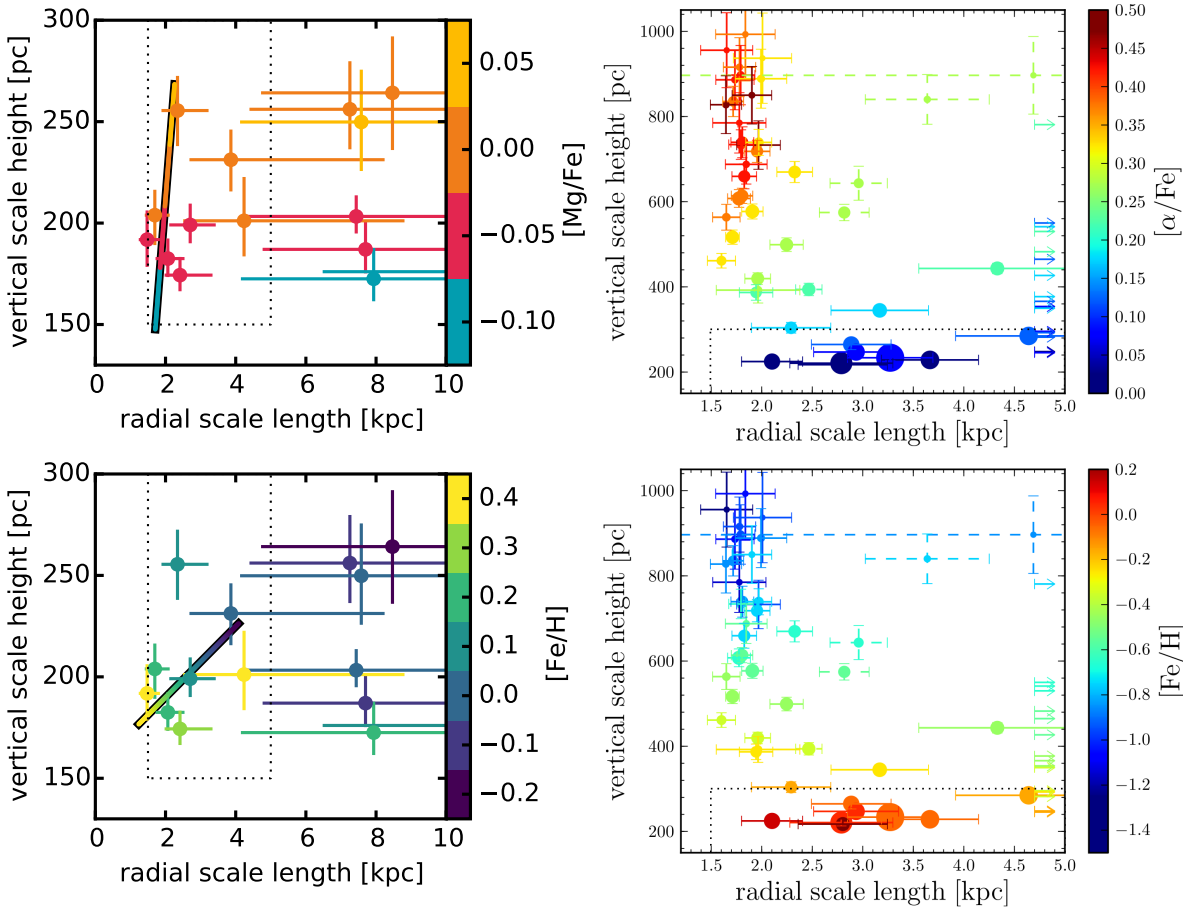
Local DM density. The best-fit NFW halo set by the fixed scale length $a_{\rm halo} = 18$ kpc, and the best-fit values in Table 4.2 for the total galaxy mass scaling $v_{\rm circ}(R_\odot) = 231.4 \pm 0.7$ km s^{-1} and the halo fraction at the Sun $f_{\rm halo}(R_\odot) = 0.22 \pm 0.02$ imply a DM density at the Sun's position of $\rho_{\rm DM,\odot} \equiv \rho_{\rm DM}(R_\odot, z_\odot) = 0.0049 \pm 0.0005$ $M_\odot$ pc^{-3}. This is consistent with recent measurements by Zhang et al. (2013a), $\rho_{\rm DM,\odot} = 0.0065 \pm 0.0023$ $M_\odot$ pc^{-3}, by Bovy & Tremaine (2012), $\rho_{\rm DM,\odot} = 0.008 \pm 0.003$ $M_\odot$ pc^{-3}, and only slightly off from Bovy & Rix (2013), $\rho_{\rm DM,\odot} = 0.008 \pm 0.0025$ $M_\odot$ pc^{-3}.

4.6.3 Findings about the Best-fit Stellar Distribution of MAPs

Our estimates for the phase-space structure of the MAPs in Figure 4.10 was derived by fitting simple exponential profiles to the data (accounting for the SF). As we only used `scipy.optimize.minimize` to estimate the best fit, we could not estimate any uncertainties. The RoadMapping fits in the previous section, however, fitted a physical DF model together with a gravitational potential in a full Bayesian framework with ten free parameters, allowing for flexibility in the model. The best-fit qDF parameters in Figure 4.12 cannot be easily interpreted in physical terms. However, together with the best-fit potential, they imply a real best-fit physical density and velocity distribution (shown in Figure 4.13). We use the convenience functions of the `galpy.quasiisothermaldf` class by Bovy (2015) to estimate for the full PDF the local exponential tracer scale length h_R, scale height h_z, the velocity dispersion σ_R and σ_z, and the exponential velocity dispersion scale lengths $h_{\sigma,R}$ and $h_{\sigma,z}$, at the Solar radius at a height of $z = 0.6$ kpc where most of the stars are located (see Figure 4.13).[28]

Tracer scale length and height. Figure 4.16 shows the correlation between the recovered tracer scale lengths and heights, and compares it with previous findings from Bovy et al. (2012d), who fitted single exponential profiles to SEGUE G-dwarfs. Bovy et al. (2016b) found that this is not a good enough model for the low-α low-metallicity MAPs, but we still compare our results to Bovy et al. (2012d), because their procedure is more comparable to ours. Note, that there is an offset between the APOGEE [Mg/Fe] scale, which we use, and the SEGUE [α/Fe] scale that they use. But the lowest-α MAPs in both data sets should be comparable to each other. We mark in Figure 4.16 the region where our and their data overlap.
There is a good qualitative agreement of the G-dwarf and RC measurements in this region, with our scale heights and lengths around $h_R \sim 3$ kpc maybe being a bit smaller, but otherwise having similar trends with abundance. Both data sets also contain at low-α and slightly sub-Solar metallicity a few MAPs that prefer very long scale lengths, which was in the study by Bovy

[28] The tracer scale length h_R was estimated at $R_\odot$ of the total tracer surface density.



(a) h_R and h_z from TGAS/RAVE RC stars.

(b) h_R and h_z from SEGUE G-dwarfs. **Figure credit:** Bovy et al. (2012d).

Figure 4.16: Correlation between the physical tracer scale length h_R and scale height h_z of different MAPs. The left panels, Figure 4.16(a), show for each of the RC MAPs in this work the radial scale length and vertical scale height with median and the 68% percentile range implied by the full PDF of the qDF and potential as derived from the given MAP. In particular, we derived the scale length from the slope of the total tracer surface density at $R_\odot$ and the scale height from the vertical slope of the density at $(R, z) = (R_\odot, 0.6 \text{ kpc})$. The upper panel color-codes the measurements by [Mg/Fe], the lower panel by [Fe/H] abundance. Overplotted, as black line with color-gradient, is a linear fit of h_R and h_z to either [Fe/H] or [Mg/Fe]. This line serves as guide to the eye to make the trend with abundance better visible. The right panels show the equivalent quantities for SEGUE G-dwarfs taken from Bovy et al. (2012d), measured directly from the data. The black dotted box in all panels mark the same region $h_z \in [150, 300]$ pc and $h_R \in [1.5, 5]$ kpc. Overall, there is a good qualitative agreement in this region. See the text in Sections 4.6.3 and 4.7.4 for more discussions.

et al. (2016b) attributed to those MAPs having a break-radius in surface density in the Solar neighborhood.

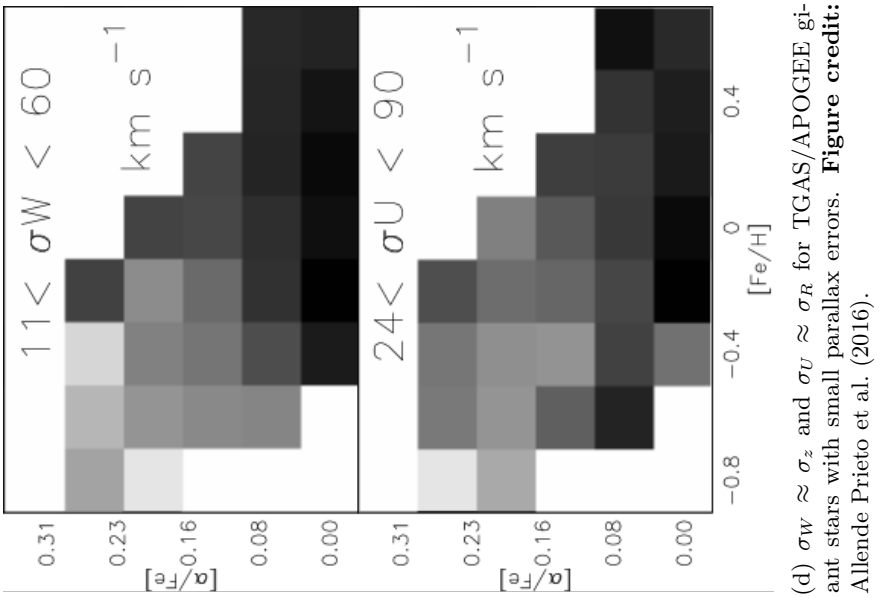
Velocity dispersions. Figure 4.17 compares the radial and vertical velocity distribution estimated for the best-fit distribution at $(R, z) = (8, 0.6)$ kpc. The Figures 4.17(b) and 4.17(c) taken from Bovy et al. (2012c) and Bovy & Rix (2013), and Figure 4.17(d) from Allende Prieto et al. (2016) are not optimal for comparison, but give a feeling for the expected range of dispersions with different abundances.

In Figure 4.17(a), there is the clear and expected trend that the velocity dispersion increases with increasing [Mg/Fe] abundance. We see again that there is, for this sample of low-α MAPs,

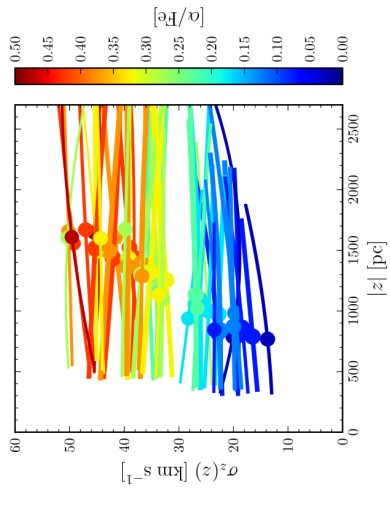
a weak but clear trend with metallicity that suggests an increase of $\sigma_R(R_\odot)$ and the velocity anisotropy σ_R/σ_z with increasing [Fe/H] (cf. also Figure 4.10).
We compare our results with previous results from the literature as derived from SEGUE G-dwarfs.
Panel 4.17(b) shows the vertical velocity dispersion profiles by Bovy et al. (2012c) at the Solar radius $R_\odot$. The low-α MAPs (blue) span a range of approximately $\sigma_z(R_\odot, 0.6 \text{ kpc}) \in [14, 24]$ km s^{-1}, which is very similar to our range of $\sigma_z(R_\odot, 0.6 \text{ kpc}) \in [18, 26]$ km s^{-1} that we also found in Figure 4.10.[29] That our vertical dispersions are by ~ 2 km s^{-1} larger, could be already an indicator why our surface density in Figure 4.15 is quite large (see Section 4.7.1). This discrepancy cannot be due to the measurement uncertainties: $\sigma_z(R_\odot)$ is dominated by the $v_{\rm los}$ measurements by RAVE with uncertainties of $\delta v_{\rm los} \lesssim 2$ km s^{-1} (see Table 1.2). Even if we we had not deconvolved the model with the measurement uncertainties, their effect on $\sigma_{z,\rm obs} = \sqrt{\sigma_z^2 + \delta v_{\rm los}^2}$ will be much smaller than this 2 km s^{-1} discrepancy.
Comparing our radial dispersion for MAPs with results from the literature is more difficult, because the radial motions in the Solar neighborhood are dominated by proper motions, which only now with TGAS became very precise for a large number of stars. In Figure 4.17(c) we show the estimates for $\sigma_{R,0}^{\rm qDF}$ from Bovy & Rix (2013) for different MAPs. The dispersions were derived from SEGUE G-dwarfs, smoothed over the abundance plane and then transformed to qDF parameters. These are close but not equal to the physical dispersion; a factor of approximately 1.1 will have to be applied to $\sigma_R^{\rm qDF}$ to get closer to the physical value (Bovy & Rix 2013, their Figure 5). Their range of approximately $\sigma_R^{\rm qDF} \in [35, 42]$ km s^{-1} $\longrightarrow \sigma_{R,0} \in [38, 46]$ km s^{-1} for the low-α MAPs agrees very well with the bulk of our $\sigma_R(R_\odot)$ measurements. We do, however, not see the trend with increasing radial velocity with [Fe/H], rather the opposite.
Figure 4.17(d) shows a recent measurement by (Allende Prieto et al. 2016) for the $\sigma_U \approx \sigma_R$ and $\sigma_W \approx \sigma_z$ velocity dispersions of the full TGAS/APOGEE giant sample depending on APOGEE [α/Fe] and [Fe/H]. Even though this is not easy to compare with our measurements and difficult to see with their adopted color-scheme, it might be possible that for the lowest-α, [Fe/H] > -0.2 MAPs there is also a slight increase in the radial velocity dispersion with metallicity. They did, however, not comment on it.
In general, APOGEE stars are located much closer to the plane of the disk than the RAVE stars in our sample, i.e., APOGEE contains more stars on near-circular orbits. This explains why σ_U and σ_W are lower for the low-α MAPs in Allende Prieto et al. (2016) as compared to our RAVE sample.

Velocity dispersion scale lengths. Figure 4.18 shows the local velocity dispersion scale lengths measured from the best-fit model at $(R, z) = (8, 0.6)$ kpc by assuming that the velocity dispersion decreases exponentially with radius. We overplot the scale lengths that were found and used in Bovy et al. (2012c) and Bovy & Rix (2013) as mean of all SEGUE G-dwarf MAPs. Bovy et al. (2012c) found no clear trend with abundance, see Figure 4.18(b). It is therefore very surprising that we find strong and clear trends of the locally measured velocity dispersion scale lengths with abundance. Especially so, because Figure 4.10 did not indicate any such trends and rather suggested that we would have difficulties properly determining the scale lengths at all, similar to h_R in Figure 4.16(a).
We will see later in Section 4.7.1 in Equation (4.87), that surface density and vertical velocity dispersion are related by $\Sigma_{\rm tot} \propto \sigma_z^2$. From this follows that $h_{\sigma,z} \sim 2R_{\rm s,disk}$, i.e., we expect very roughly a $h_{\sigma,z} \sim 6$ kpc.
At Solar [Fe/H] and [Mg/Fe], we do recover the previously found $h_{\sigma,z} \sim 7$ kpc (see, e.g., Bovy et al. 2012c; Sanders & Binney 2015b). Our range for $h_{\sigma,z} \in [4, 14]$ kpc in Figure 4.18(a) corresponds

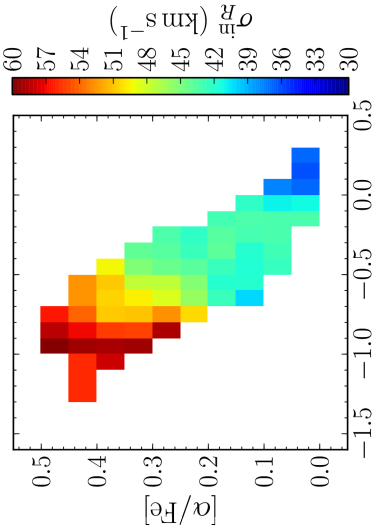
[29] It also agrees with the $\sigma_z(R_\odot, z = 0.6 \text{ kpc}) \sim 20$ km s^{-1} by Bienaymé et al. (2014), their Figure 7, for stars with metallicity [M/H]> -0.15.



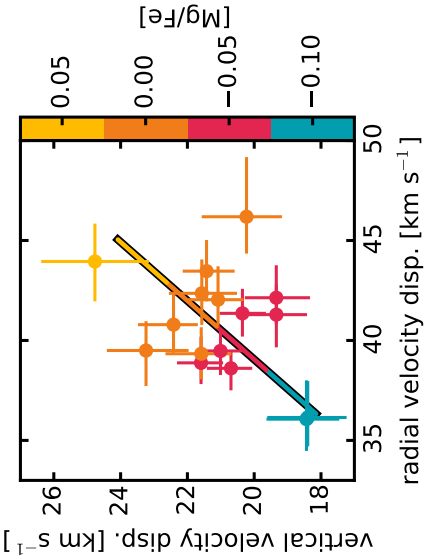
(d) $\sigma_W \approx \sigma_z$ and $\sigma_U \approx \sigma_R$ for TGAS/APOGEE giant stars with small parallax errors. **Figure credit:** Allende Prieto et al. (2016).

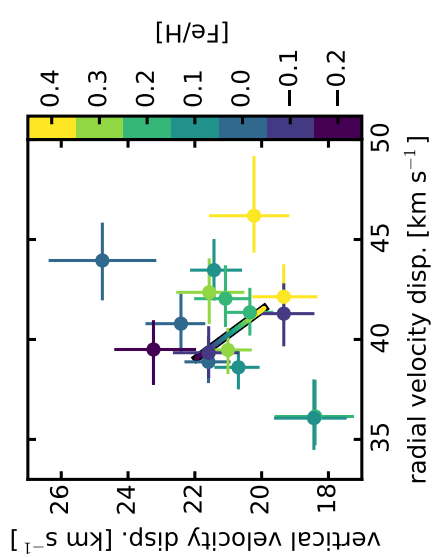


(b) $\sigma_z(R_\odot, z)$ from SEGUE G-dwarfs. **Figure credit:** Bovy et al. (2012c).



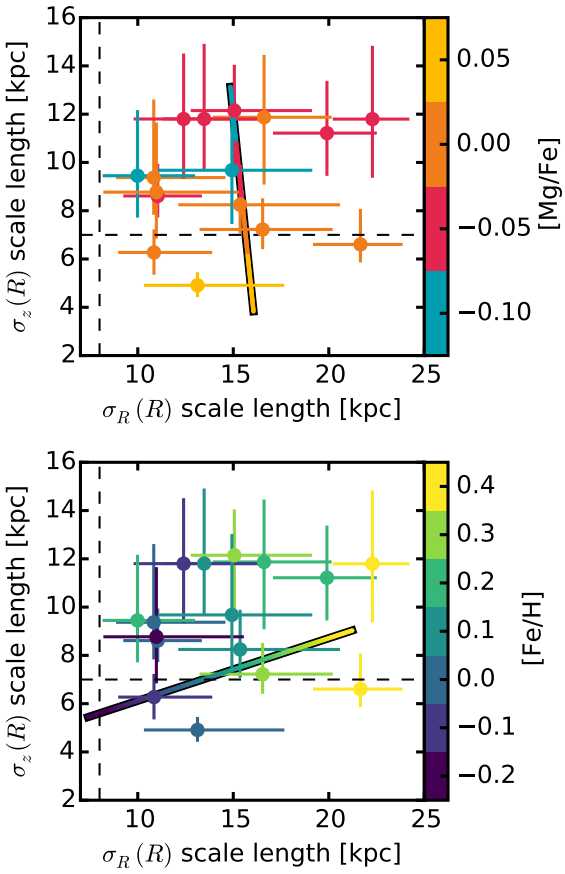
(c) $\sigma_{R,0}^{\rm qDF}$ estimated from SEGUE G-dwarfs. **Figure credit:** Bovy & Rix (2013).



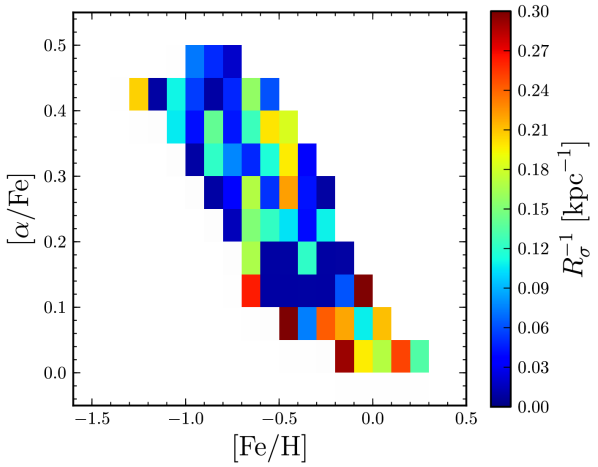


(a) σ_R and σ_z at $(R, z) = (R_\odot, 0.6$ kpc) from TGAS/RAVE RC stars.

Figure 4.17: Correlation between the physical tracer velocity dispersion in radial and vertical direction, σ_R and σ_z at $(R, z) = (R_\odot, 0.6$ kpc) of different MAPs. The left panels, Figure 4.17(a), show analogous to Figure 4.16(b) the physical velocity dispersion that is implied by the combination of the best-fit potential and qDF in Figure 4.12 depending on the chemical abundances of the respective MAP. We compare this with previous measurements from the literature in the middle and right panels. Note, that the σ_R measurements in Panel 4.17(c) by Bovy & Rix (2013) are (i) smoothed and (ii) qDF parameters that are close but not equal to the physical velocity dispersion. We discuss similarities and differences in the text in Sections 4.6.3 and 4.7.4. In Panel 4.17(d) the dispersion was measured for the (U, V, W) velocities from the whole TGAS/APOGEE giant data set (not only at the Sun), the minimum (black) and maximum (white) values of the color-scale are given in the panels.



(a) $h_{\sigma,z}$ vs. $h_{\sigma,R}$ in this work from TGAS/RAVE RC stars.



(b) $R_\sigma^{-1} \equiv h_{\sigma,z}^{-1}$ from SEGUE G-dwarfs. **Figure credit:** Bovy et al. (2012c).

Figure 4.18: Velocity dispersion scale lengths in radial and vertical direction, measured locally from the best-fit PDF of each MAP at $(R, z) = (8, 0.6)$ kpc by assuming that the velocity dispersion is decreasing exponentially with radius with a single scale length. We find in Panel 4.18(a) a surprising trend of the scale lengths with abundances. Panel 4.18(b) shows the inverse vertical velocity dispersion scale lengths measured from SEGUE G-dwarfs that show no trend with abundance at all. The range of observed $h_{\sigma,z}$ agrees, however, very well for the low-α MAPs and at Solar abundance we recover the same $h_{\sigma,z} \sim 7$ kpc. If these unexpected trends with abundance are real or systematic effects, remains to be seen.

to $h_{\sigma,z}^{-1} \in [0.07, 0.25]$ kpc^{-1} on the color-bar in Figure 4.18(b) (bright-blue to bright-red). The lowest-α MAPs in Figure 4.18(b) indeed do have $h_{\sigma,z}^{-1}$ in this range.
All our $h_{\sigma,R}$ measurements are much larger than the 8 kpc used in Bovy & Rix (2013). We note, however, that this was only a coarse estimate by Bovy & Rix (2013) based on less precise proper motions.
If our trends of the local velocity dispersion scale lengths with abundance are physically meaningful, or a systematic effect related to the specific shape of the qDF, remains an open question for now and will need to be investigated in the future.

4.7 Discussion and Summary

In the following, we will discuss several aspects of our results. In Section 4.7.1, we attempt to track down the reason for the overestimation of the surface density. In Section 4.7.2, we compare, as an additional test, the results for two methods to combine the different MAP potential results, and discuss precision and accuracy. In Section 4.7.3, we compare our Galactic potential

measurements in more detail with previous estimates. In Section 4.7.4, we set our findings on the stellar phase-space structure in a galaxy formation and evolution context. In Section 4.7.5, we will give a sneak peek on how the RoadMapping results can be useful in identifying disk substructure in the future. In Section 4.7.6, we give an overview of the caveats of the modeling and improvements that should and will be made in the future. In Section 4.7.7, we lay out how some of the problems will be resolved with *Gaia* DR2.

4.7.1 What is the Reason for the Overestimation of the Surface Mass Density?

Estimating the surface density directly from the data. The high $\Sigma_{\mathrm{tot,1.1kpc}}$ of our best-fit model is quite surprising. In the following we use heuristic arguments to show that it is indeed the data (together with the SF) that prefers such high surface densities. Let's start by assuming a vertically isothermal stellar population with vertical velocity dispersion $\sigma_z(z) = \text{const.}$, no vertical streaming motions, and a velocity ellipsoid aligned with the cylindrical coordinate frame. The vertical Jeans equation (Equation (4.222b) in Binney & Tremaine 2008)[30]

$$\frac{1}{R}\frac{\partial\left(R\rho_*\langle v_R v_z\rangle\right)}{\partial R} + \frac{\partial\left(\rho_*\langle v^2\rangle\right)}{\partial z} + \rho_*\frac{\partial\Phi}{\partial z} = 0 \tag{4.79}$$

then becomes (with $\langle v^2\rangle = \langle v\rangle^2 + \sigma_z^2 = \sigma_z^2$, $\langle v_R v_z\rangle = 0$[31], and the vertical gravitational force $F_z = -\partial\Phi/\partial z$)

$$\frac{\partial \ln\rho_*}{\partial z} \times \sigma_z^2 = F_z(z). \tag{4.80}$$

Assuming now that the population does not have a perfectly exponential density profile (because otherwise the vertical force would be independent of z) but can at least at $\bar{z}$ be approximated by an exponential $\rho_*(z) \approx \rho_{*,0}\exp\left(-|z|/h_z\right)$, we arrive at

$$F_z(\bar{z}) \approx -\frac{\sigma_z^2}{h_z}. \tag{4.81}$$

In the second step, we start with the Poisson equation in cylindrical coordinates for an axisymmetric galaxy (Eqation (3.88) in Binney & Tremaine 2008)

$$\frac{1}{R}\frac{\partial}{\partial R}\underbrace{\left(R\frac{\partial\Phi}{\partial R}\right)}_{\simeq v_{\mathrm{circ}}} + \frac{\partial^2\Phi}{\partial z^2} = 4\pi G\rho_{\mathrm{tot}}. \tag{4.82}$$

For a flat rotation curve $\partial v_{\mathrm{circ}}/\partial R = 0$, we get

$$\begin{aligned}
\frac{\partial^2\Phi}{\partial z^2} &= 4\pi G\rho_{\mathrm{tot}}(z) && (4.83)\\
\Rightarrow -\int_0^{\bar{z}} \mathrm{d}z\ \frac{\partial F_z}{\partial z} &= 4\pi G\int_0^{\bar{z}} \mathrm{d}z\ \rho_{\mathrm{tot}}(z) && (4.84)\\
\Leftrightarrow -\left[F_z(\bar{z}) - \underbrace{F_z(0)}_{=0}\right] &= 4\pi G \times \frac{1}{2}\Sigma_{\mathrm{tot}}(|z| \le \bar{z}) && (4.85)\\
\Leftrightarrow \Sigma_{\mathrm{tot}}(|z| \le \bar{z}) &= -\frac{F_z(\bar{z})}{2\pi G}. && (4.86)
\end{aligned}$$

[30] The Jeans equations are the first moments of the collisionless Boltzmann equation and relate the motions of a collisionless stellar distribution to the gravitational potential the stars move in.

[31] Even if the condition that the velocity ellipsoid is aligned with the coordinate axes is not fulfilled, the first term is still small, following §4.9.3 in (Binney & Tremaine 2008). This was also estimated by Zhang et al. (2013b) for the SEGUE G-dwarf sample.

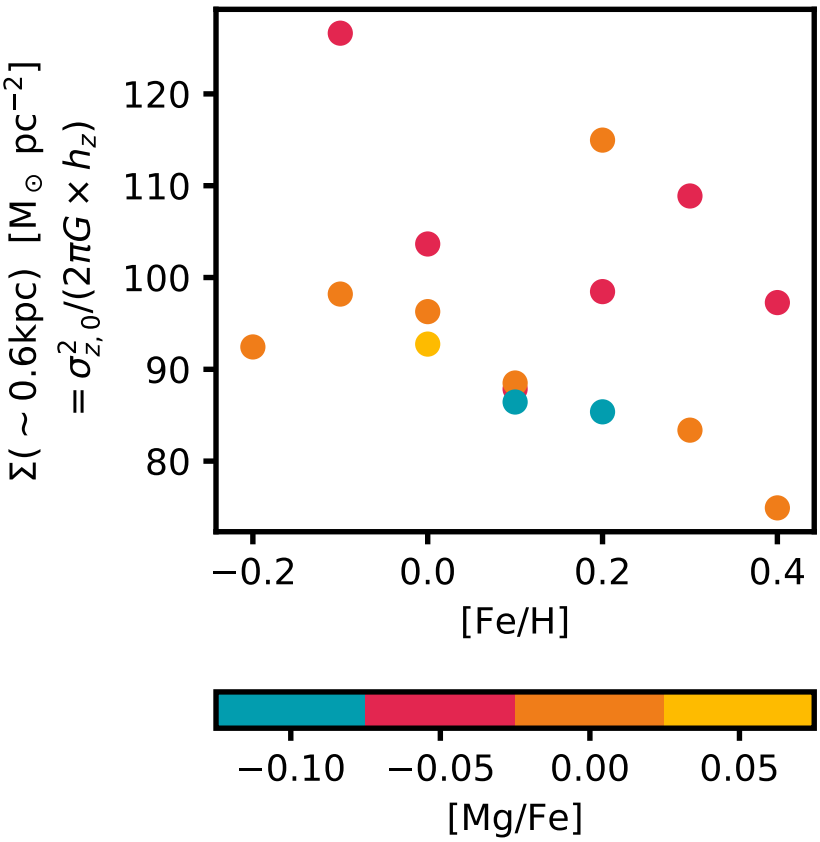


Figure 4.19: Estimates for the total surface mass density at approximately $(R, z) \sim (8, 0.6)$ kpc as estimated from combining the vertical Jeans equation with the Poisson equation and assuming an isothermal velocity dispersion in Equation (4.87). This formula was evaluated for each MAP using the quasi-isothermal vertical velocity dispersion $\sigma_{z,0}$ and scale height h_z in Figure 4.10 derived from fitting exponential profiles to the MAP data in the SF. This estimate suggests that the total surface mass density that we measure with RoadMapping in the MW is indeed driven by the data.

Taken together with Equation (4.81) this becomes

$$\Sigma_{\text{tot}}(|z| \leq \bar{z}) = \frac{\sigma_z^2}{2\pi G \times h_z}. \tag{4.87}$$

Figure 4.10 provides measurements from exponential density and velocity dispersion profile fits to the MAP data. The dispersion profile assumed that the velocity dispersion is isothermal with height and $\sigma_{z,0}$ was measured at the Sun. The median of the data sample is $\bar{z} \approx 0.6$ kpc (see Figure 4.13), and from fitting an exponential to the density around this height, we derived the scale heights h_z given in Figure 4.10. By inserting this into Equation (4.87), we get some coarse heuristic estimates for the surface density at the Sun within $\bar{z} \sim 0.6$ kpc. Figure 4.19 summarizes these estimates for all MAPs. Only one of the MAPs seems to favour a surface density of 70 $M_\odot$ pc^{-2} at $\bar{z} \sim 0.6$ kpc, most others give estimates between $80 - 100$ $M_\odot$ pc^{-2} at $\bar{z} \sim 0.6$ kpc. This is consistent with the high value for $\Sigma_{\text{tot,0.6kpc}} = 85 \pm 2$ $M_\odot$ pc^{-2} that we got from RoadMapping (see Figure 4.22).

From this simple estimate follows, that it is indeed the data (in combination with the SF) that favors the high total surface density.

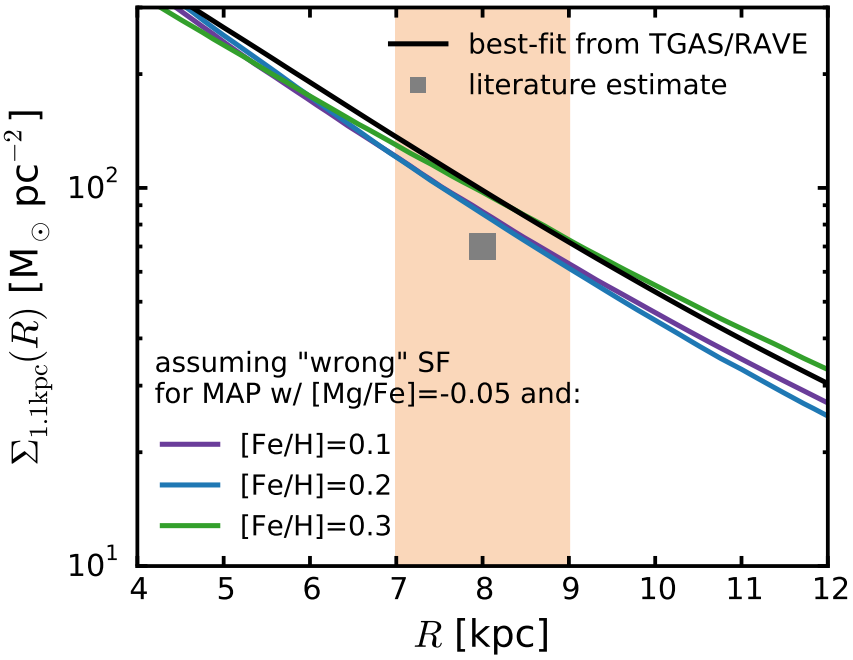


Figure 4.20: This figure tests if misjudgments in the SF could be responsible for the discrepancy of our surface density estimate at the Sun of 100 $M_\odot$ pc^{-2} and the literature estimate of 70 $M_\odot$ pc^{-2}. For this we used in the SF in Section 4.3.3, but assumed that the absolute magnitudes of RC stars in (4.36) and (4.37) were fainter by 0.3 mag. This is a strong misjudgment which is expected to decrease the surface density. This figure demonstrates that this is, at least for two of the example MAPs, indeed the case. However, even this strong misjudgment cannot fully explain the discrepancy. (We only show the median profile for each MAP to make the plot more readable. The black line is the best-fit in Table 4.2.)

Testing for a misjudgment of the SF. RoadMapping is well-tested and had no problems recovering the true surface density for the simulation in Chapter 3. From the previous paragraph we found that it appears to be the data in combination with the SF that favors the large surface densities. We therefore assume that the discrepancy is not due to the modeling machinery. Are maybe the SF or the RC distances wrong such that they bias the potential result? We chose a very simple approach to test this. We created a "wrong" SF by assuming that the RC absolute I-band and G-band magnitudes in Equations (4.36) and (4.37) were 0.3 mag fainter. This should lead to lower completeness at larger distance, cause an overestimation of the tracer density at larger distances, and therefore to measurements of higher h_z and lower surface density. 0.3 mag corresponds to 15% difference in distances—a pretty strong assumption. We are confident that we have neither misjudged the SF nor the RC distances by such a large amount. But assuming we had, by how much would the surface density reduce? Figure 4.20 shows the surface density profiles for three MAPs derived with RoadMapping when using this "wrong" SF. The recovered surface density is now lower for two of the MAPs, as expected, but it still has $\Sigma_{1.1\mathrm{kpc}}(R_\odot) \sim 85 \pm 7$ $M_\odot$. From this follows that even a hypothetical misjudgment of the SF of 15% in distances could not fully explain why we get such a large surface density. There is no other conclusion than that it is indeed the data alone that drives us towards the high surface density.

4.7.2 Precision vs. Accuracy in RoadMapping Modeling

In Section 4.6.2, we have summarized how Bovy & Rix (2013) have combined their different MAP measurements into one single surface density profile. As can be seen in Figure 1.20, some of their

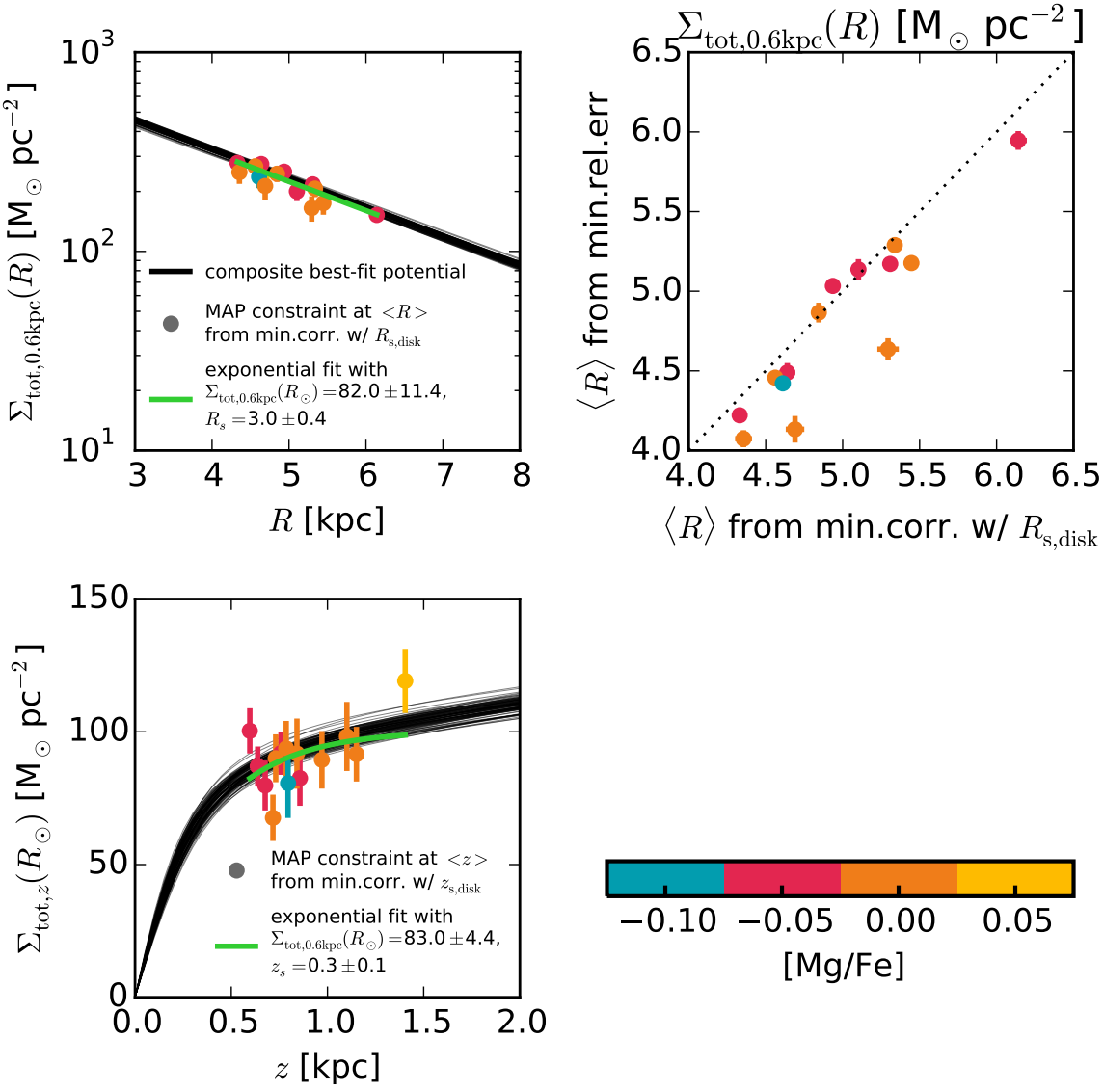


Figure 4.21: Comparison between two methods to combine the different MAP measurements of the MW surface density profile. The upper left panel shows the radial surface density profile at $z = 0.6$ kpc, the lower left panel the vertical surface density profile at $R = 8$ kpc. The black curve represents our composite best-fit model in Table 4.2 from multiplying the potential PDFs which we trust in the range $R \in [7, 8]$ kpc based on the physical and orbital location of the data. The colored dots are constraints from different MAPs, each at one single best radius or height. The best radius $\langle R\rangle$ was determined by minimizing the correlation between the surface density and the disk scale length $R_{\rm s,disk}$. The best height $\langle z\rangle$ was derived analogously, but from the correlation with $z_{\rm s,disk}$. The green line is an exponential fit to the dots and agrees well with the composite model in black; the best fit parameters are given in the respective panels. The upper right panel compares the best radius derived from the minimum correlation with the radius at which the relative uncertainty on the surface density is minimal. Both radii agree quite well with each other.

MAPs gave the best constraints at $R \in [4, 6]$ kpc, while the data was located at $R \in [6, 12]$ kpc (Bovy et al. 2012d). In Figure 2.18 in Chapter 2, and Section 3.4.2.2 in Chapter 3 we saw that the most accurate constraints on the potential were always where most of the data was located. We suspect that the explanation lies in the difference between precision and accuracy of the

potential constraint.
The upper right panel in Figure 4.21 shows on the x-axis the best radius $\langle R \rangle$ that we have determined from minimizing the correlation coefficient between the surface density profile and the disk scale length,

$$C(R) \equiv \frac{\mathrm{cov}\left[\Sigma_{\mathrm{tot,0.6kpc}}(R), R_{\mathrm{s,disk}}\right]}{\sqrt{\mathrm{var}\left[\Sigma_{\mathrm{tot,0.6kpc}}(R)\right] \cdot \mathrm{var}\left[R_{\mathrm{s,disk}}\right]}} \tag{4.88}$$

for the distribution of 6400 MCMC walker positions in the posterior PDF of the potential model. The y-axis shows the best radius $\langle R \rangle$ determined from minimizing the relative surface density error

$$S(R) \equiv \left| \frac{\mathrm{std}\left[\Sigma_{\mathrm{tot,0.6kpc}}(R)\right]}{\mathrm{mean}\left[\Sigma_{\mathrm{tot,0.6kpc}}(R)\right]} \right|. \tag{4.89}$$

To estimate the uncertainty on $\langle R \rangle$ due to the sparse sampling of the PDF, we have used the bootstrap algorithm and have determined $\langle R \rangle$ 120-times for 6400 potentials drawn with replacement from the PDF. The upper right panel in Figure 4.21 shows that the uncertainties of $\langle R \rangle$ from minimizing $C(R)$ and $S(R)$ are quite small, which means that the MCMC sampling of the PDF is sufficient. More importantly, this figure demonstrates that the radius of smallest correlation with $R_{\mathrm{s,disk}}$ is related to the radius of highest precision. We do not see any correlation with [Mg/Fe].
Bovy & Rix (2013) proceeded similarly and found that their minimum-correlation radius is close to the mean orbital radius of the population (see their Figure 12). This is not true anymore in the RoadMapping analysis of this work when comparing Figure 4.11 with Figure 4.21: The mean orbital radii for our cool thin disk populations are close to where the peak of the data is located, but $\langle R \rangle$ is a few kiloparsecs smaller and outside of the survey volume. In so far, the argumentation by Bovy & Rix (2013) ("best constraint at mean orbital radius") is not at tension with our findings ("most accurate constraint where the data is") for this data set.
If the minimum-correlation/highest-precision radius in our study is at all meaningful in the sense of Bovy & Rix (2013), remains an open question.
The upper left panel in Figure 4.21 compares the composite best fit surface density profile (black; Table 4.2; from multiplying all PDFs) to the individual MAP measurements at $\langle R \rangle$. The lower panel shows the same for the vertical surface density profile, but the best $\langle z \rangle$ was calculated from minimizing the correlation with $z_{\mathrm{s,disk}}$. Here, we see a weak trend that MAPs with higher [Mg/Fe] and higher scale heights have their minimum correlation at higher $\langle z \rangle$.
Apart from some scatter, both strategies to combine the MAP measurements give consistent results. We trust, however, only the composite result in the region $R \in [7, 8]$ kpc and $z \sim 0.6$ kpc where the data is located and sensitive.

4.7.3 Discussion of the Potential Estimates and Implications

Successes. Our disk scale length of $R_{\mathrm{s,kpc}} = 3.01 \pm 0.05$ kpc is very nicely consistent with the thin disk scale length previously measured by other studies, e.g., 3.0 ± 0.2 kpc by McMillan (2011). Bland-Hawthorn & Gerhard (2016) compiled from 15 different studies a best estimate of 2.6 ± 0.5 kpc. Also our measurement of $v_{\mathrm{circ}}(R_\odot) = 231.4 \pm 0.7$ km s^{-1} agrees within 5% with previous findings, e.g. 238 ± 15 km s^{-1} as compiled by Bland-Hawthorn & Gerhard (2016) from the literature, 239 ± 5 km s^{-1} by McMillan (2011), or 218 ± 6 km s^{-1} by Bovy et al. (2012a) from the APOGEE project. The recovered contribution of a spherical halo, the local DM density, and, related to this, the disk maximality is in agreement with previous findings as well, as we already discussed in Section 4.6.2.
All potential parameters were recovered with very high statistical precision. This indicates that now, in the era of *Gaia* where we have access highly precise measurements of many stars, statistical uncertainties do not play a major role in dynamical modeling anymore. Today, all

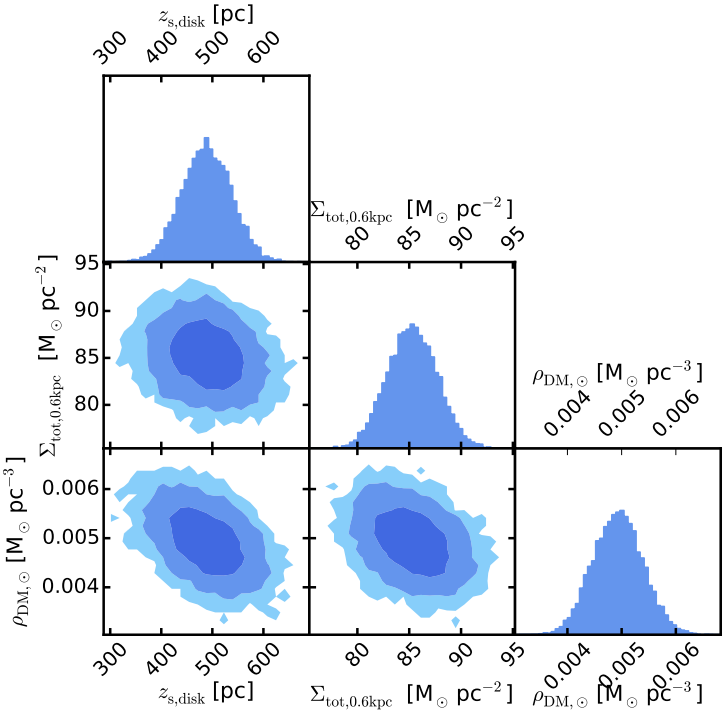


Figure 4.22: Corner plot showing the PDF for vertical disk height, $z_{\rm s,disk}$, total surface mass density at $|z| = 0.6$ kpc, $\Sigma_{\rm tot,0.6kpc}$, and the local DM density at the Solar position, $\rho_{\rm DM,\odot}$. This PDF is implied by the composite best-fit gravitational potential model in Table 4.2, derived from multiplying all MAPs PDFs. The resulting best-fit multivariate Gaussian was then MC sampled, and for each MC sample we derived the quantities shown here. The correlations between the quantities suggest, that the high $z_{\rm s,disk}$ and $\Sigma_{\rm tot,0.6kpc}$, and the low $\rho_{\rm DM,\odot}$ that we derived with respect to literature values, are related.

measurements will be dominated by systematic uncertainties. Because systematics are very hard to quantify, this is both a reason to celebrate and a challenge for this and future modeling approaches.

Discrepancies with previous findings. Our recovered disk scale height $z_{\rm s,disk} = 490 \pm 50$ pc is larger than in the literature Bland-Hawthorn & Gerhard (2016). The surface density within 1.1 kpc, $\Sigma_{\rm tot,1.1kpc} = 98 \pm 3$ M$_\odot$ pc^{-2}, is especially overestimated compared with previous findings; a large range of studies have independently and consistently recovered $\Sigma_{\rm tot,1.1kpc} \sim 70$ M$_\odot$ pc^{-2} (Kuijken & Gilmore 1991; Holmberg & Flynn 2004; Catena & Ullio 2010; Zhang et al. 2013b; Bovy & Rix 2013; Piffl et al. 2014; Bienaymé et al. 2014; see also review by Read (2014)). Siebert et al. (2003) found a slightly higher value of $\Sigma_{\rm tot,1.1kpc} = 85^{+32}_{-13}$ M$_\odot$ pc^{-2}; due to their large uncertainties it is consistent with both our measurement and the 70 M$_\odot$ pc^{-2}. At $z = 0.6$ kpc, where we expect our data to give the most reliable constraints, our $\Sigma_{\rm tot,0.6kpc} = 85 \pm 2$ M$_\odot$ pc^{-2} is still very massive.
While our local $\rho_{\rm DM,\odot}$ is consistent with previous measurements (see Section 4.6.2), it is still comparably low, which might be related to our disk being unusually massive. Figure 4.22 demonstrates that the low $\rho_{\rm DM,\odot}$ is weakly correlated with the high $z_{\rm s,disk}$, $\Sigma_{\rm tot,0.6kpc}$. It is possible that our decomposition of the overall potential constraint into dark and stellar matter is very model dependent and hinges on how good the vertical disk profile of the model reproduces the true vertical profile of the baryonic mass.

Comment on methodology. Piffl et al. (2014) and Bienaymé et al. (2014) have used orbit-based DF modeling of RAVE stars—methodologically very similar to the approach in this chapter.

However, the former has avoided the problem of including the SF by binning the stars spatially and fitting only the velocity distributions. The latter has taken into account the completeness when fitting the vertical density profile in spatial bins and even considered different MAPs in metallicity, but did only consider vertical motions.
Even though there are some discrepancies of our results with previous findings, we stress that no other dynamical modeling attempt of the MW has been so rigorous as our RoadMapping modeling, considering that we fitted 5D data simultaneously with potential and DF, fully accounting for selection effects. By applying it to TGAS data, we have also used the best data that is available so far.

Implications. If we considered our results were true, what implications would that have?
The baryonic surface density in the solar neighborhood is estimated to be $\Sigma_b \sim 50$ M$_\odot$ pc^{-2} (Kuijken & Gilmore 1989a; and more recently as sum of $\Sigma_{\rm gas} \sim 13$ M$_\odot$ pc^{-2}, $\Sigma_{*,\rm remnants} \sim 7$ M$_\odot$ pc^{-2}, and $\Sigma_{*,visible} \sim 29$ M$_\odot$ pc^{-2} (Flynn et al. 2006), the latter being in agreement with the $\Sigma_* \sim 30$ M$_\odot$ pc^{-2} determined by Bovy et al. (2012b)).
Combining our measurement of the total surface mass density at 1.1 kpc with the estimate for the baryonic column density, this allows for quite a lot of DM in the disk: ~ 50 M$_\odot$ pc^{-2}. As our model has a maximum disk and the spherical halo has a relatively low contribution to $v_{\rm circ}$ (see Figure 4.15), this would suggest that there might be space for the contribution of a DM disk to the local budget. If this was indeed the case, this would be a very important constraint on any galaxy formation theory (see Section 1.2.5).

4.7.4 Discussion of the Phase-space Structure of MAPs in a Galaxy Formation Context

Spatial distribution. We have been able to confirm previous findings about the spatial structure of our MAPs in the thin disk (see Section 1.2.3).
If we consider [Mg/Fe] as age indicator (e.g., Haywood et al. 2013), we found that younger populations are more confined to the plane of the disk, i.e., have lower scale heights h_z (see Figures 4.16, upper panels), and we have a smooth transition of scale heights between $\sim 170 - 270$ pc, in concordance with the findings of $h_z \gtrsim 200$ pc for low-α MAPs by Bovy et al. (2012d, 2016b).[32] As suggested in Bovy et al. (2012b, 2016b), this points towards the existence of a continuous internal heating mechanism in the disk of the MW.
We found no trend with [Mg/Fe] with tracer scale length, however a weak but significant trend with [Fe/H], suggesting that the more metal-rich thin disk MAPs are more centrally concentrated and closer to the disk, also in agreement with Bovy et al. (2012b, 2016b). Especially for the sub-solar metallicity MAPs, we found very long and unconstrained scale lengths. Given, that we only fitted exponential radial density profiles, that our survey volume is within $R \in [7, 9]$, and that all MAPs with [Fe/H] $\leq +0.0$ have their break-radii in $R_{\rm break} > 7$ kpc according to Bovy et al. (2016b) in Figure 1.8, this is not surprising and as expected. As suggested by Bovy et al. (2016b), this could be due to stars of a given metallicity being mainly formed at a peak-radius, where the metallicity is determined by the equilibrium of local enrichment (which is slower further out in the Galaxy) and the inflow of metal-poor gas. Subsequent radial mixing and migration could lead to the observed donut-like spatial distributions (Mackereth et al. 2017; see also Section 1.2.3).

[32] There might be an unknown offset between the SEGUE α (used in Bovy et al. 2012d; [α/Fe] ≤ 0), the APOGEE [O+Mg+Si+S+Ca]/5 (used in Bovy et al. 2016b; [([O+Mg+Si+S+Ca])/5/Fe] > -0.07), and our APOGEE/RAVE/*Cannon* Mg ([Mg/Fe] > -0.125), so we cannot fully determine, if the fact that we get slightly smaller scale heights is because we reach down to lower α and continue the trend of lower h_z with lower α, or if it is some systematic but small discrepancy that also lead to the overestimation of the surface density.

That stars with higher metallicity appear to be located more closely to the disk, could be due to the fact that the gas density in the plane of the disk is highest, therefore also the star-formation efficiency and the enrichment due to core-collapse SN Type II (see Section 1.1.4).

Velocity distribution. In the velocity distribution, we reproduced the known trend that α-younger stars have lower velocity dispersions at the Sun (see upper panel in Figure 4.17(a)), i.e., are on more circular orbits. This could be due to the fact that they have been formed more recently from the thin-layer of cool gas in the Galactic plane. The increase in the velocity dispersions for the older populations might be due to internal heating mechanisms, for example due to the spiral arms.
What was surprising, was the weak but significant trend of the radial velocity dispersion at given [Mg/Fe] increasing with rising metallicity in the lower panel of Figure 4.17(a) and also in Figure 4.10. To check that this trend is real, we have employed a similar strategy as in Figure 4.10 to the cross-match of APOGEE and TGAS stars. For the low-α MAPs we also found this trend, albeit weaker than for RAVE stars.
We attempt three speculative non-quantitative explanations. We consider a very small range of metallicities and according to Minchev et al. (2017) the MAPs in this regime are no mono-age populations. The large scatter between metallicity and age has also been measured by, e.g., Edvardsson et al. (1993); Nordström et al. (2004). The observed σ_R-vs.-[Fe/H] relation might therefore not be a trend with age, but rather related to the spatial distribution of the population:

(a) A selection effect. The radially binned velocities suggest that the dispersion is very noisy with radius. That is the reason why we could not measure h_σ in Figure 4.10, and h_σ is effectively flat over the data range. If we can believe Figure 4.18(a), this is especially true for $h_{\sigma,R}$. Then the measured σ_R is less constrained by the actual dispersion at the Sun, but by the average dispersion of the sample. More metal-rich MAPs live closer to the GC (see Figures 4.16(a) and 4.11). Closer to the GC the velocity dispersion is larger. And this is the weak σ_R-vs.-[Fe/H] trend that we see.

(b) A signature of blurring. The ISM near the Solar circle is currently at [Fe/H] $\lesssim 0$. Any super-Solar metallicity stars currently observed near the Sun are expected to have been born at smaller Galactocentric radii. If they are on circular orbits they must have been affected by radial migration induced at the co-rotating resonance (Section 1.4.5). Otherwise, they are expected to be on highly eccentric orbits and currently close to apocenter (*churning*; Kordopatis et al. 2015a; Section 1.2.3). The larger [Fe/H], the smaller the birth radii of the stars, the larger their velocity dispersion, the more stars on highly eccentric orbits can reach into the Solar neighborhood. Several studies found, however, a flat relation between [Fe/H] and eccentricity (Lee et al. 2011; Kordopatis et al. 2015a) or σ_R (Wojno et al. 2016) in the thin disk. We note, that if we averaged over [Mg/Fe] in Figure 4.17(a) most of this signature would also vanish and result in a flat trend.

(c) A signature of scattering at the Outer Lindblad resonance. The more metal-rich stars are kinematically cooler and more concentrated in the plane, which makes them more vulnerable to perturbations. They also live at smaller radii, which put them closer to the Outer Lindblad Resonance, which is expected somewhat inside of the Solar radius (Dehnen 2000). Radial migration triggered at the Outer Lindblad Resonance increases the eccentricity of the orbit (Lynden-Bell & Kalnajs 1972; see also Figure 1.17 and Section 1.4.5). In that way, resonant scattering might be responsible for the increase in radial velocity dispersion with metallicity.

All of this is pure speculation and should be also compared to N-body simulations of MW-like galaxies.

The velocity dispersion scale lengths have not been investigated much in the literature, most likely due to the lack of good proper motions. In Figure 4.18, we have compared our $h_{\sigma,z}$ with the very noisy estimates by Bovy et al. (2012c) and could at least confirm the range of recovered $h_{\sigma,z} \in [4, 14]$ kpc. We found very clear trends from our best-fit model with the velocity dispersion scale lengths, which did, however, not show up in the data directly (see Figure 4.10). This makes us suspicious that this could be a systematic effect induced by the specific assumed shape of the qDF. This will need to be investigated further in the future, and we do not attempt to find a physical explanation for it now.

4.7.5 Substructure in Action Space

In Figure 4.23, we compare—analogous to Figure 4.13—the one-dimensional action distributions for three example MAPs estimated in the best-fit potential for each MAP with the distribution implied by the best-fit qDF in the SF. While overall the fit is quite good, there are discrepancies. The strongest one is the overdensity in the data with respect to the model at $L_z \sim 1900$ kpc km s^{-1}. We suspect that this is a direct consequence of the less-than-perfect fit of the radial distribution in Figure 4.13, which we attributed to the qDF in its current form not being able to fit the flat spatial distribution of these MAPs. There are also smaller overdensities, for example in the [Mg/Fe] = -0.1 MAP (blue) at $J_R \sim 30$ kpc km s^{-1}. This could be potentially interesting, because it means there are more stars with eccentric orbits than what we would expect from an axisymmetric, smooth model distribution. Unfortunately, in the current case with only between 150-600 stars per MAP, we cannot investigate if this would be also a statistically signifiant overdensity when looking at the full 3D distribution, i.e., an excess of stars on similar orbits.

It does, however, point towards a possible application of our results in the future. As soon as we have a trustworthy potential model and reliable SF, the range of best-fit qDFs for different MAP will provide a SF-corrected model for a smooth Galactic disk in terms of orbits and abundances. This will provide us with a better handle on how to find and define substructure in the Galactic disk, and therefore strongly help with any chemical tagging attempts (cf. Hogg et al. 2016). So far, similar approaches were restricted to the Galactic stellar halo, which basically only consists of stellar sub-structure without any considerable smooth background distribution (e.g., Sanderson et al. 2015; Helmi et al. 2017; Myeong et al. 2017).

4.7.6 Caveats of the Modeling and Future Improvements

There are several caveats and restrictions in the RoadMapping modeling of the TGAS/RAVE data in this work.

Data. We have assumed that the RAVE-on stellar labels ($\log g, T_{\rm eff}$, [Fe/H], [Mg/H]) have no uncertainties. While we expect them to be small (see Section 4.2.3), they still might play a role in contaminating the sample with non-RC stars for which the RC distance estimates might not be very good, and in the contamination of the MAPs due to stars from neighboring MAPs.

We note, that if the latter contamination was unexpectedly high, we would observe an increase in surface density: Mixing two populations with different scale heights and fitting a single exponential, leads to an effective scale height that is smaller than both true scale heights (analogous to Section 2.3.5 for the scale lenghts). Following Equation (4.87), this directly translates into an overestimation of $\Sigma_{\rm tot}$.

For the de-reddening of the magnitudes and colors we use only a 2D extinction map (Schlegel et al. 1998). As we considered only stars away from the plane, we believe that this did not significantly increase the uncertainties on the distance estimates.

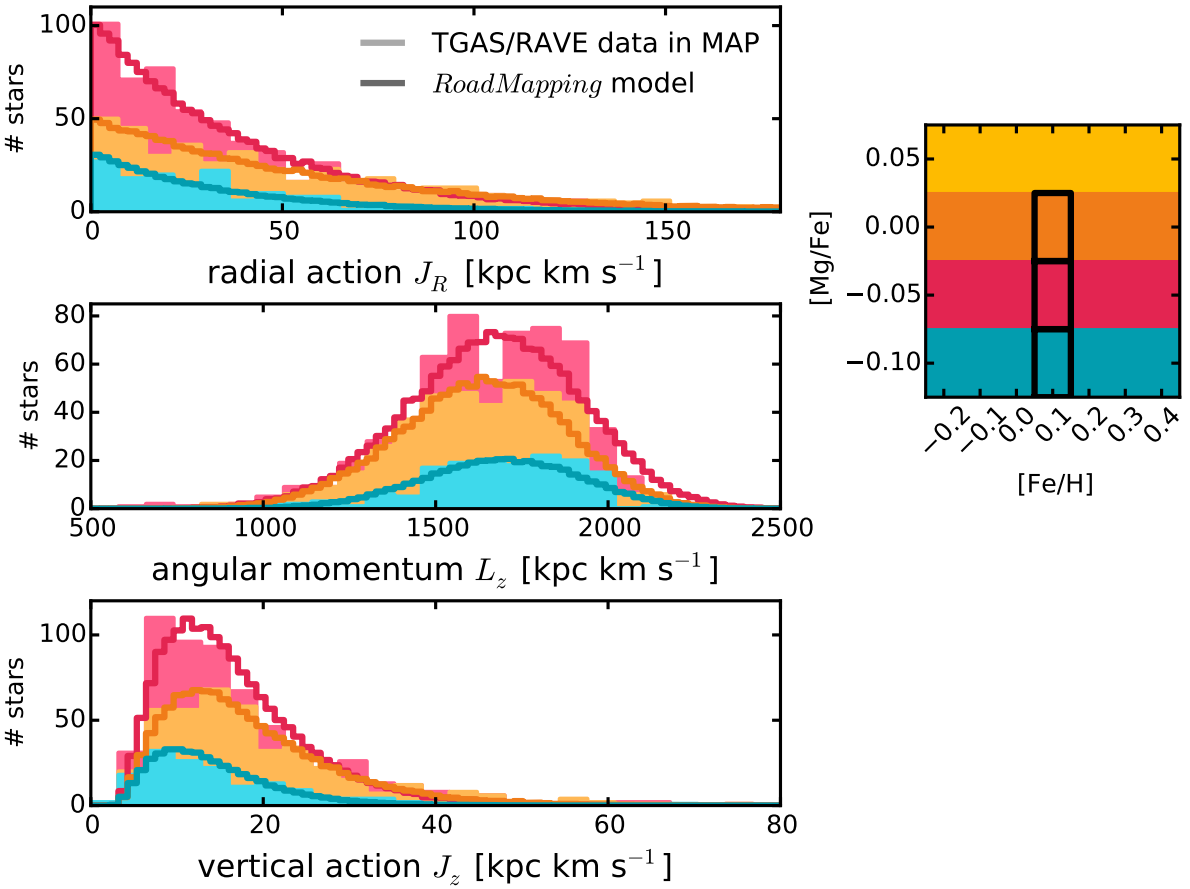


Figure 4.23: Distribution of data actions in the best-fit potentials compared to the distribution predicted by the best-fit qDF (in Figure4.12) within the SF, for the same MAPs as in Figure 4.13. This demonstrates that (i) the results of this work could be used to detect sub-structure in the disk as overdensities in the orbit-abundance space and (ii) that for the MAPs shown here the radial profile of the qDF, expressed as a function of the guiding star radius $R_g \sim L_z \times v_{\rm circ}$ and therefore L_z, declines much faster than the data, as expected from the findings by Bovy et al. (2016b).

SF. Several of our assumptions in setting up the SF for TGAS in Section 4.3.2 might be too simplified, in particular the assumption that the full *Gaia* DR1 data set (secondary catalog) is a good parent sample for TGAS (primary catalog). Figure 4.6(a) clearly shows signatures of the *Gaia* scanning law, where stars with less than 5 scans were not included in the data set. While we do not expect this to be a game changer (see also the "misjudging the SF" test in Section 4.7.1), it is a something that can and should be improved.

The SF does not use any extinction map. Because we excluded the plane of the disk of the modeling, the effect might be small, but on the other hand we remove a lot of stars at smaller latitudes that could have helped in better determining the vertical tracer profile on which also the surface density depends strongly.

We apply quality cuts to the data set (in particular quality flags from RAVE and RAVE-on) that are not modeled as part of the SF and are also not all necessary. In Section 4.2.2 we assumed that these flags would remove only random stars. It turned out that there was indeed a weak bias towards removing stars from preferentially larger distances by applying these unnecessary RAVE quality cuts, which were not accounted for in the SF and could therefore lead to misjudgments of the underlying spatial distribution. We explicitly tested for three MAPs with RoadMapping that even if we applied the RAVE quality flags in a consistent way to the data selection and SF setup, the results in this chapter are not or only very insignificantly affected.

The effect of the unmodeled RAVE-on quality constraints remains unclear but is not expected to be larger than for RAVE.

Solar position. One of the major assumptions in this work is to assume that the solar position and velocity within the Galaxy are perfectly known. This affects both the transformation of the data observables and the SF into Galactocentric cylindrical coordinates. The very precise value we found for $v_{\rm circ}(R_\odot)$ in Section 4.6.2 is expected to depend on the exact choice of $R_\odot$, here $R_\odot = 8$ kpc, and solar peculiar motion. We leave this caveat uninvestigated in this work, but will revisit it in the future.

Model. Our potential model spans with four free model parameters still a relatively small space of possible potentials. We also fix the bulge and the spherical halo's scale length to relatively arbitrary estimates, because we cannot hope to constrain them. That we only use one disk component while the MW's baryonic disk consists of three disk components, thin and thick stellar disk and gas disk, is a strong restriction. We have, however, run a few tests with a Miyamoto-Nagai disk instead of the 3MN exponential disk in Section 4.5.1 and got very similar potential constraints.
As discussed in Section 4.5.2, we have good reason to believe that the exponential profile of the qDF in its current form is not a good model for our low-α MAPs. How strongly this affects the modeling will need to be investigated in the future.

Likelihood. We have assumed a prior on the flatness of the rotation curve on the modeling (see Section 4.5.3). While this might seem restrictive, we found that by not applying this prior, we got the same potential constraints (at least for the MAPs with most of the stars). The data seems therefore to indeed favor a flat rotation curve. This caveat is therefore not a problem.
The one large approximation that RoadMapping uses is to assume that the 3D positions of the stars are perfectly known. Large distance errors could therefore introduce biases. Our SF has two sharp distance edges (see Section 4.3.3). We have not investigated how well RoadMapping behaves with significant distance errors around this additional inner edge which is needed to take into account the bright-limit of the data. If the assumption that the RC distances are very good is true, the effect is not large.

Improvements in the SF. Since the first RoadMapping modeling attempts at the beginning of 2017, there have been several advancements in the literature: Bovy (2017) has published a version of the TGAS SF that avoids regions on the sky where the *Gaia* completeness is expected to be low, and expresses the completeness as function of magnitude and also of color. The RAVE SF by Wojno et al. (2017) is now in a format that is consistent with the `HEALPIX` pixelation scheme of *Gaia*. Hawkins et al. (2017) has provided the first calibration of the *Gaia* G-band absolute magnitude of RC stars and found $M_G = +0.44 \pm 0.01$ (cf. our estimate in Equation (4.37)). We plan very soon to put all of this together and to add a 3D extinction map by, for example, Green et al. (2015) and Bovy et al. (2016a), to have an improved version of the SF for TGAS/RAVE. This will also allow to use more stars closer to the disk which increases the spatial coverage.

Improvements in the data selection. Johanna Coronado (MPIA) is currently working on a model to assign photometric distances to giant stars based on TGAS parallaxes (which we are not using in this work at all). We are planing to apply RoadMapping to her data set as soon as possible.
The selection and distance estimations of RC stars in Section 4.2.4 could also be improved by using 3D instead of 2D dust maps to de-redden the color. We would employ this together with

the distance estimates by Astraatmadja & Bailer-Jones (2016), who used a prior on the most likely 3D position of a given star within a Galaxy model to determine from the inverse TGAS parallax with its asymmetric uncertainties a realistic distance estimate.

Upshot. Overall, the data does currently not contain enough stars with enough spatial coverage to allow for a more flexible model with more free potential and qDF parameters. That the true distance uncertainties are unknown and only estimated to be 5% following Bovy et al. (2014), might introduce unknown systematics. We are, however, optimistic that some of these issues can be addressed in the next few months given the research efforts on *Gaia* data by other scientists. And last but not least, all these problems will not be issues anymore for *Gaia* DR2.

4.7.7 What will Improve with *Gaia* DR2?

From the $\sim 250,000$ stars in the TGAS/RAVE overlap, we used, in the end, only 2%. All other stars were rejected because (i) they were not RC stars and provided therefore no precise distance measurements, (ii) they had to be excluded because of the SF footprint[33], or (iii) they did not fall into a well-populated region in the abundance plane.
Fortunately, the former two aspects will not deprive us anymore of the pleasure to use the full wealth of *Gaia* with the second data release. As soon as the data is comprised of more than one year of measurements, the parallax measurements should be much better for 10^9 stars than those in TGAS. We will not have to use standard candles anymore.
The effect of the scanning law on the data completeness will be strongly reduced, because more regions on the sky will have achieved more than just 5 scans. This will simplify the SF considerably and will not force us to restrict the SF to "good" regions of the sky (Bovy 2017).
If we continue to combine RAVE with the *Gaia* data, we will not considerably increase the survey volume, because, as we saw in Section 4.3.3, it was mostly the RAVE footprint and magnitude range that restricted the volume and not the (also small) TGAS survey volume which is set by *Tycho-2*'s brightness limit. *Gaia* DR2 will contain $> 10^9$ stars down to $G \sim 20$ mag. However, if we combined *Gaia* DR2 with APOGEE, we would vastly increase the spatial coverage of the data, approximately $R \in [4, 14]$ kpc, $|z| \in [0, 3]$ kpc (Bovy et al. 2014), and also have precise radial velocities and chemical abundances thanks to APOGEE's high-resolution spectra (see Table 1.2). The SF of APOGEE is also well-understood (Bovy et al. 2014).
There even might be the possibility to combine the APOGEE/*Gaia* DR2 overlap with stars that are only in *Gaia*, because DR2 will also provide the first set of median radial velocities for stars brighter than $G_{\rm RVS} = 12$ mag (ESA 2017a). The lack of chemical abundances might make this, however, less feasible.
Even if the parallaxes of DR2 will not be precise enough for RoadMapping, there is still the option to implement the convolution of the likelihood normalisation in RoadMapping with the distance uncertainties. This will be at a very expensive price in terms of computation speed. However, it is the spatial extent of *Gaia* DR2 that will be unprecedented and important for RoadMapping (see Figure 4.24).

4.7.8 Summary

We have presented a RoadMapping analysis of RC stars in the overlap of the *Gaia*-TGAS/RAVE survey, from which we have measured the MW's gravitational potential and the phase-space and abundance structure of the thin disk in the Solar neighborhood.

[33]The TGAS footprint by Bovy (2017) rejects 52% of the sky. Even if we included more of the disk in highly extincted regions, we will only have a slightly larger number of stars per MAP than we have in our current version of the SF.

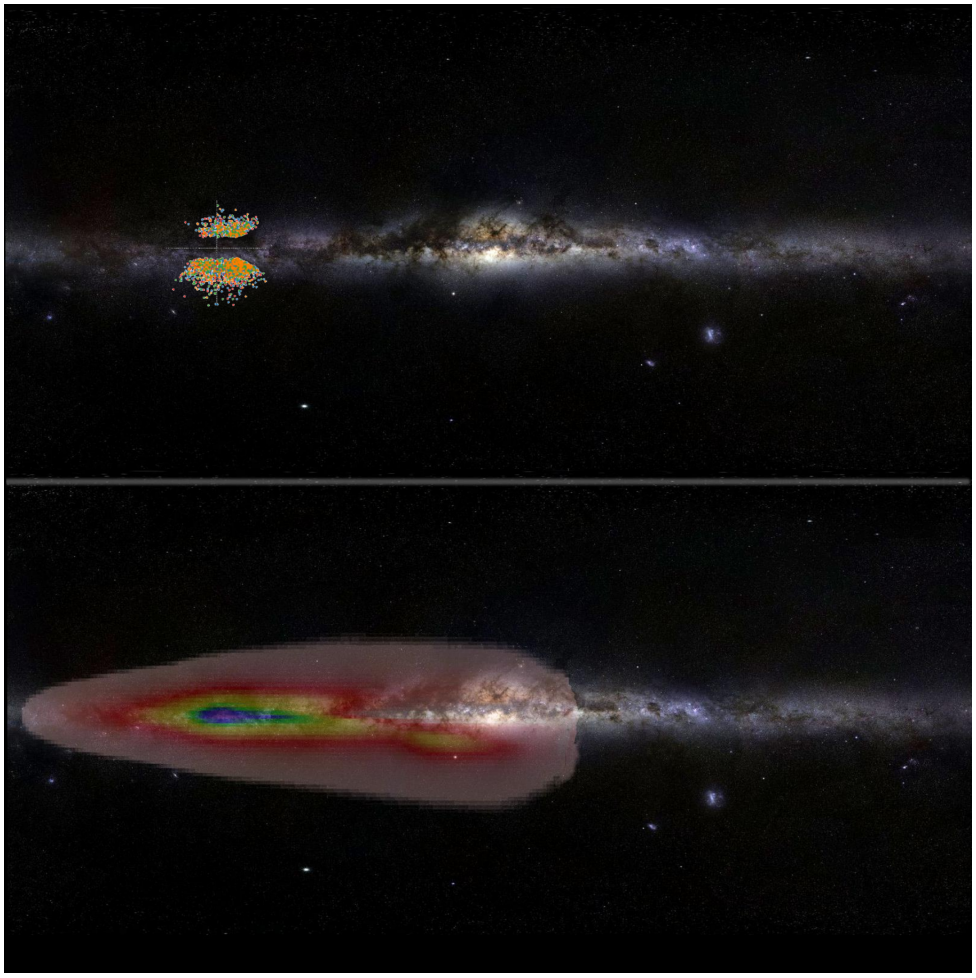


Figure 4.24: Comparison of the extent of our data set within the MW (upper panel; cf. Figure 4.9), and the expected final extent of the *Gaia* survey (lower panel). **Figure credit:** background: ESO/S. Brunier; lower panel: X. Luri & the DPAC-CU2.

RoadMapping requires 6D phase-space coordinates of stars in the MW. *Gaia*-TGAS provided 2D positions and proper motions (Lindegren et al. 2016), and RAVE line-of-sight velocities (Kunder et al. 2017). Because the parallax uncertainties of TGAS are too big for RoadMapping, we used RC standard candles. We selected RC giants on the basis of RAVE-on (Casey et al. 2017) stellar labels and assigned photometric distances following Bovy et al. (2014).
We considered 16 MAPs of ~ 5000 thin disk stars in the chemical abundance range $[\mathrm{Fe/H}] \in [-0.2, 0.4]$ and $[\mathrm{Mg/Fe}] \in [-0.1, 0.05]$ (Casey et al. 2017). We fitted simultaneously an exponential disk, NFW halo, and Hernquist bulge potential model (with four free parameters), the qDF as stellar orbit DF (with all five parameters free), and a halo outlier model separately to each MAP. We set up and took into account a SF for RC stars in both RAVE (Wojno et al. 2017) and TGAS. We constructed the TGAS SF from the primary and secondary data set of *Gaia* DR1. Apart from a weak prior on the flatness of the rotation curve from Bovy & Rix (2013), the Solar position, and the halo scale length from McMillan (2011), we have not included any previous measurements in the modeling (as opposed to many similar studies).
All MAPs gave independently potential constraints that were consistent with each other. Taken together, they provided new and highly precise constraints on the disk scale length and the circular velocity at the Sun, $R_{\mathrm{s,disk}} = 3.01 \pm 0.05$ kpc and $v_{\mathrm{circ}}(R_\odot) = 231.4 \pm 0.7$ kms s^{-1}. Our estimate for the local DM density, $\rho_{\mathrm{DM},\odot} = 0.0049 \pm 0.0005$ M$_\odot$, is consistent with previous estimates, and we also confirmed that the MW has a maximum disk. The disk scale height,

$z_{\mathrm{s,disk}} = 490 \pm 50$ pc, and especially the total surface mass density, $\Sigma_{\mathrm{tot,1.1kpc}} = 98 \pm 3$ $\mathrm{M}_\odot$ pc^{-2}, that our models recovered, were, however, larger than previous findings. We traced this issue back and found that it appears to be the data that favors these unusual high values.
We confirmed previous findings about the phase-space structure of MAPs in the thin disk, consistent with current theories of disk formation. In addition, we found a new and surprising inverse trend of the metallicity with the radial velocity dispersion.
The main caveat of our RoadMapping modeling was the assumption of radially exponential tracer density profiles. Bovy et al. (2016b) showed that the thin disk populations have, in fact, break radii within our survey volume.
Even though there are some discrepancies of our results with previous findings, we stress that no other dynamical modeling attempt of the MW has ever been so throughout and rigorous as our RoadMapping modeling. With TGAS+RAVE+RAVE-on, this RoadMapping analysis has also used the best high-dimensional stellar phase-space data ever available.
This application was a preliminary analysis using simplified assumptions for SF, potential and DF model. Given that this study was mostly a proof of concept and a teaser for what we will be able to achieve with *Gaia* DR2, our constraints for the MW's gravitational potential were already very ecouraging. With future improvements in the DF shape and SF, and especially with the *Gaia* DR2 data with its large spatial coverage and precise parallaxes, RoadMapping promises exciting new insights into the structure and formation history of the MW.

CHAPTER 5

Conclusion

We conclude this PhD thesis with an overview of possible future work and prospects for RoadMapping modeling of the MW disk in Section 5.1, and summarize the results and findings of this work in Section 5.2.

5.1 Future Prospects

Our application of RoadMapping to *Gaia* DR1-TGAS and RAVE data in Chapter 4 resulted in some encouraging and some surprising findings about the Galactic potential and stellar phase-space distribution. They will require further investigations. In the long-term, RoadMapping can help to study several different aspects in the MW beyond the obvious goal of getting better constraints on the axisymmetric Galactic potential with each new *Gaia* data release.

Further development of modeling ingredients by means of N-body simulations. Some modeling aspects of RoadMapping might be worth to go back to N-body simulations as in Chapter 3 for further investigation.

(i) The simple qDF with its exponential density profile in Section 1.5.2 was used in many studies about the Galactic disk (Binney & McMillan 2011; Ting et al. 2013; Bovy & Rix 2013; Piffl et al. 2014; Sanders & Binney 2015b; Trick et al. 2016a, 2017). Recently, Bovy et al. (2016b), Minchev et al. (2017), and Mackereth et al. (2017) have shown that in the low-α disk we might have to go beyond the simple exponential profile and explicitly include break radii and flaring in the modeling. There exists, however, no functional form for an action-based DF to capture this structure, yet. The development of such a DF could and should be accompanied by tests on data from N-body simulations (e.g., Grand et al. 2017).

(ii) In Chapter 3, we have investigated the effect of strong spiral arms on RoadMapping modeling in detail. The galaxy simulation by D'Onghia et al. (2013) did not contain a bar. While the dynamical effect in terms of resonances on the disk distribution is expected to be qualitatively similar, it might be interesting to apply RoadMapping to a simulation with a galactic bar similar to the MW bar, which has its outer Lindblad resonance close to the Solar radius.

(iii) If we plan to detect disk substructure using RoadMapping models, N-body simulations could be a pragmatic way to test how well we need to recover the potential in order to

identify clumps in orbit space (cf., e.g., McMillan & Binney 2008) and what kind of metric in orbit space and clustering algorithm is most advisable to use (cf., e.g., Hogg et al. 2016).

(iv) It would be worth to check explicitly how results from different dynamical modeling methods, in particular RoadMapping and Jeans modeling, compare to each other.

Medium-term RoadMapping modeling goals with *Gaia* DR2. The *Gaia* DR2 in April 2018 will remain state-of-the-art until the end of 2020 when DR3 is released. The key results for the potential and orbit DF with RoadMapping are therefore expected from DR2 data.
In Section 4.7.7, we have already mentioned some of the improvements that we expect with DR2. The increased spatial extent (see Figure 4.24) will also lead to a higher coverage in the space of abundances and ages as more thick disk stars will enter the catalog cross-match with, e.g., APOGEE. The MAPs that can be analyzed with RoadMapping will span a wider range considering their mean orbital radii and heights as compared to the DR1/RAVE sample in this work (see Figure 4.11). This will allow us to further exploit and investigate the possibilities to combine different potential constraints from different MAPs to get best-fit models beyond the assumed potential shape.
As the halo potential is rather difficult to constrain with stars located mostly in the stellar disk, we could include potential measurements from halo streams (e.g., Koposov et al. 2010; Law et al. 2009) as prior knowledge in RoadMapping to get the overall halo shape right.
An important new approach would be to apply RoadMapping instead of to MAPs, to mono-age populations (Bird et al. 2013; Martig et al. 2014, 2016; Minchev et al. 2014, 2017; Ness et al. 2016), or mono-age-metallicity populations (Mackereth et al. 2017). The orbit distribution for these populations will also inform research on radial migration.

Studying the disk in orbit space. One application of the RoadMapping results is to compare the smooth axisymmetric best-fit model of the disk—in the form of a DF times the spatial SF—with the data in both configuration and orbit space for MAPs and/or mono-age populations. This will allow us to detect substructure with respect to the smooth background distribution. Substructure connected to potential perturbations (like spiral arms) should show up in each sub-population as signature in orbit (see, e.g., Figure 3.8) and real space. Substructure connected to disrupted stellar groups should show up in only a few MAPs of similar abundances and would be locally confined in action space rather than in real space (see Section 4.7.5).

Mass contrast of the local spiral arm. The pitch-angle of the Orion arm is only $\sim 10^\circ - 15^\circ$ (Reid et al. 2014). Locally, a torus-shaped density ring around the GC is therefore a good enough first approximation for a spiral arm. By adding such a component to the RoadMapping potential model, we would still keep the modeling axisymmetric (even though a more accurate action estimation method might be needed), but should be able to fit and constrain, in addition to the overall exponential disk profile, the matter of the local spiral arm (see also Section 1.2.5).

Recovering non-axisymmetric structures in the potential from modeling small volumes. As we saw in Chapter 3 (Sections 3.4.2.2 and 3.4.2.4), RoadMapping made a very good attempt at constraining the local potential for volumes as small as $r_{\max} = 500$ pc or 1 kpc. It recovered, for example, the higher surface density in volumes dominated by spiral arms, or correctly estimated the circular velocity at the median radial position of the stars (see Figure 3.10).
At a time after the final *Gaia* data release, when we have data of high accuracy covering a large proportion of the disk, we could make use of this interesting property of RoadMapping modeling. We could split the data set not only into different MAPs analogous to Bovy & Rix (2013), but also into different spatial bins in the (x, y) plane of the MW and model each of the

smaller volumes separately. This approach would only probe the local potential, even when using an axisymmetric potential model, and should be sensitive to the overdensities induced by the spiral arms. In this way, it should be possible to build up a non-axisymmetric map of the MW potential—albeit with very large spatial pixels—with constraints from dynamical modeling only.

Full chemo-dynamical modeling. For reasons of galaxy and chemical evolution, the DF properties are astrophysically linked between different MAPs (e.g., Sanders & Binney 2015b; see also Sections 1.1 and 1.2). In its current implementation, RoadMapping treats all MAPs as independent and does not exploit such correlations. Ultimately, the goal is to perform a consistent chemodynamical modeling that simultaneously fits the potential and DF($\boldsymbol{J}$, [X/H]) (where [X/H] is [Fe/H] and other elements either referenced to H or Fe, i.e., [X/H] denotes the whole abundance space) with a full likelihood analysis. This has not yet been attempted with RoadMapping because the behavior is quite complex.

5.2 Summary

This PhD thesis was dedicated to the motivation, development, characterization, testing, and application of RoadMapping, a dynamical modeling technique for the MW to constrain the Galactic gravitational potential and stellar orbit distribution in the Galactic disk.
The potential informs us about the total matter distribution in the Galaxy and constrains therefore the nature of DM. The chemo-orbital DF teaches us about the assembly process of the stellar disk, the star formation history, and subsequent (secular) evolution. Both are crucial constraints on galaxy formation theory as laid out in Sections 1.1-1.3.
Axisymmetric dynamical modeling of discrete stellar tracers using action-based orbit DFs is a powerful and physically well-motivated technique to learn more about the Galaxy (see Sections 1.4-1.5). RoadMapping proceeds by simultaneously fitting an action-based DF and gravitational potential to the individual 6D phase-space coordinates of stellar MAPs in the MW disk (Section 1.5 and 2.2). It employs a full-likelihood framework, taking into account the survey SF and measurement uncertainties. RoadMapping builds on previous work by Binney & McMillan (2011), Binney (2012a), and Bovy (2015), and was first applied by Bovy & Rix (2013) to SEGUE G-dwarfs.

RoadMapping modeling and the breakdown of its assumptions. In Chapter 2, which is based on our work in Trick et al. (2016a), we have developed and improved RoadMapping to be able to flexibly incorporate a wide range of different potential, DF, and SF models, as well as data with different characteristics. The code was optimized for large data sets and for robust and fast finding of the best-fit using a combination of nested-grid search and MCMC. Chapter 2 used a large suite of mock data sets for differential test cases not only to investigate RoadMapping's modeling characteristics, but also to specifically test how robust the potential and DF constraints of RoadMapping are under the break-down of several modeling assumptions. Our key results were:

(i) If the MW's true potential is not included in the assumed model potential family, we can—in the axisymmetric case—still find a robust estimate for the potential, with only $\lesssim 10\%$ difference in surface density within $|z| \leq 1.1$ kpc inside the observed volume.

(ii) Modest systematic differences between the true and model DF are inconsequential. For instance, when binning stars to define sub-populations with simple DFs, binning errors do not affect the modeling as long as the DF parameters of neighboring bins differ by $< 20\%$.

In addition, RoadMapping ensures unbiased potential estimates for either

(iii) small misjudgments of the spatial SF ($\lesssim 15\%$ at the survey volume's edge),

(iv) if distances are known to within 10% (at a distance of 3 kpc), or

(v) if proper motion uncertainties are known within 10% or are smaller than $\delta\mu \lesssim 1$ mas yr^{-1}.

Challenges are the rapidly increasing computational costs for large sample sizes.

The influence of spiral arms in RoadMapping modeling. The fundamental assumption of RoadMapping is that the Galaxy is axisymmetric. This is required for the potential model to allow fast action calculations, and to use a simple DF that is a function of the three actions only. In reality, the MW is a spiral galaxy and far from axisymmetric. In order to explicitly test for this caveat and to challenge RoadMapping with a realistic use case, we consider an N-body simulation of a disk-dominated galaxy without bar but strong spiral arms from D'Onghia et al. (2013). We draw stellar positions and velocities from survey volumes of different sizes and positions with respect to the spiral arms and apply RoadMapping to these data sets. This investigation, which we have also published as Trick et al. (2017), lead to the following main conclusions:

(i) The potential constraints are very robust, even though we use (a) the most simple action-based DF, the qDF by Binney & McMillan (2011), and (b) a potential model family with Miyamoto-Nagai disk of which we know that it deviates slightly from the true exponential disk profile.

(ii) The best-fit RoadMapping model always recovers the correct gravitational forces where most of the stars that entered the analysis are located, even for small survey volumes.

From exploring different survey volumes, our conclusions are:

(iii) For data from large survey volumes, RoadMapping finds axisymmetric models that average well over the spiral arms. Unsurprisingly, the models are slightly biased by the excess of stars in the spiral arms.

(iv) Gravitational potential models derived from survey volumes with at least $r_{\max} = 3$ kpc around the Sun can be reliably extrapolated to larger volumes. However, a large radial survey extent, $r_{\max} \sim 5$ kpc, is needed to correctly recover the halo scale length.

(v) In general, the recovery and extrapolability of potentials inferred from data sets that were drawn from inter-arm regions appear to be better than those of data sets drawn from spiral arms.

Overall, the findings from Chapters 2 and 3 imply that building axisymmetric models for the Galaxy with *Gaia* data will lead to sensible and robust approximations of the MW's potential.

RoadMapping modeling of the MW with stars from *Gaia* and RAVE. Chapter 4, which has not been published yet, presents the first RoadMapping application to real MW data that uses the full 6D information of the stellar phase-space coordinates, flexible potential and DF models, a robust fitting process, and—most of all—an unprecedented data set combining sky positions and proper motions from *Gaia*-TGAS (Lindegren et al. 2016) with RAVE line-of-sight-velocities (Kunder et al. 2017) and chemical abundances from RAVE-on (Casey et al. 2017). To ensure no more than 10% distance errors, we selected RC standard candles following Bovy et al. (2014). We constructed a purely spatial SF for this sample by making use of the constant absolute magnitude of RC stars. The data covers approximately 1 kpc around the Sun. Approximately 5000 low-α stars in 16 MAPs in the [Fe/H]-vs.-[Mg/Fe] plane were analyzed with RoadMapping. Our key results about the gravitational potential are:

(i) All MAPs give consistent results on the parameters of the MW's gravitational potential.

(ii) By combining the individual results from all MAPs, we constrain the disk scale length, the circular velocity and the DM density at the Sun to high precision and within the errors of previous results: $R_{\rm s,disk} = 3.01 \pm 0.05$ kpc, $v_{\rm circ}(R_\odot) = 231.4 \pm 0.7$ km s^{-1}, and $\rho_{\rm DM,\odot} = 0.0049 \pm 0.0005$ M$_\odot$ pc^{-3}.

(iv) We confirm that the MW has a maximal disk (Sackett 1997), with $(v_{\rm circ,disk}/v_{\rm circ,tot})|_{R=2.2R_{\rm s,disk}} = 85\% \pm 1\%$.

(v) The surface density and the disk scale height are overestimated compared to previous results: $\Sigma_{\rm tot,1.1kpc}(R_\odot) = 98 \pm 3$ M$_\odot$pc^{-2} and $z_{\rm s,disk} = 490 \pm 50$ pc in our modeling as opposed to $\Sigma_{\rm tot,1.1kpc}(R_\odot) = 70 \pm 5$ M$_\odot$pc^{-2} and $z_{\rm s,disk} = 300^{+150}_{-80}$ pc in the literature (see, e.g., Bland-Hawthorn & Gerhard 2016). We have traced this overestimation back to be caused by the data.

We also investigated the DF of MAPs.

(vi) We confirmed previous findings on the spatial structure of low-α MAPs: Their scale heights decrease with decreasing [Mg/Fe] and increasing [Fe/H]. Low-metallicity MAPs have almost flat radial distributions within the survey volume. The density of high-metallicity MAPs decreases locally with a radial scale length of ~ 2 kpc. This picture is consistent with inside-out disk formation and gradual heating processes.

(vii) The vertical velocity dispersion at the Sun, $\sigma_{z,0}$, decreases with decreasing [Mg/Fe] and increasing [Fe/H], as expected. We found, however, that at given [Mg/Fe] the radial velocity dispersion, $\sigma_{R,0}$, increases with increasing [Fe/H]. This might be either a selection effect or a physical signature of radial mixing.

Our preliminary application of RoadMapping to *Gaia* DR1 data was overall successful and indicates that the spatial coverage of subsequent *Gaia* DRs will allow unprecedented insight into the MW's matter and stellar distribution and therefore provide important constraints on galaxy formation.

Thank you

Hans-Walter Rix, for giving me the opportunity to work on this exciting research project with you. Thank you for your scientific insight and support when I was stuck, for giving me freedom when things were going well, and for always knowing what was required when.

Jo Bovy, for being the best "pen pal" advisor I could wish for. Your answers to my emails were always prompt and incredibly helpful, no matter if it was about a small coding problem, or a major scientific confusion. It was and is a pleasure to work with you.

Glenn van de Ven, for your advice during the last five years, both on science and science life.

Christina Eilers, Jo Bovy, Johanna Coronado, Martin Trick, Michael Walther, Rima Schüßler, Sven Buder, Tobias Buck, Tobias Ostermann, for proofreading of small or large portions of my PhD thesis. Thank you!

The Max-Planck-Institute for Astronomy and the "Milky Way" and "Galaxy Dynamics" research groups, for providing such a friendly and scientifically inspiring environment. I enjoyed being part of MPIA from 2012 to 2017.

Carola Jordan, for your help with all administration related things.

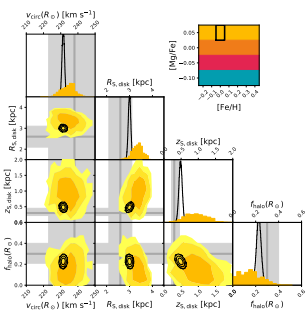
The European Research Council, for funding under the European Union's Seventh Framework Programme (FP 7) ERC Grant Agreement n. [321035].

IMPRS-HD, the International Max Planck Research School for Astronomy and Cosmic Physics at the University of Heidelberg, for support.

Rima, Christina, Michi, for a great time in Heidelberg, at MPIA, and outside of it.

My family, for your support, love, fun, and inspiration.

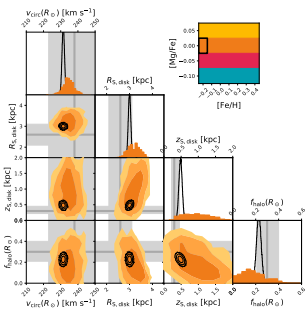
Tobi, for everything.



Abbreviations

2MASS	Two Micron All Sky Survey . . . 29
3MN	Three Miyamoto-Nagai . . . 148
ADQL	Astronomical Data Query Language . . . 139
AGB	Asymptotic Giant Branch . . . 134
AGN	Active Galactic Nucleus . . . 9
APOGEE	Sloan Digital Sky Survey III's Apache Point Observatory Galactic Evolution Experiment . . . 16
BH	Black Hole . . . 6
CCDM	Collisionless Cold Dark Matter . . . 4
CMB	Cosmic Microwave Background . . . 3
CMD	Color-Magnitude Diagram . . . 134
DF	Distribution Function . . . 25
DM	Dark Matter . . . 2
DR	Data Release . . . 29
ESA	European Space Agency . . . 127
GALAH	GALactic Archaeology with HERMES
GC	Galactic Center . . . 13
Hipparcos	High Precision Parallax Collecting Satellite . . . 28
HTPM	Hundred Thousand Proper Motions . . . 127
IMF	Initial Mass Function . . . 11
ISM	Interstellar Medium . . . 11
LAMOST	Large Sky Area Multi-Object Fibre Spectroscopic Telescope
LEGUE	LAMOST Experiment for Galactic Understanding and Exploration
MACHO	Massive Compact Halo Object . . . 6
MAP	Mono-Abundance Population . . . 17
MC	Monte Carlo . . . 57
MCMC	Monte Carlo Markov Chain . . . 63
MW	Milky Way . . . xi
NFW	Navarro, Frenk, and White . . . 149
NIR	Near-Infrared . . . 22
PARSEC	Padova and Trieste Stellar Evolution Code . . . 132

PDF Probability Distribution Function 59
qDF quasi-isothermal Distribution Function 45
RAVE RAdial Velocity Experiment 29
RC Red Clump 16
RGB Red Giant Branch 132
RoadMapping Recovery of the Orbit Action Distribution of Mono-Abundance Populations and Potential INference for our Galaxy xi
RVS Radial Velocity Spectrometer 29
SEGUE Sloan Extension for Galactic Understanding and Exploration 18
SDSS Sloan Digital Sky Survey 21
SF Selection Function 28
SMBH Supermassive Black Hole 6
SN Supernova 3
TGAS Tycho-Gaia Astrometric Solution 29
WIMP Weakly Interacting Massive Particle 5



Software

In this work, I have used the following software:

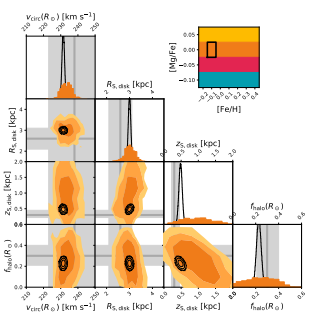
- `emcee` (Foreman-Mackey et al. 2013)
- `galpy`, version 1.1 and 1.2 (Bovy 2015)
- `Matplotlib` (Hunter 2007)
- `numpy` (`http://www.numpy.org/`)
- Read-out routines for `GADGET-3` (Springel 2005)
- `scipy` (Jones et al. 2001–2017)

Publications by Wilma Trick

Trick, W. H., Bovy, J., D'Onghia, E., and Rix, H.-W. "Action-based Dynamical Modeling for the Milky Way Disk: The Influence of Spiral Arms". *Astrophysical Journal*, **839**, 61, 2017. **Note:** Chapter 3 of this PhD thesis is based on this publication.

Trick, W. H., van de Ven, G., and Dutton, A. A. "A spiral galaxy's mass distribution uncovered through lensing and dynamics". *Monthly Notices of the Royal Astronomical Society*, **463**, 3151-3168, 2016b.

Trick, W. H., Bovy, J., and Rix, H.-W. "Action-Based Dynamical Modeling for the Milky Way Disk". *Astrophysical Journal*, **830**, 97, 2016a. **Note:** Chapter 2 of this PhD thesis is based on this publication.



Bibliography

Abadi, M. G., Navarro, J. F., Steinmetz, M., & Eke, V. R. "Simulations of Galaxy Formation in a Λ Cold Dark Matter Universe. II. The Fine Structure of Simulated Galactic Disks". *Astrophysical Journal*, **597**, 21–34, 2003.

Abazajian, K., Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., Anderson, S. F., et al. "The First Data Release of the Sloan Digital Sky Survey". *Astronomical Journal*, **126**, 2081–2086, 2003.

Aguerri, J. A. L., Balcells, M., & Peletier, R. F. "Growth of galactic bulges by mergers. I. Dense satellites". *Astronomy and Astrophysics*, **367**, 428–442, 2001.

Ahn, C. P., Alexandroff, R., Allende Prieto, C., Anders, F., Anderson, S. F., et al. "The Tenth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-III Apache Point Observatory Galactic Evolution Experiment". *Astrophysical Journal, Supplement*, **211**, 17, 2014.

Alam, S., Albareti, F. D., Allende Prieto, C., Anders, F., Anderson, S. F., et al. "The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III". *Astrophysical Journal, Supplement*, **219**, 12, 2015.

Alcock, C., Allsman, R. A., Alves, D., Axelrod, T. S., Becker, A. C., et al. "The MACHO Project Large Magellanic Cloud Microlensing Results from the First Two Years and the Nature of the Galactic Dark Halo". *Astrophysical Journal*, **486**, 697–726, 1997.

Allende Prieto, C., Kawata, D., & Cropper, M. "The rotation-metallicity relation for the Galactic disk as measured in the Gaia DR1 TGAS and APOGEE data". *Astronomy and Astrophysics*, **596**, A98, 2016.

Alpher, R. A., Bethe, H., & Gamow, G. "The Origin of Chemical Elements". *Physical Review*, **73**, 803–804, 1948.

Alpher, R. A. & Herman, R. "Evolution of the Universe". *Nature*, **162**, 774–775, 1948.

Alves, D. R. "K-Band Calibration of the Red Clump Luminosity". *Astrophysical Journal*, **539**, 732–741, 2000.

Anders, F., Chiappini, C., Santiago, B. X., Rocha-Pinto, H. J., Girardi, L., et al. "Chemodynamics of the Milky Way. I. The first year of APOGEE data". *Astronomy and Astrophysics*, **564**, A115, 2014.

Arenou, F., Luri, X., Babusiaux, C., Fabricius, C., Helmi, A., et al. "Gaia Data Release 1. Catalogue validation". *Astronomy and Astrophysics*, **599**, A50, 2017.

Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. "The Chemical Composition of the Sun". *Annual Review of Astronomy and Astrophysics*, **47**, 481–522, 2009.

Astraatmadja, T. L. & Bailer-Jones, C. A. L. "Estimating Distances from Parallaxes. III. Distances of Two Million Stars in the Gaia DR1 Catalogue". *Astrophysical Journal*, **833**, 119, 2016.

Athanassoula, E. “On the nature of bulges in general and of box/peanut bulges in particular: input from N-body simulations”. *Monthly Notices of the Royal Astronomical Society*, **358**, 1477–1488, 2005.

Balbinot, E. & Gieles, M. “The devil is in the tails: the role of globular cluster mass evolution on stream properties”. *ArXiv e-prints*, 2017.

Balcells, M. & Peletier, R. F. “Colors and color gradients in bulges of galaxies”. *Astronomical Journal*, **107**, 135–152, 1994.

Balick, B. & Brown, R. L. “Intense sub-arcsecond structure in the galactic center”. *Astrophysical Journal*, **194**, 265–270, 1974.

Bardeen, J. M., Steinhardt, P. J., & Turner, M. S. “Spontaneous creation of almost scale-free density perturbations in an inflationary universe”. *Physical Review D*, **28**, 679–693, 1983.

Barros, D. A., Lépine, J. R. D., & Dias, W. S. “Models for the 3D axisymmetric gravitational potential of the Milky Way galaxy. A detailed modelling of the Galactic disk”. *Astronomy and Astrophysics*, **593**, A108, 2016.

Batsleer, P. & Dejonghe, H. “The generation of Stackel potentials for galactic dynamical modeling”. *Astronomy and Astrophysics*, **287**, 43–54, 1994.

Behroozi, P. S., Wechsler, R. H., & Conroy, C. “The Average Star Formation Histories of Galaxies in Dark Matter Halos from z = 0-8”. *Astrophysical Journal*, **770**, 57, 2013.

Bell, E. F. & de Jong, R. S. “Stellar Mass-to-Light Ratios and the Tully-Fisher Relation”. *Astrophysical Journal*, **550**, 212–229, 2001.

Bell, E. F., Zucker, D. B., Belokurov, V., Sharma, S., Johnston, K. V., et al. “The Accretion Origin of the Milky Way’s Stellar Halo”. *Astrophysical Journal*, **680**, 295-311, 2008.

Belokurov, V., Zucker, D. B., Evans, N. W., Gilmore, G., Vidrih, S., et al. “The Field of Streams: Sagittarius and Its Siblings”. *Astrophysical Journal, Letters*, **642**, L137–L140, 2006.

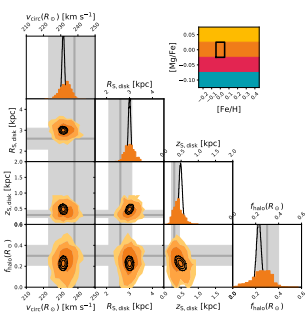
Bender, R., Burstein, D., & Faber, S. M. “Dynamically hot galaxies. I - Structural properties”. *Astrophysical Journal*, **399**, 462–477, 1992.

Bensby, T., Feltzing, S., & Oey, M. S. “Exploring the Milky Way stellar disk. A detailed elemental abundance study of 714 F and G dwarf stars in the solar neighbourhood”. *Astronomy and Astrophysics*, **562**, A71, 2014.

Bienaymé, O., Famaey, B., Siebert, A., Freeman, K. C., Gibson, B. K., et al. “Weighing the local dark matter with RAVE red clump stars”. *Astronomy and Astrophysics*, **571**, A92, 2014.

Binney, J. “Distribution functions for the Milky Way”. *Monthly Notices of the Royal Astronomical Society*, **401**, 2318–2330, 2010.

Binney, J. “Extracting science from surveys of our Galaxy”. *Pramana*, **77**, 39–52, 2011.



Binney, J. “Actions for axisymmetric potentials”. *Monthly Notices of the Royal Astronomical Society*, **426**, 1324–1327, 2012a.

Binney, J. “More dynamical models of our Galaxy”. *Monthly Notices of the Royal Astronomical Society*, **426**, 1328–1337, 2012b.

Binney, J. “Dynamics for galactic archaeology”. *NewAR*, **57**, 29–51, 2013.

Binney, J., Gerhard, O., & Spergel, D. “The photometric structure of the inner Galaxy”. *Monthly Notices of the Royal Astronomical Society*, **288**, 365–374, 1997.

Binney, J., Gerhard, O. E., Stark, A. A., Bally, J., & Uchida, K. I. "Understanding the kinematics of Galactic centre gas". *Monthly Notices of the Royal Astronomical Society*, **252**, 210–218, 1991.

Binney, J. & Lacey, C. "The diffusion of stars through phase space". *Monthly Notices of the Royal Astronomical Society*, **230**, 597–627, 1988.

Binney, J. & McMillan, P. "Models of our Galaxy - II". *Monthly Notices of the Royal Astronomical Society*, **413**, 1889–1898, 2011.

Binney, J. & McMillan, P. J. "Torus mapper: a code for dynamical models of galaxies". *Monthly Notices of the Royal Astronomical Society*, **456**, 1982–1998, 2016.

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

Binney, J. J. & Evans, N. W. "Cuspy dark matter haloes and the Galaxy". *Monthly Notices of the Royal Astronomical Society*, **327**, L27–L31, 2001.

Bird, J. C., Kazantzidis, S., & Weinberg, D. H. "Radial mixing in galactic discs: the effects of disc structure and satellite bombardment". *Monthly Notices of the Royal Astronomical Society*, **420**, 913–925, 2012.

Bird, J. C., Kazantzidis, S., Weinberg, D. H., Guedes, J., Callegari, S., et al. "Inside out and Upside down: Tracing the Assembly of a Simulated Disk Galaxy Using Mono-age Stellar Populations". *Astrophysical Journal*, **773**, 43, 2013.

Birnboim, Y. & Dekel, A. "Virial shocks in galactic haloes?" *Monthly Notices of the Royal Astronomical Society*, **345**, 349–364, 2003.

Bissantz, N., Debattista, V. P., & Gerhard, O. "Large-Scale Model of the Milky Way: Stellar Kinematics and the Microlensing Event Timescale Distribution in the Galactic Bulge". *Astrophysical Journal, Letters*, **601**, L155–L158, 2004.

Blaauw, A. "Star formation in the Orion Arm". In "The Milky Way Galaxy", , edited by van Woerden, H., Allen, R. J., & Burton, W. B., volume 106 of *IAU Symposium*, pp. 335–341, 1985.

Bland-Hawthorn, J. & Gerhard, O. "The Galaxy in Context: Structural, Kinematic, and Integrated Properties". *Annual Review of Astronomy and Astrophysics*, **54**, 529–596, 2016.

Blitz, L. & Spergel, D. N. "Direct evidence for a bar at the Galactic center". *Astrophysical Journal*, **379**, 631–638, 1991.

Bode, P., Ostriker, J. P., & Turok, N. "Halo Formation in Warm Dark Matter Models". *Astrophysical Journal*, **556**, 93–107, 2001.

Bonaca, A., Conroy, C., Wetzel, A., Hopkins, P. F., & Keres, D. "Gaia reveals a metal-rich in-situ component of the local stellar halo". *ArXiv e-prints*, 2017.

Bournaud, F., Elmegreen, B. G., & Martig, M. "The Thick Disks of Spiral Galaxies as Relics from Gas-rich, Turbulent, Clumpy Disks at High Redshift". *Astrophysical Journal, Letters*, **707**, L1–L5, 2009.

Bovy, J. "Dynamical Modeling of Tidal Streams". *Astrophysical Journal*, **795**, 95, 2014.

Bovy, J. "galpy: A python Library for Galactic Dynamics". *Astrophysical Journal, Supplement*, **216**, 29, 2015.

Bovy, J. "The Chemical Homogeneity of Open Clusters". *Astrophysical Journal*, **817**, 49, 2016.

Bovy, J. "Stellar Inventory of the Solar Neighborhood using Gaia DR1". *ArXiv e-prints*, 2017.

Bovy, J., Allende Prieto, C., Beers, T. C., Bizyaev, D., da Costa, L. N., et al. "The Milky Way's Circular-velocity Curve between 4 and 14 kpc from APOGEE data". *Astrophysical Journal*, **759**, 131, 2012a.

Bovy, J., Bird, J. C., García Pérez, A. E., Majewski, S. R., Nidever, D. L., et al. "The Power Spectrum of the Milky Way: Velocity Fluctuations in the Galactic Disk". *Astrophysical Journal*, **800**, 83, 2015.

Bovy, J. & Hogg, D. W. "The Velocity Distribution of Nearby Stars from Hipparcos Data. II. The Nature of the Low-velocity Moving Groups". *Astrophysical Journal*, **717**, 617–639, 2010.

Bovy, J., Hogg, D. W., & Roweis, S. T. "The Velocity Distribution of Nearby Stars from Hipparcos Data. I. The Significance of the Moving Groups". *Astrophysical Journal*, **700**, 1794–1819, 2009.

Bovy, J., Kawata, D., & Hunt, J. A. S. "Made-to-measure modeling of observed galaxy dynamics". *ArXiv e-prints*, 2017.

Bovy, J., Nidever, D. L., Rix, H.-W., Girardi, L., Zasowski, G., et al. "The APOGEE Red-clump Catalog: Precise Distances, Velocities, and High-resolution Elemental Abundances over a Large Area of the Milky Way's Disk". *Astrophysical Journal*, **790**, 127, 2014.

Bovy, J. & Rix, H.-W. "A Direct Dynamical Measurement of the Milky Way's Disk Surface Density Profile, Disk Scale Length, and Dark Matter Profile at 4 kpc <˜ R <˜ 9 kpc". *Astrophysical Journal*, **779**, 115, 2013.

Bovy, J., Rix, H.-W., Green, G. M., Schlafly, E. F., & Finkbeiner, D. P. "On Galactic Density Modeling in the Presence of Dust Extinction". *Astrophysical Journal*, **818**, 130, 2016a.

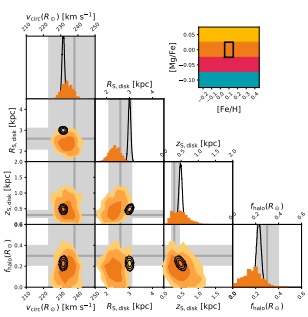
Bovy, J., Rix, H.-W., & Hogg, D. W. "The Milky Way Has No Distinct Thick Disk". *Astrophysical Journal*, **751**, 131, 2012b.

Bovy, J., Rix, H.-W., Hogg, D. W., Beers, T. C., Lee, Y. S., et al. "The Vertical Motions of Mono-abundance Sub-populations in the Milky Way Disk". *Astrophysical Journal*, **755**, 115, 2012c.

Bovy, J., Rix, H.-W., Liu, C., Hogg, D. W., Beers, T. C., et al. "The Spatial Structure of Mono-abundance Sub-populations of the Milky Way Disk". *Astrophysical Journal*, **753**, 148, 2012d.

Bovy, J., Rix, H.-W., Schlafly, E. F., Nidever, D. L., Holtzman, J. A., et al. "The Stellar Population Structure of the Galactic Disk". *Astrophysical Journal*, **823**, 30, 2016b.

Bovy, J. & Tremaine, S. "On the Local Dark Matter Density". *Astrophysical Journal*, **756**, 89, 2012.



Bressan, A., Marigo, P., Girardi, L., Salasnich, B., Dal Cero, C., et al. "PARSEC: stellar tracks and isochrones with the PAdova and TRieste Stellar Evolution Code". *Monthly Notices of the Royal Astronomical Society*, **427**, 127–145, 2012.

Brink, J., Geyer, M., & Hinderer, T. "Astrophysics of resonant orbits in the Kerr metric". *Physical Review D*, **91**(8), 083001, 2015.

Büdenbender, A., van de Ven, G., & Watkins, L. L. "The tilt of the velocity ellipsoid in the Milky Way disc". *Monthly Notices of the Royal Astronomical Society*, **452**, 956–968, 2015.

Buote, D. A. & Canizares, C. R. "X-ray constraints on the shape of the dark matter in five Abell clusters". *Astrophysical Journal*, **400**, 385–397, 1992.

Burke, B. F. "Systematic distortion of the outer regions of the galaxy." *Astronomical Journal*, **62**, 90, 1957.

Burkert, A. M. & D'Onghia, E. "Galaxy Formation and the Cosmological Angular Momentum Problem". In "Penetrating Bars Through Masks of Cosmic Dust", , edited by Block, D. L., Puerari, I., Freeman, K. C., Groess, R., & Block, E. K., volume 319 of *Astrophysics and Space Science Library*, p. 341, 2004.

Burles, S., Nollett, K. M., & Turner, M. S. "Big Bang Nucleosynthesis Predictions for Precision Cosmology". *Astrophysical Journal, Letters*, **552**, L1–L5, 2001.

Burnett, B. & Binney, J. "Stellar distances from spectroscopic observations: a new technique". *Monthly Notices of the Royal Astronomical Society*, **407**, 339–354, 2010.

Cannon, R. D. "Red giants in old open clusters." *Monthly Notices of the Royal Astronomical Society*, **150**, 111–135, 1970.

Cantat-Gaudin, T., Donati, P., Vallenari, A., Sordo, R., Bragaglia, A., et al. "Abundances and kinematics for ten anticentre open clusters". *Astronomy and Astrophysics*, **588**, A120, 2016.

Carney, B. W. "The Origins of the Halo and the Thick Disk". In "Dynamics of Star Clusters and the Milky Way", , edited by Deiters, S., Fuchs, B., Just, A., Spurzem, R., & Wielen, R., volume 228 of *Astronomical Society of the Pacific Conference Series*, p. 255, 2001.

Carr, B. J., Bond, J. R., & Arnett, W. D. "Cosmological consequences of Population III stars". *Astrophysical Journal*, **277**, 445–469, 1984.

Carroll, B. & Ostlie, D. *An Introduction to Modern Astrophysics*. Pearson Addison-Wesley, 2007.

Casagrande, L., Schönrich, R., Asplund, M., Cassisi, S., Ramírez, I., et al. "New constraints on the chemical evolution of the solar neighbourhood and Galactic disc(s). Improved astrophysical parameters for the Geneva-Copenhagen Survey". *Astronomy and Astrophysics*, **530**, A138, 2011.

Casey, A. R., Hawkins, K., Hogg, D. W., Ness, M., Rix, H.-W., et al. "The RAVE-on Catalog of Stellar Atmospheric Parameters and Chemical Abundances for Chemo-dynamic Studies in the Gaia Era". *Astrophysical Journal*, **840**, 59, 2017.

Catena, R. & Ullio, P. "A novel determination of the local dark matter density". *Journal of Cosmology and Astroparticle Physics*, **8**, 004, 2010.

Chabrier, G. "The Galactic Disk Mass Budget. I. Stellar Mass Function and Density". *Astrophysical Journal*, **554**, 1274–1281, 2001.

Chiappini, C., Matteucci, F., & Romano, D. "Abundance Gradients and the Formation of the Milky Way". *Astrophysical Journal*, **554**, 1044–1058, 2001.

Churchwell, E., Babler, B. L., Meade, M. R., Whitney, B. A., Benjamin, R., et al. "The Spitzer/GLIMPSE Surveys: A New View of the Milky Way". *Publications of the Astronomical Society of the Pacific*, **121**, 213–230, 2009.

Clemens, D. P. "Massachusetts-Stony Brook Galactic plane CO survey - The Galactic disk rotation curve". *Astrophysical Journal*, **295**, 422–428, 1985.

Clowe, D., Bradač, M., Gonzalez, A. H., Markevitch, M., Randall, S. W., et al. "A Direct Empirical Proof of the Existence of Dark Matter". *Astrophysical Journal, Letters*, **648**, L109–L113, 2006.

Cole, S., Aragon-Salamanca, A., Frenk, C. S., Navarro, J. F., & Zepf, S. E. "A Recipe for Galaxy Formation". *Monthly Notices of the Royal Astronomical Society*, **271**, 781, 1994.

Combes, F., Debbasch, F., Friedli, D., & Pfenniger, D. "Box and peanut shapes generated by stellar bars". *Astronomy and Astrophysics*, **233**, 82–95, 1990.

Conselice, C. J. "The Evolution of Galaxy Structure Over Cosmic Time". *Annual Review of Astronomy and Astrophysics*, **52**, 291–337, 2014.

Courteau, S. & Rix, H.-W. "Maximal Disks and the Tully-Fisher Relation". *Astrophysical Journal*, **513**, 561–571, 1999.

Cui, X.-Q., Zhao, Y.-H., Chu, Y.-Q., Li, G.-P., Li, Q., et al. "The Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST)". *Research in Astronomy and Astrophysics*, **12**, 1197–1242, 2012.

Dahlen, T., Strolger, L.-G., & Riess, A. G. "The Extended HST Supernova Survey: The Rate of SNe Ia at z ¿ 1.4 Remains Low". *Astrophysical Journal*, **681**, 462-469, 2008.

Dalcanton, J. J., Spergel, D. N., & Summers, F. J. "The Formation of Disk Galaxies". *Astrophysical Journal*, **482**, 659–676, 1997.

Das, P. & Binney, J. "Characterising stellar halo populations I: An extended distribution function for halo K giants". *Monthly Notices of the Royal Astronomical Society*, 2016.

Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. "The evolution of large-scale structure in a universe dominated by cold dark matter". *Astrophysical Journal*, **292**, 371–394, 1985.

de Bernardis, P., Ade, P. A. R., Bock, J. J., Bond, J. R., Borrill, J., et al. "A flat Universe from high-resolution maps of the cosmic microwave background radiation". *Nature*, **404**, 955–959, 2000.

de Bruijne, J. H. J. "Science performance of Gaia, ESA's space-astrometry mission". *Astrophysics and Space Science*, **341**, 31–41, 2012.

de Bruijne, J. H. J., Rygl, K. L. J., & Antoja, T. "Gaia Astrometric Science Performance - Post-Launch Predictions". In "EAS Publications Series", volume 67 of *EAS Publications Series*, pp. 23–29, 2014.

de Grijs, R., Kregel, M., & Wesson, K. H. "Radially truncated galactic discs". *Monthly Notices of the Royal Astronomical Society*, **324**, 1074–1086, 2001.

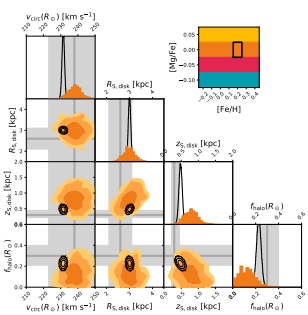
de Jong, R. S. "Near-infrared and optical broadband surface photometry of 86 face-on disk dominated galaxies. III. The statistics of the disk and bulge parameters." *Astronomy and Astrophysics*, **313**, 45–64, 1996a.

de Jong, R. S. "Near-infrared and optical broadband surface photometry of 86 face-on disk dominated galaxies. IV. Using color profiles to study stellar and dust content of galaxies." *Astronomy and Astrophysics*, **313**, 377–395, 1996b.

de la Vega, A., Quillen, A. C., Carlin, J. L., Chakrabarti, S., & D'Onghia, E. "Phase wrapping of epicyclic perturbations in the Wobbly Galaxy". *Monthly Notices of the Royal Astronomical Society*, **454**, 933–945, 2015.

de Lorenzi, F., Debattista, V. P., Gerhard, O., & Sambhus, N. "NMAGIC: a fast parallel implementation of a χ^2-made-to-measure algorithm for modelling observational data". *Monthly Notices of the Royal Astronomical Society*, **376**, 71–88, 2007.

De Silva, G. M., Freeman, K. C., Bland-Hawthorn, J., Martell, S., de Boer, E. W., et al. "The GALAH survey: scientific motivation". *Monthly Notices of the Royal Astronomical Society*, **449**, 2604–2617, 2015.



de Zeeuw, T. “Elliptical galaxies with separable potentials”. *Monthly Notices of the Royal Astronomical Society*, **216**, 273–334, 1985.

Dehnen, W. “The Distribution of Nearby Stars in Velocity Space Inferred from HIPPARCOS Data”. *Astronomical Journal*, **115**, 2384–2396, 1998.

Dehnen, W. “The Effect of the Outer Lindblad Resonance of the Galactic Bar on the Local Stellar Velocity Distribution”. *Astronomical Journal*, **119**, 800–812, 2000.

Dehnen, W. & Binney, J. “Mass models of the Milky Way”. *Monthly Notices of the Royal Astronomical Society*, **294**, 429, 1998.

Dejonghe, H. & de Zeeuw, T. “A simple dynamical model for stars in the Galactic halo”. *Astrophysical Journal*, **329**, 720–728, 1988.

Dekel, A. & Birnboim, Y. “Galaxy bimodality due to cold flows and shock heating”. *Monthly Notices of the Royal Astronomical Society*, **368**, 2–20, 2006.

Delaunay, M. “Note sur les Iné Lunaires às Pédues à l’action perturbatrice de Vé nus”. *Monthly Notices of the Royal Astronomical Society*, **21**, 60, 1860.

DENIS Consortium. “VizieR Online Data Catalog: The DENIS database (DENIS Consortium, 2005)”. *VizieR Online Data Catalog*, **2263**, 2005.

Dicke, R. H., Peebles, P. J. E., Roll, P. G., & Wilkinson, D. T. “Cosmic Black-Body Radiation.” *Astrophysical Journal*, **142**, 414–419, 1965.

D’Onghia, E. “Disk-stability Constraints on the Number of Arms in Spiral Galaxies”. *Astrophysical Journal, Letters*, **808**, L8, 2015.

D’Onghia, E., Vogelsberger, M., & Hernquist, L. “Self-perpetuating Spiral Arms in Disk Galaxies”. *Astrophysical Journal*, **766**, 34, 2013.

Doroshkevich, A. G. “Spatial structure of perturbations and origin of galactic rotation in fluctuation theory”. *Astrophysics*, **6**, 320–330, 1970.

Dotter, A., Conroy, C., Cargile, P., & Asplund, M. “The Influence of Atomic Diffusion on Stellar Ages and Chemical Tagging”. *Astrophysical Journal*, **840**, 99, 2017.

Drimmel, R. “Evidence for a two-armed spiral in the Milky Way”. *Astronomy and Astrophysics*, **358**, L13–L16, 2000.

Drimmel, R. & Spergel, D. N. “Three-dimensional Structure of the Milky Way Disk: The Distribution of Stars and Dust beyond 0.35 R_{solar}”. *Astrophysical Journal*, **556**, 181–202, 2001.

Dubinski, J. & Carlberg, R. G. “The structure of cold dark matter halos”. *Astrophysical Journal*, **378**, 496–503, 1991.

Dutton, A. A. “On the origin of exponential galaxy discs”. *Monthly Notices of the Royal Astronomical Society*, **396**, 121–140, 2009.

Edvardsson, B., Andersen, J., Gustafsson, B., Lambert, D. L., Nissen, P. E., et al. “The Chemical Evolution of the Galactic Disk - Part One - Analysis and Results”. *Astronomy and Astrophysics*, **275**, 101, 1993.

Efstathiou, G. & Jones, B. J. T. “The rotation of galaxies - Numerical investigations of the tidal torque theory”. *Monthly Notices of the Royal Astronomical Society*, **186**, 133–144, 1979.

Eggen, O. J., Lynden-Bell, D., & Sandage, A. R. “Evidence from the motions of old stars that the Galaxy collapsed.” *Astrophysical Journal*, **136**, 748, 1962.

Eisenstein, D. J., Weinberg, D. H., Agol, E., Aihara, H., Allende Prieto, C., et al. "SDSS-III: Massive Spectroscopic Surveys of the Distant Universe, the Milky Way, and Extra-Solar Planetary Systems". *Astronomical Journal*, **142**, 72, 2011.

ESA (editor). *The HIPPARCOS and TYCHO catalogues. Astrometric and photometric star catalogues derived from the ESA HIPPARCOS Space Astrometry Mission*, volume 1200 of *ESA Special Publication*, 1997.

ESA. "Gaia Data Release Scenario". `https://www.cosmos.esa.int/web/gaia/release`, 2017a. Accessed: 2017-05-15.

ESA. "Gaia Science Performance". `https://www.cosmos.esa.int/web/gaia/science-performance`, 2017b. Accessed: 2017-05-15.

ESA. "The Gaia Archive". `https://gea.esac.esa.int/archive/`, 2017c. Accessed: 2017-05-15.

Fall, S. M. & Efstathiou, G. "Formation and rotation of disc galaxies with haloes". *Monthly Notices of the Royal Astronomical Society*, **193**, 189–206, 1980.

Famaey, B. "Gaia and the dynamics of the Galaxy". In "SF2A-2012: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics", , edited by Boissier, S., de Laverny, P., Nardetto, N., Samadi, R., Valls-Gabaud, D., et al., pp. 15–24, 2012.

Famaey, B. & Dejonghe, H. "Three-component Stäckel potentials satisfying recent estimates of Milky Way parameters". *Monthly Notices of the Royal Astronomical Society*, **340**, 752–762, 2003.

Famaey, B., Jorissen, A., Luri, X., Mayor, M., Udry, S., et al. "Local kinematics of K and M giants from CORAVEL/Hipparcos/Tycho-2 data. Revisiting the concept of superclusters". *Astronomy and Astrophysics*, **430**, 165–186, 2005.

Famaey, B., Siebert, A., & Jorissen, A. "On the age heterogeneity of the Pleiades, Hyades, and Sirius moving groups". *Astronomy and Astrophysics*, **483**, 453–459, 2008.

Fan, J., Katz, A., Randall, L., & Reece, M. "Double-Disk Dark Matter". *Physics of the Dark Universe*, **2**, 139–156, 2013.

Fardal, M. A., Huang, S., & Weinberg, M. D. "Generation of mock tidal streams". *Monthly Notices of the Royal Astronomical Society*, **452**, 301–319, 2015.

Feldmeier-Krause, A., Zhu, L., Neumayer, N., van de Ven, G., de Zeeuw, P. T., et al. "Triaxial orbit-based modelling of the Milky Way Nuclear Star Cluster". *Monthly Notices of the Royal Astronomical Society*, **466**, 4040–4052, 2017.

Feng, J. L. "Dark Matter Candidates from Particle Physics and Methods of Detection". *Annual Review of Astronomy and Astrophysics*, **48**, 495–545, 2010.

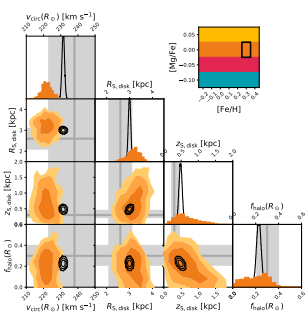
Ferrarese, L. & Ford, H. "Supermassive Black Holes in Galactic Nuclei: Past, Present and Future Research". *Space Science Reviews*, **116**, 523–624, 2005.

Ferrarese, L. & Merritt, D. "A Fundamental Relation between Supermassive Black Holes and Their Host Galaxies". *Astrophysical Journal, Letters*, **539**, L9–L12, 2000.

Fitzpatrick, R. *Plasma Physics: An Introduction*. CRC Press, 2014.

Flynn, C., Holmberg, J., Portinari, L., Fuchs, B., & Jahreiß, H. "On the mass-to-light ratio of the local Galactic disc and the optical luminosity of the Galaxy". *Monthly Notices of the Royal Astronomical Society*, **372**, 1149–1160, 2006.

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. "emcee: The MCMC Hammer". *Publications of the Astronomical Society of the Pacific*, **125**, 306–312, 2013.



Fouvry, J.-B., Binney, J., & Pichon, C. "Self-gravity, Resonances, and Orbital Diffusion in Stellar Disks". *Astrophysical Journal*, **806**, 117, 2015.

Frebel, A. & Norris, J. E. "Near-Field Cosmology with Extremely Metal-Poor Stars". *Annual Review of Astronomy and Astrophysics*, **53**, 631–688, 2015.

Freeman, K. & Bland-Hawthorn, J. "The New Galaxy: Signatures of Its Formation". *Annual Review of Astronomy and Astrophysics*, **40**, 487–537, 2002.

Freeman, K. C. "On the Disks of Spiral and S0 Galaxies". *Astrophysical Journal*, **160**, 811, 1970.

Friedmann, A. "Über die Krümmung des Raumes". *Zeitschrift fur Physik*, **10**, 377–386, 1922.

Fuhrmann, K. "Nearby stars of the Galactic disk and halo". *Astronomy and Astrophysics*, **338**, 161–183, 1998.

Fux, R. "Order and chaos in the local disc stellar kinematics induced by the Galactic bar". *Astronomy and Astrophysics*, **373**, 511–535, 2001.

Gaia Collaboration, Brown, A. G. A., Vallenari, A., Prusti, T., de Bruijne, J. H. J., et al. "Gaia Data Release 1. Summary of the astrometric, photometric, and survey properties". *Astronomy and Astrophysics*, **595**, A2, 2016a.

Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., Brown, A. G. A., Vallenari, A., et al. "The Gaia mission". *Astronomy and Astrophysics*, **595**, A1, 2016b.

Gaia@AIP. "Getting started âĂŞ Examples of Gaia DR1 MySQL queries". `https://gaia.aip.de/cms/documentation/getting-started-examples-of-gaia-dr1-mysql-queries/`, 2016. Accessed: 2017-05-16.

Gamow, G. "The Origin of Elements and the Separation of Galaxies". *Physical Review*, **74**, 505–506, 1948.

Garbari, S., Liu, C., Read, J. I., & Lake, G. "A new determination of the local dark matter density from the kinematics of K dwarfs". *Monthly Notices of the Royal Astronomical Society*, **425**, 1445–1458, 2012.

Georgelin, Y. M. & Georgelin, Y. P. "The spiral structure of our Galaxy determined from H II regions". *Astronomy and Astrophysics*, **49**, 57–79, 1976.

Gerhard, O. "Pattern speeds in the Milky Way." *Memorie della Societa Astronomica Italiana Supplementi*, **18**, 185, 2011.

Gerhard, O. E. "Dynamics of the Galaxy". In "Galaxy Dynamics - A Rutgers Symposium", , edited by Merritt, D. R., Valluri, M., & Sellwood, J. A., volume 182 of *Astronomical Society of the Pacific Conference Series*, 1999.

Ghez, A. M., Salim, S., Weinberg, N. N., Lu, J. R., Do, T., et al. "Measuring Distance and Properties of the Milky Way's Central Supermassive Black Hole with Stellar Orbits". *Astrophysical Journal*, **689**, 1044-1062, 2008.

Gillessen, S., Eisenhauer, F., Fritz, T. K., Bartko, H., Dodds-Eden, K., et al. "The Orbit of the Star S2 Around SGR A* from Very Large Telescope and Keck Data". *Astrophysical Journal, Letters*, **707**, L114–L117, 2009a.

Gillessen, S., Eisenhauer, F., Trippe, S., Alexander, T., Genzel, R., et al. "Monitoring Stellar Orbits Around the Massive Black Hole in the Galactic Center". *Astrophysical Journal*, **692**, 1075–1109, 2009b.

Gilmore, G., Randich, S., Asplund, M., Binney, J., Bonifacio, P., et al. "The Gaia-ESO Public Spectroscopic Survey". *The Messenger*, **147**, 25–31, 2012.

Gilmore, G. & Reid, N. "New light on faint stars. III - Galactic structure towards the South Pole and the Galactic thick disc". *Monthly Notices of the Royal Astronomical Society*, **202**, 1025–1047, 1983.

Gilmore, G. & Wyse, R. F. G. "Chemical evolution with bursts of star formation - Element ratios in dwarf galaxies". *Astrophysical Journal, Letters*, **367**, L55–L58, 1991.

Gilmore, G., Wyse, R. F. G., & Jones, J. B. "A determination of the thick disk chemical abundance distribution: Implications for galaxy evolution". *Astronomical Journal*, **109**, 1095–1111, 1995.

Gilmore, G., Wyse, R. F. G., & Kuijken, K. "Kinematics, chemistry, and structure of the Galaxy". *Annual Review of Astronomy and Astrophysics*, **27**, 555–627, 1989.

Girardi, L. "Red Clump Stars". *Annual Review of Astronomy and Astrophysics*, **54**, 95–133, 2016.

Girardi, L. & Salaris, M. "Population effects on the red giant clump absolute magnitude, and distance determinations to nearby galaxies". *Monthly Notices of the Royal Astronomical Society*, **323**, 109–129, 2001.

Glatt, K., Grebel, E. K., Sabbi, E., Gallagher, J. S., III, Nota, A., et al. "Age Determination of Six Intermediate-Age Small Magellanic Cloud Star Clusters with HST/ACS". *Astronomical Journal*, **136**, 1703-1727, 2008.

Gómez, F. A., White, S. D. M., Marinacci, F., Slater, C. T., Grand, R. J. J., et al. "A fully cosmological model of a Monoceros-like ring". *Monthly Notices of the Royal Astronomical Society*, **456**, 2779–2793, 2016.

Górski, K. M., Hivon, E., Banday, A. J., Wandelt, B. D., Hansen, F. K., et al. "HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere". *Astrophysical Journal*, **622**, 759–771, 2005.

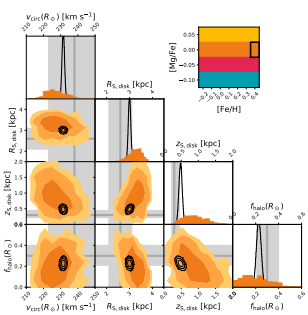
Gott, J. R., III, Jurić, M., Schlegel, D., Hoyle, F., Vogeley, M., et al. "A Map of the Universe". *Astrophysical Journal*, **624**, 463–484, 2005.

Gould, A. & Kollmeier, J. A. "Proper-Motion Catalog from SDSS ∩ USNO-B". *Astrophysical Journal, Supplement*, **152**, 103–111, 2004.

Graham, A. W. "An Investigation into the Prominence of Spiral Galaxy Bulges". *Astronomical Journal*, **121**, 820–840, 2001.

Grand, R. J. J., Gómez, F. A., Marinacci, F., Pakmor, R., Springel, V., et al. "The Auriga Project: the properties and formation mechanisms of disc galaxies across cosmic time". *Monthly Notices of the Royal Astronomical Society*, **467**, 179–207, 2017.

Grand, R. J. J., Kawata, D., & Cropper, M. "The dynamics of stars around spiral arms". *Monthly Notices of the Royal Astronomical Society*, **421**, 1529–1538, 2012.



Gratton, R. G., Carretta, E., Matteucci, F., & Sneden, C. "Abundances of light elements in metal-poor stars. IV. [Fe/O] and [Fe/Mg] ratios and the history of star formation in the solar neighborhood". *Astronomy and Astrophysics*, **358**, 671–681, 2000.

Green, G. M., Schlafly, E. F., Finkbeiner, D. P., Rix, H.-W., Martin, N., et al. "A Three-dimensional Map of Milky Way Dust". *Astrophysical Journal*, **810**, 25, 2015.

Gregory, S. A. & Thompson, L. A. "The Coma/A1367 supercluster and its environs". *Astrophysical Journal*, **222**, 784–799, 1978.

Grocholski, A. J. & Sarajedini, A. "WIYN Open Cluster Study. X. The K-Band Magnitude of the Red Clump as a Distance Indicator". *Astronomical Journal*, **123**, 1603–1612, 2002.

Groenewegen, M. A. T. "The red clump absolute magnitude based on revised Hipparcos parallaxes". *Astronomy and Astrophysics*, **488**, 935–941, 2008.

Guth, A. H. "Inflationary universe: A possible solution to the horizon and flatness problems". *Physical Review D*, **23**, 347–356, 1981.

Guth, A. H. & Pi, S.-Y. "Fluctuations in the new inflationary universe". *Physical Review Letters*, **49**, 1110–1113, 1982.

Hambly, N. C., MacGillivray, H. T., Read, M. A., Tritton, S. B., Thomson, E. B., et al. "The SuperCOSMOS Sky Survey - I. Introduction and description". *Monthly Notices of the Royal Astronomical Society*, **326**, 1279–1294, 2001.

Hammer, F., Puech, M., Chemin, L., Flores, H., & Lehnert, M. D. "The Milky Way, an Exceptionally Quiet Galaxy: Implications for the Formation of Spiral Galaxies". *Astrophysical Journal*, **662**, 322–334, 2007.

Hammersley, P. L., Garzón, F., Mahoney, T. J., López-Corredoira, M., & Torres, M. A. P. "Detection of the old stellar component of the major Galactic bar". *Monthly Notices of the Royal Astronomical Society*, **317**, L45–L49, 2000.

Häring, N. & Rix, H.-W. "On the Black Hole Mass-Bulge Mass Relation". *Astrophysical Journal, Letters*, **604**, L89–L92, 2004.

Harrison, E. R. "Fluctuations at the Threshold of Classical Cosmology". *Physical Review D*, **1**, 2726–2730, 1970.

Hawking, S. W. "The development of irregularities in a single bubble inflationary universe". *Physics Letters B*, **115**, 295–297, 1982.

Hawkins, K., Leistedt, B., Bovy, J., & Hogg, D. W. "Red clump stars and Gaia: Calibration of the standard candle using a hierarchical probabilistic model". *ArXiv e-prints*, 2017.

Hayden, M. R., Bovy, J., Holtzman, J. A., Nidever, D. L., Bird, J. C., et al. "Chemical Cartography with APOGEE: Metallicity Distribution Functions and the Chemical Structure of the Milky Way Disk". *Astrophysical Journal*, **808**, 132, 2015.

Haywood, M., Di Matteo, P., Lehnert, M. D., Katz, D., & Gómez, A. "The age structure of stellar populations in the solar vicinity. Clues of a two-phase formation history of the Milky Way disk". *Astronomy and Astrophysics*, **560**, A109, 2013.

Helmi, A., Veljanoski, J., Breddels, M. A., Tian, H., & Sales, L. V. "A box full of chocolates: The rich structure of the nearby stellar halo revealed by Gaia and RAVE". *Astronomy and Astrophysics*, **598**, A58, 2017.

Helmi, A., White, S. D. M., de Zeeuw, P. T., & Zhao, H. "Debris streams in the solar neighbourhood as relicts from the formation of the Milky Way". *Nature*, **402**, 53–55, 1999.

Henon, M. "L'amas isochrone: I". *Annales d'Astrophysique*, **22**, 126, 1959.

Hernquist, L. "An analytical model for spherical galaxies and bulges". *Astrophysical Journal*, **356**, 359–364, 1990.

Herpich, J., Tremaine, S., & Rix, H.-W. "Galactic disc profiles and a universal angular momentum distribution from statistical physics". *Monthly Notices of the Royal Astronomical Society*, **467**, 5022–5032, 2017.

Høg, E., Fabricius, C., Makarov, V. V., Urban, S., Corbin, T., et al. "The Tycho-2 catalogue of the 2.5 million brightest stars". *Astronomy and Astrophysics*, **355**, L27–L30, 2000.

Hog, E., Fabricius, C., Makarov, V. V., Urban, S., Corbin, T., et al. "VizieR Online Data Catalog: The Tycho-2 Catalogue (Hog+ 2000)". *VizieR Online Data Catalog*, **1259**, 2000.

Hogg, D. W., Bovy, J., & Lang, D. "Data analysis recipes: Fitting a model to data". *ArXiv e-prints*, 2010.

Hogg, D. W., Casey, A. R., Ness, M., Rix, H.-W., Foreman-Mackey, D., et al. "Chemical Tagging Can Work: Identification of Stellar Phase-space Structures Purely by Chemical-abundance Similarity". *Astrophysical Journal*, **833**, 262, 2016.

Holmberg, J. & Flynn, C. "The local surface density of disc matter mapped by Hipparcos". *Monthly Notices of the Royal Astronomical Society*, **352**, 440–446, 2004.

Holmberg, J., Nordström, B., & Andersen, J. "The Geneva-Copenhagen survey of the solar neighbourhood. III. Improved distances, ages, and kinematics". *Astronomy and Astrophysics*, **501**, 941–947, 2009.

Holtzman, J. A., Shetrone, M., Johnson, J. A., Allende Prieto, C., Anders, F., et al. "Abundances, Stellar Parameters, and Spectra from the SDSS-III/APOGEE Survey". *Astronomical Journal*, **150**, 148, 2015.

Howes, L. M., Casey, A. R., Asplund, M., Keller, S. C., Yong, D., et al. "Extremely metal-poor stars from the cosmic dawn in the bulge of the Milky Way". *Nature*, **527**, 484–487, 2015.

Hoyle, F. "Problems of Cosmical Aerodynamics: Proceedings of the Symposium on the Motion of Gaseous Masses of Cosmical Dimensions Held at Paris, August 16 - 19, 1949". Central Air Documents Office, 1951.

Hoyle, F. & Tayler, R. J. "The Mystery of the Cosmic Helium Abundance". *Nature*, **203**, 1108–1110, 1964.

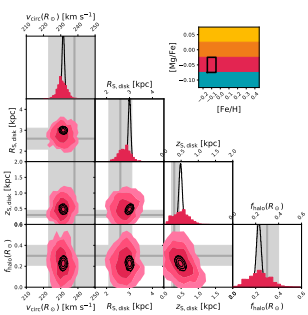
Hubble, E. "A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae". *Proceedings of the National Academy of Science*, **15**, 168–173, 1929.

Huber, D., Bryson, S. T., Haas, M. R., Barclay, T., Barentsen, G., et al. "The K2 Ecliptic Plane Input Catalog (EPIC) and Stellar Classifications of 138,600 Targets in Campaigns 1-8". *Astrophysical Journal, Supplement*, **224**, 2, 2016.

Hunt, J. A. S., Bovy, J., & Carlberg, R. G. "Detection of a Dearth of Stars with Zero Angular Momentum in the Solar Neighborhood". *Astrophysical Journal, Letters*, **832**, L25, 2016.

Hunt, J. A. S. & Kawata, D. "Disc galaxy modelling with a particle-by-particle made-to-measure method". *Monthly Notices of the Royal Astronomical Society*, **430**, 1928–1939, 2013.

Hunt, J. A. S. & Kawata, D. "M2M modelling of the Galactic disc via PRIMAL: fitting to Gaia error added data". *Monthly Notices of the Royal Astronomical Society*, **443**, 2112–2125, 2014.



Hunter, J. D. "Matplotlib: A 2D graphics environment". *Computing In Science & Engineering*, **9**(3), 90–95, 2007.

Ibata, R. A., Gilmore, G., & Irwin, M. J. "A dwarf satellite galaxy in Sagittarius". *Nature*, **370**, 194–196, 1994.

Ivezić, Ž., Monet, D. G., Bond, N., Jurić, M., Sesar, B., et al. "Astrometry with digital sky surveys: from SDSS to LSST". In "A Giant Step: from Milli- to Micro-arcsecond Astrometry", , edited by Jin, W. J., Platais, I., & Perryman, M. A. C., volume 248 of *IAU Symposium*, pp. 537–543, 2008.

Ivezic, Z., Tyson, J. A., Abel, B., Acosta, E., Allsman, R., et al. "LSST: from Science Drivers to Reference Design and Anticipated Data Products". *ArXiv e-prints*, 2008.

Jablonka, P., Martin, P., & Arimoto, N. "The Luminosity-Metallicity Relation for Bulges of Spiral Galaxies". *Astronomical Journal*, **112**, 1415, 1996.

Jeans, J. H. "The Stability of a Spherical Nebula". *Philosophical Transactions of the Royal Society of London Series A*, **199**, 1–53, 1902.

Jeans, J. H. "On the theory of star-streaming and the structure of the universe". *Monthly Notices of the Royal Astronomical Society*, **76**, 70–84, 1915.

Jones, E., Oliphant, T., Peterson, P., et al. "SciPy: Open source scientific tools for Python", 2001–2017. [Online; accessed ¡today¿].

Jordi, C., Gebran, M., Carrasco, J. M., de Bruijne, J., Voss, H., et al. "Gaia broad band photometry". *Astronomy and Astrophysics*, **523**, A48, 2010.

Jordi, C., Høg, E., Brown, A. G. A., Lindegren, L., Bailer-Jones, C. A. L., et al. "The design and performance of the Gaia photometric system". *Monthly Notices of the Royal Astronomical Society*, **367**, 290–314, 2006.

Jungman, G., Kamionkowski, M., & Griest, K. "Supersymmetric dark matter". *Physics Reports*, **267**, 195–373, 1996.

Jurić, M., Ivezić, Ž., Brooks, A., Lupton, R. H., Schlegel, D., et al. "The Milky Way Tomography with SDSS. I. Stellar Number Density Distribution". *Astrophysical Journal*, **673**, 864-914, 2008.

Kaiser, N. "On the spatial correlations of Abell clusters". *Astrophysical Journal, Letters*, **284**, L9–L12, 1984.

Kaiser, N., Aussel, H., Burke, B. E., Boesgaard, H., Chambers, K., et al. "Pan-STARRS: A Large Synoptic Survey Telescope Array". In "Survey and Other Telescope Technologies and Discoveries", , edited by Tyson, J. A. & Wolff, S., volume 4836 of *Proceedings of the SPIE*, pp. 154–164, 2002.

Kaiser, N., Burgett, W., Chambers, K., Denneau, L., Heasley, J., et al. "The Pan-STARRS wide-field optical/NIR imaging survey". In "Ground-based and Airborne Telescopes III", volume 7733 of *Proceedings of the SPIE*, p. 77330E, 2010.

Kaiser, N. & Squires, G. "Mapping the dark matter with weak gravitational lensing". *Astrophysical Journal*, **404**, 441–450, 1993.

Katz, D., Munari, U., Cropper, M., Zwitter, T., Thévenin, F., et al. "Spectroscopic survey of the Galaxy with Gaia- I. Design and performance of the Radial Velocity Spectrometer". *Monthly Notices of the Royal Astronomical Society*, **354**, 1223–1238, 2004.

Kauffmann, G., Colberg, J. M., Diaferio, A., & White, S. D. M. "Clustering of galaxies in a hierarchical universe - I. Methods and results at z=0". *Monthly Notices of the Royal Astronomical Society*, **303**, 188–206, 1999.

Kauffmann, G., White, S. D. M., & Guiderdoni, B. "The Formation and Evolution of Galaxies Within Merging Dark Matter Haloes". *Monthly Notices of the Royal Astronomical Society*, **264**, 201, 1993.

Kennicutt, R. C., Jr. "The Global Schmidt Law in Star-forming Galaxies". *Astrophysical Journal*, **498**, 541–552, 1998.

Kepner, J. V. "Inside-out Galaxy Formation". *Astrophysical Journal*, **520**, 59–66, 1999.

Kerr, F. J. & Lynden-Bell, D. "Review of galactic constants". *Monthly Notices of the Royal Astronomical Society*, **221**, 1023–1038, 1986.

Kirkman, D., Tytler, D., Suzuki, N., O'Meara, J. M., & Lubin, D. "The Cosmological Baryon Density from the Deuterium-to-Hydrogen Ratio in QSO Absorption Systems: D/H toward Q1243+3047". *Astrophysical Journal, Supplement*, **149**, 1–28, 2003.

Klement, R., Fuchs, B., & Rix, H.-W. "Identifying Stellar Streams in the First RAVE Public Data Release". *Astrophysical Journal*, **685**, 261-271, 2008.

Kleyna, J. T., Wilkinson, M. I., Gilmore, G., & Evans, N. W. "A Dynamical Fossil in the Ursa Minor Dwarf Spheroidal Galaxy". *Astrophysical Journal, Letters*, **588**, L21–L24, 2003.

Klypin, A., Kravtsov, A. V., Valenzuela, O., & Prada, F. "Where Are the Missing Galactic Satellites?" *Astrophysical Journal*, **522**, 82–92, 1999.

Klypin, A., Zhao, H., & Somerville, R. S. "ΛCDM-based Models for the Milky Way and M31. I. Dynamical Models". *Astrophysical Journal*, **573**, 597–613, 2002.

Koposov, S. E., Rix, H.-W., & Hogg, D. W. "Constraining the Milky Way Potential with a Six-Dimensional Phase-Space Map of the GD-1 Stellar Stream". *Astrophysical Journal*, **712**, 260–273, 2010.

Kordopatis, G., Binney, J., Gilmore, G., Wyse, R. F. G., Belokurov, V., et al. "The rich are different: evidence from the RAVE survey for stellar radial migration". *Monthly Notices of the Royal Astronomical Society*, **447**, 3526–3535, 2015a.

Kordopatis, G. & RAVE Collaboration. "The Radial Velocity Experiment (RAVE): Fourth Data Release". In "American Astronomical Society Meeting Abstracts #223", volume 223 of *American Astronomical Society Meeting Abstracts*, p. 346.02, 2014.

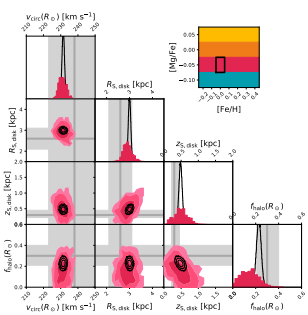
Kordopatis, G., Recio-Blanco, A., de Laverny, P., Gilmore, G., Hill, V., et al. "A spectroscopic survey of thick disc stars outside the solar neighbourhood". *Astronomy and Astrophysics*, **535**, A107, 2011.

Kordopatis, G., Wyse, R. F. G., Chiappini, C., Minchev, I., Anders, F., et al. "Cardinal kinematics - I. Rotation fields of the APOGEE survey". *Monthly Notices of the Royal Astronomical Society*, **467**, 469–489, 2017.

Kordopatis, G., Wyse, R. F. G., Gilmore, G., Recio-Blanco, A., de Laverny, P., et al. "The Gaia-ESO Survey: characterisation of the [α/Fe] sequences in the Milky Way discs". *Astronomy and Astrophysics*, **582**, A122, 2015b.

Kormendy, J. & Richstone, D. "Inward Bound—The Search For Supermassive Black Holes In Galactic Nuclei". *Annual Review of Astronomy and Astrophysics*, **33**, 581, 1995.

Kramer, E. D. & Randall, L. "Updated Kinematic Constraints on a Dark Disk". *Astrophysical Journal*, **824**, 116, 2016.



Kuijken, K. & Gilmore, G. "The Mass Distribution in the Galactic Disc - II - Determination of the Surface Mass Density of the Galactic Disc Near the Sun". *Monthly Notices of the Royal Astronomical Society*, **239**, 605–649, 1989a.

Kuijken, K. & Gilmore, G. "The mass distribution in the galactic disc. I - A technique to determine the integral surface mass density of the disc near the sun." *Monthly Notices of the Royal Astronomical Society*, **239**, 571–603, 1989b.

Kuijken, K. & Gilmore, G. "The galactic disk surface mass density and the Galactic force K(z) at Z = 1.1 kiloparsecs". *Astrophysical Journal, Letters*, **367**, L9–L13, 1991.

Kunder, A., Kordopatis, G., Steinmetz, M., Zwitter, T., McMillan, P. J., et al. "The Radial Velocity Experiment (RAVE): Fifth Data Release". *Astronomical Journal*, **153**, 75, 2017.

Lacey, C. G. & Ostriker, J. P. "Massive black holes in galactic halos?" *Astrophysical Journal*, **299**, 633–652, 1985.

Law, D. R., Majewski, S. R., & Johnston, K. V. "Evidence for a Triaxial Milky Way Dark Matter Halo from the Sagittarius Stellar Tidal Stream". *Astrophysical Journal, Letters*, **703**, L67–L71, 2009.

Lee, Y. S., Beers, T. C., An, D., Ivezić, Ž., Just, A., et al. "Formation and Evolution of the Disk System of the Milky Way: [α/Fe] Ratios and Kinematics of the SEGUE G-dwarf Sample". *Astrophysical Journal*, **738**, 187, 2011.

Lin, C. C. & Shu, F. H. "On the Spiral Structure of Disk Galaxies." *Astrophysical Journal*, **140**, 646, 1964.

Lin, D. N. C. & Pringle, J. E. "A viscosity prescription for a self-gravitating accretion disc". *Monthly Notices of the Royal Astronomical Society*, **225**, 607–613, 1987.

Lindegren, L., Lammers, U., Bastian, U., Hernández, J., Klioner, S., et al. "Gaia Data Release 1. Astrometry: one billion positions, two million proper motions and parallaxes". *Astronomy and Astrophysics*, **595**, A4, 2016.

Liszt, H. S. & Burton, W. B. "The gas distribution in the central region of the Galaxy. III - A barlike model of the inner-Galaxy gas based on improved H I data". *Astrophysical Journal*, **236**, 779–797, 1980.

Loebman, S. R., Roškar, R., Debattista, V. P., Ivezić, Ž., Quinn, T. R., et al. "The Genesis of the Milky Way's Thick Disk Via Stellar Migration". *Astrophysical Journal*, **737**, 8, 2011.

Long, R. J., Mao, S., Shen, J., & Wang, Y. "Made-to-measure galaxy models - III. Modelling with Milky Way observations". *Monthly Notices of the Royal Astronomical Society*, **428**, 3478–3486, 2013.

Lynden-Bell, D. & Kalnajs, A. J. "On the generating mechanism of spiral structure". *Monthly Notices of the Royal Astronomical Society*, **157**, 1, 1972.

MacArthur, L. A., Courteau, S., & Holtzman, J. A. "Structure of Disk-dominated Galaxies. I. Bulge/Disk Parameters, Simulations, and Secular Evolution". *Astrophysical Journal*, **582**, 689–722, 2003.

Mackereth, J. T., Bovy, J., Schiavon, R. P., Zasowski, G., Cunha, K., et al. "The age-metallicity structure of the Milky Way disk". *ArXiv e-prints*, 2017.

Madau, P. & Dickinson, M. "Cosmic Star-Formation History". *Annual Review of Astronomy and Astrophysics*, **52**, 415–486, 2014.

Magorrian, J. "Bayes versus the virial theorem: inferring the potential of a galaxy from a kinematical snapshot". *Monthly Notices of the Royal Astronomical Society*, **437**, 2230–2248, 2014.

Magorrian, J., Tremaine, S., Richstone, D., Bender, R., Bower, G., et al. "The Demography of Massive Dark Objects in Galaxy Centers". *Astronomical Journal*, **115**, 2285–2305, 1998.

Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., Allende Prieto, C., Barkhouser, R., et al. "The Apache Point Observatory Galactic Evolution Experiment (APOGEE)". *ArXiv e-prints*, 2015.

Majewski, S. R., Skrutskie, M. F., Weinberg, M. D., & Ostheimer, J. C. "A Two Micron All Sky Survey View of the Sagittarius Dwarf Galaxy. I. Morphology of the Sagittarius Core and Tidal Arms". *Astrophysical Journal*, **599**, 1082–1115, 2003.

Maji, M., Zhu, Q., Marinacci, F., & Li, Y. "Is there a disk of satellites around the Milky Way?" *ArXiv e-prints*, 2017.

Maoz, D. *Astrophysics in a Nutshell*. In a Nutshell. Princeton University Press, 2011.

Maoz, D., Mannucci, F., Li, W., Filippenko, A. V., Della Valle, M., et al. "Nearby supernova rates from the Lick Observatory Supernova Search - IV. A recovery method for the delay-time distribution". *Monthly Notices of the Royal Astronomical Society*, **412**, 1508–1521, 2011.

Marchetti, T., Rossi, E. M., Kordopatis, G., Brown, A. G. A., Rimoldi, A., et al. "An artificial neural network to discover Hypervelocity stars: Candidates in Gaia DR1/TGAS". *ArXiv e-prints*, 2017.

Markevitch, M., Gonzalez, A. H., David, L., Vikhlinin, A., Murray, S., et al. "A Textbook Example of a Bow Shock in the Merging Galaxy Cluster 1E 0657-56". *Astrophysical Journal, Letters*, **567**, L27–L31, 2002.

Martig, M., Fouesneau, M., Rix, H.-W., Ness, M., Mészáros, S., et al. "Red giant masses and ages derived from carbon and nitrogen abundances". *Monthly Notices of the Royal Astronomical Society*, **456**, 3655–3670, 2016.

Martig, M., Minchev, I., & Flynn, C. "Dissecting simulated disc galaxies - I. The structure of mono-age populations". *Monthly Notices of the Royal Astronomical Society*, **442**, 2474–2486, 2014.

Martos, M., Hernandez, X., Yáñez, M., Moreno, E., & Pichardo, B. "A plausible Galactic spiral pattern and its rotation speed". *Monthly Notices of the Royal Astronomical Society*, **350**, L47–L51, 2004.

Masseron, T. & Gilmore, G. "Carbon, nitrogen and α-element abundances determine the formation sequence of the Galactic thick and thin discs". *Monthly Notices of the Royal Astronomical Society*, **453**, 1855–1866, 2015.

Math Explorer's Club. "Topology and Geometry of Surfaces: Flat Life". `http://www.math.cornell.edu/%7Emec/Winter2009/Victor/part1.htm`, 2009. Accessed: 2017-06-28.

Mather, J. C., Cheng, E. S., Eplee, R. E., Jr., Isaacman, R. B., Meyer, S. S., et al. "A preliminary measurement of the cosmic microwave background spectrum by the Cosmic Background Explorer (COBE) satellite". *Astrophysical Journal, Letters*, **354**, L37–L40, 1990.

Matteucci, F. & Francois, P. "Galactic chemical evolution - Abundance gradients of individual elements". *Monthly Notices of the Royal Astronomical Society*, **239**, 885–904, 1989.

Matteucci, F. & Recchi, S. "On the Typical Timescale for the Chemical Enrichment from Type Ia Supernovae in Galaxies". *Astrophysical Journal*, **558**, 351–358, 2001.

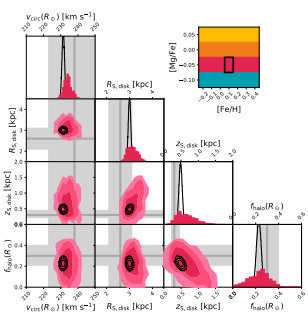
McCall, M. L. "On Determining Extinction from Reddening". *Astronomical Journal*, **128**, 2144–2169, 2004.

McClure-Griffiths, N. M. & Dickey, J. M. "Milky Way Kinematics. I. Measurements at the Subcentral Point of the Fourth Quadrant". *Astrophysical Journal*, **671**, 427–438, 2007.

McMillan, P. J. "Mass models of the Milky Way". *Monthly Notices of the Royal Astronomical Society*, **414**, 2446–2457, 2011.

McMillan, P. J. & Binney, J. "Analysing surveys of our Galaxy - I. Basic astrometric data". *Monthly Notices of the Royal Astronomical Society*, **419**, 2251–2263, 2012.

McMillan, P. J. & Binney, J. J. "Disassembling the Galaxy with angle-action coordinates". *Monthly Notices of the Royal Astronomical Society*, **390**, 429–437, 2008.



McMillan, P. J. & Binney, J. J. "Analysing surveys of our Galaxy - II. Determining the potential". *Monthly Notices of the Royal Astronomical Society*, **433**, 1411–1424, 2013.

McWilliam, A. "Abundance Ratios and Galactic Chemical Evolution". *Annual Review of Astronomy and Astrophysics*, **35**, 503–556, 1997.

Merritt, D. "Elliptical Galaxy Dynamics". *Publications of the Astronomical Society of the Pacific*, **111**, 129–168, 1999.

Mestel, L. "On the galactic law of rotation". *Monthly Notices of the Royal Astronomical Society*, **126**, 553, 1963.

Michalik, D. "Tycho-Gaia and beyond: The special character of Gaia data release 1", 2016.

Michalik, D., Lindegren, L., & Hobbs, D. "The Tycho-Gaia astrometric solution . How to get 2.5 million parallaxes with less than one year of Gaia data". *Astronomy and Astrophysics*, **574**, A115, 2015a.

Michalik, D., Lindegren, L., Hobbs, D., & Butkevich, A. G. "Gaia astrometry for stars with too few observations. A Bayesian approach". *Astronomy and Astrophysics*, **583**, A68, 2015b.

Michalik, D., Lindegren, L., Hobbs, D., & Lammers, U. "Joint astrometric solution of HIPPARCOS and Gaia. A recipe for the Hundred Thousand Proper Motions project". *Astronomy and Astrophysics*, **571**, A85, 2014.

Mignard, F. "The Hundred Thousand Proper Motions Project", 2009. GAIA-C3-TN-OCA-FM-040.

Milgrom, M. & Bekenstein, J. "The modified Newtonian dynamics as an alternative to hidden matter". In "Dark matter in the universe", , edited by Kormendy, J. & Knapp, G. R., volume 117 of *IAU Symposium*, pp. 319–330, 1987.

Minchev, I. "Constraining the Milky Way assembly history with Galactic Archaeology". *ArXiv e-prints*, 2017.

Minchev, I., Chiappini, C., & Martig, M. "Chemodynamical evolution of the Milky Way disk. II. Variations with Galactic radius and height above the disk plane". *Astronomy and Astrophysics*, **572**, A92, 2014.

Minchev, I., Famaey, B., Combes, F., Di Matteo, P., Mouhcine, M., et al. "Radial migration in galactic disks caused by resonance overlap of multiple patterns: Self-consistent simulations". *Astronomy and Astrophysics*, **527**, A147, 2011.

Minchev, I., Famaey, B., Quillen, A. C., Dehnen, W., Martig, M., et al. "Radial migration does little for Galactic disc thickening". *Astronomy and Astrophysics*, **548**, A127, 2012.

Minchev, I., Martig, M., Streich, D., Scannapieco, C., de Jong, R. S., et al. "On the Formation of Galactic Thick Disks". *Astrophysical Journal, Letters*, **804**, L9, 2015.

Minchev, I., Steinmetz, M., Chiappini, C., Martig, M., Anders, F., et al. "The Relationship between Mono-abundance and Mono-age Stellar Populations in the Milky Way Disk". *Astrophysical Journal*, **834**, 27, 2017.

Miyamoto, M. & Nagai, R. "Three-dimensional models for the distribution of mass in galaxies". *Publications of the Astronomical Society of Japan*, **27**, 533–543, 1975.

Mo, H., van den Bosch, F., & White, S. *Galaxy Formation and Evolution*. Galaxy Formation and Evolution. Cambridge University Press, 2010.

Mo, H. J., Mao, S., & White, S. D. M. "The formation of galactic discs". *Monthly Notices of the Royal Astronomical Society*, **295**, 319–336, 1998.

Momany, Y., Zaggia, S., Gilmore, G., Piotto, G., Carraro, G., et al. "Outer structure of the Galactic warp and flare: explaining the Canis Major over-density". *Astronomy and Astrophysics*, **451**, 515–538, 2006.

Monari, G., Famaey, B., & Siebert, A. "Modelling the Galactic disc: perturbed distribution functions in the presence of spiral arms". *Monthly Notices of the Royal Astronomical Society*, **457**, 2569–2582, 2016.

Monari, G., Famaey, B., Siebert, A., Duchateau, A., Lorscheider, T., et al. "Staying away from the bar: the local dynamical signature of slow and fast bars in the Milky Way". *Monthly Notices of the Royal Astronomical Society*, **465**, 1443–1453, 2017a.

Monari, G., Kawata, D., Hunt, J. A. S., & Famaey, B. "Tracing the Hercules stream with Gaia and LAMOST: new evidence for a fast bar in the Milky Way". *Monthly Notices of the Royal Astronomical Society*, **466**, L113–L117, 2017b.

Monet, D. G., Levine, S. E., Canzian, B., Ables, H. D., Bird, A. R., et al. "The USNO-B Catalog". *Astronomical Journal*, **125**, 984–993, 2003.

Moore, B., Ghigna, S., Governato, F., Lake, G., Quinn, T., et al. "Dark Matter Substructure within Galactic Halos". *Astrophysical Journal, Letters*, **524**, L19–L22, 1999.

Moore, B., Governato, F., Quinn, T., Stadel, J., & Lake, G. "Resolving the Structure of Cold Dark Matter Halos". *Astrophysical Journal, Letters*, **499**, L5–L8, 1998.

Morgan, W. W., Whitford, A. E., & Code, A. D. "Studies in Galactic Structure. I. a Preliminary Determination of the Space Distribution of the Blue Giants." *Astrophysical Journal*, **118**, 318, 1953.

Moster, B. P., Naab, T., & White, S. D. M. "Galactic star formation and accretion histories from matching galaxies to dark matter haloes". *Monthly Notices of the Royal Astronomical Society*, **428**, 3121–3138, 2013.

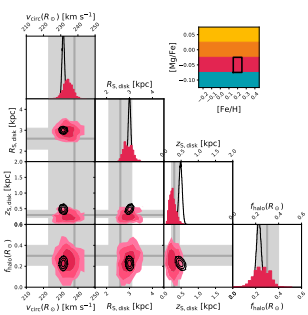
Moster, B. P., Somerville, R. S., Maulbetsch, C., van den Bosch, F. C., Macciò, A. V., et al. "Constraints on the Relationship between Stellar Mass and Halo Mass at Low and High Redshift". *Astrophysical Journal*, **710**, 903–923, 2010.

Munn, J. A., Monet, D. G., Levine, S. E., Canzian, B., Pier, J. R., et al. "An Improved Proper-Motion Catalog Combining USNO-B and the Sloan Digital Sky Survey". *Astronomical Journal*, **127**, 3034–3042, 2004.

Munn, J. A., Monet, D. G., Levine, S. E., Canzian, B., Pier, J. R., et al. "Erratum: "an Improved Proper-Motion Catalog Combining Usno-B and the Sloan Digital Sky Survey" (2004, AJ, 127, 3034)". *Astronomical Journal*, **136**, 895, 2008.

Myeong, G. C., Evans, N. W., Belokurov, V., Koposov, S. E., & Sanders, J. L. "A halo substructure in Gaia Data Release 1". *Monthly Notices of the Royal Astronomical Society*, **469**, L78–L82, 2017.

Naab, T. & Ostriker, J. P. "Theoretical Challenges in Galaxy Formation". *ArXiv e-prints*, 2016.



Nataf, D. M., Gould, A., Fouqué, P., Gonzalez, O. A., Johnson, J. A., et al. "Reddening and Extinction toward the Galactic Bulge from OGLE-III: The Inner Milky Way's R_V ~ 2.5 Extinction Curve". *Astrophysical Journal*, **769**, 88, 2013.

Nataf, D. M., Udalski, A., Gould, A., Fouqué, P., & Stanek, K. Z. "The Split Red Clump of the Galactic Bulge from OGLE-III". *Astrophysical Journal, Letters*, **721**, L28–L32, 2010.

Navarro, J. F., Frenk, C. S., & White, S. D. M. "The Structure of Cold Dark Matter Halos". *Astrophysical Journal*, **462**, 563, 1996.

Navarro, J. F., Frenk, C. S., & White, S. D. M. "A Universal Density Profile from Hierarchical Clustering". *Astrophysical Journal*, **490**, 493–508, 1997.

Navarro, J. F., Helmi, A., & Freeman, K. C. "The Extragalactic Origin of the Arcturus Group". *Astrophysical Journal, Letters*, **601**, L43–L46, 2004.

Navarro, J. F. & Steinmetz, M. "Dark Halo and Disk Galaxy Scaling Laws in Hierarchical Universes". *Astrophysical Journal*, **538**, 477–488, 2000.

Ness, M., Freeman, K., Athanassoula, E., Wylie-de-Boer, E., Bland-Hawthorn, J., et al. "ARGOS - III. Stellar populations in the Galactic bulge of the Milky Way". *Monthly Notices of the Royal Astronomical Society*, **430**, 836–857, 2013.

Ness, M., Hogg, D. W., Rix, H.-W., Ho, A. Y. Q., & Zasowski, G. "The Cannon: A data-driven approach to Stellar Label Determination". *Astrophysical Journal*, **808**, 16, 2015.

Ness, M., Hogg, D. W., Rix, H.-W., Martig, M., Pinsonneault, M. H., et al. "Spectroscopic Determination of Masses (and Implied Ages) for Red Giants". *Astrophysical Journal*, **823**, 114, 2016.

Newberg, H. J., Yanny, B., Grebel, E. K., Hennessy, G., Ivezić, Ž., et al. "Sagittarius Tidal Debris 90 Kiloparsecs from the Galactic Center". *Astrophysical Journal, Letters*, **596**, L191–L194, 2003.

Newberg, H. J., Yanny, B., Rockosi, C., Grebel, E. K., Rix, H.-W., et al. "The Ghost of Sagittarius and Lumps in the Halo of the Milky Way". *Astrophysical Journal*, **569**, 245–274, 2002.

Nidever, D. L., Bovy, J., Bird, J. C., Andrews, B. H., Hayden, M., et al. "Tracing Chemical Evolution over the Extent of the Milky Way's Disk with APOGEE Red Clump Stars". *Astrophysical Journal*, **796**, 38, 2014.

Nidever, D. L., Zasowski, G., & Majewski, S. R. "Lifting the Dusty Veil with Near- and Mid-infrared Photometry. III. Two-dimensional Extinction Maps of the Galactic Midplane Using the Rayleigh-Jeans Color Excess Method". *Astrophysical Journal, Supplement*, **201**, 35, 2012.

Nordström, B., Mayor, M., Andersen, J., Holmberg, J., Pont, F., et al. "The Geneva-Copenhagen survey of the Solar neighbourhood. Ages, metallicities, and kinematic properties of ˜14 000 F and G dwarfs". *Astronomy and Astrophysics*, **418**, 989–1019, 2004.

Norris, J. "Population studies - The nature of the thick disk". *Astrophysical Journal, Letters*, **314**, L39–L43, 1987.

O'Leary, R. M., Kistler, M. D., Kerr, M., & Dexter, J. "Young and Millisecond Pulsar GeV Gamma-ray Fluxes from the Galactic Center and Beyond". *ArXiv e-prints*, 2016.

Oort, J. H. "The force exerted by the stellar system in the direction perpendicular to the galactic plane and some related problems". *Bulletin Astronomical Institute of the Netherlands*, **6**, 249, 1932.

Ortolani, S., Renzini, A., Gilmozzi, R., Marconi, G., Barbuy, B., et al. "Near-coeval formation of the Galactic bulge and halo inferred from globular cluster ages". *Nature*, **377**, 701–704, 1995.

Ostriker, J. P. & Steinhardt, P. "New Light on Dark Matter". *Science*, **300**, 1909–1914, 2003.

Paczynski, B. "Gravitational microlensing by the galactic halo". *Astrophysical Journal*, **304**, 1–5, 1986.

Paczyński, B. & Stanek, K. Z. "Galactocentric Distance with the Optical Gravitational Lensing Experiment and HIPPARCOS Red Clump Stars". *Astrophysical Journal, Letters*, **494**, L219–L222, 1998.

Peebles, P. J. E. "Rotation of Galaxies and the Gravitational Instability Picture". *Astronomy and Astrophysics*, **11**, 377, 1971.

Penzias, A. A. & Wilson, R. W. "A Measurement of Excess Antenna Temperature at 4080 Mc/s." *Astrophysical Journal*, **142**, 419–421, 1965.

Perlmutter, S., Aldering, G., Goldhaber, G., Knop, R. A., Nugent, P., et al. "Measurements of Ω and Λ from 42 High-Redshift Supernovae". *Astrophysical Journal*, **517**, 565–586, 1999.

Perryman, M. A. C. "Overview of the Gaia Mission". In "Astrometry in the Age of the Next Generation of Large Telescopes", , edited by Seidelmann, P. K. & Monet, A. K. B., volume 338 of *Astronomical Society of the Pacific Conference Series*, p. 3, 2005.

Perryman, M. A. C., de Boer, K. S., Gilmore, G., Høg, E., Lattanzi, M. G., et al. "GAIA: Composition, formation and evolution of the Galaxy". *Astronomy and Astrophysics*, **369**, 339–363, 2001.

Perryman, M. A. C., Lindegren, L., Kovalevsky, J., Hoeg, E., Bastian, U., et al. "The HIPPARCOS Catalogue". *Astronomy and Astrophysics*, **323**, L49–L52, 1997.

Pierce, M. J. & Tully, R. B. "Luminosity-line width relations and the extragalactic distance scale. I - Absolute calibration". *Astrophysical Journal*, **387**, 47–55, 1992.

Pietrzyński, G., Gieren, W., & Udalski, A. "The Araucaria Project: Dependence of Mean K, J, and I Absolute Magnitudes of Red Clump Stars on Metallicity and Age". *Astronomical Journal*, **125**, 2494–2501, 2003.

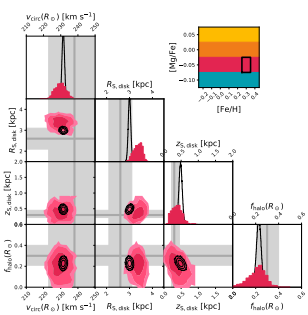
Piffl, T., Binney, J., McMillan, P. J., Steinmetz, M., Helmi, A., et al. "Constraining the Galaxy's dark halo with RAVE stars". *Monthly Notices of the Royal Astronomical Society*, **445**, 3133–3151, 2014.

Pinsonneault, M. H., Elsworth, Y., Epstein, C., Hekker, S., Mészáros, S., et al. "The APOKASC Catalog: An Asteroseismic and Spectroscopic Joint Survey of Targets in the Kepler Fields". *Astrophysical Journal, Supplement*, **215**, 19, 2014.

Planck Collaboration, Ade, P. A. R., Aghanim, N., Arnaud, M., Ashdown, M., et al. "Planck 2015 results. XIII. Cosmological parameters". *Astronomy and Astrophysics*, **594**, A13, 2016.

Poggio, E., Drimmel, R., Smart, R. L., Spagna, A., & Lattanzi, M. G. "The kinematic signature of the Galactic warp in Gaia DR1. I. The Hipparcos subsample". *Astronomy and Astrophysics*, **601**, A115, 2017.

Pohlen, M., Dettmar, R.-J., & Lütticke, R. "Cut-off radii of galactic disks . A new statistical study on the truncation of galactic disks". *Astronomy and Astrophysics*, **357**, L1–L4, 2000.



Portail, M., Wegg, C., Gerhard, O., & Martinez-Valpuesta, I. "Made-to-measure models of the Galactic box/peanut bulge: stellar and total mass in the bulge region". *Monthly Notices of the Royal Astronomical Society*, **448**, 713–731, 2015.

Press, W. H. & Schechter, P. "Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation". *Astrophysical Journal*, **187**, 425–438, 1974.

Price-Whelan, A. M., Johnston, K. V., Sheffield, A. A., Laporte, C. F. P., & Sesar, B. "A reinterpretation of the Triangulum-Andromeda stellar clouds: a population of halo stars kicked out of the Galactic disc". *Monthly Notices of the Royal Astronomical Society*, **452**, 676–685, 2015.

Pringle, J. E. "Accretion discs in astrophysics". *Annual Review of Astronomy and Astrophysics*, **19**, 137–162, 1981.

Prochaska, J. X., Naumov, S. O., Carney, B. W., McWilliam, A., & Wolfe, A. M. "The Galactic Thick Disk Stellar Abundances". *Astronomical Journal*, **120**, 2513–2549, 2000.

Quillen, A. C. "Chaos Caused by Resonance Overlap in the Solar Neighborhood: Spiral Structure at the Bar's Outer Lindblad Resonance". *Astronomical Journal*, **125**, 785–793, 2003.

Quillen, A. C. & Minchev, I. "The Effect of Spiral Structure on the Stellar Velocity Distribution in the Solar Neighborhood". *Astronomical Journal*, **130**, 576–585, 2005.

Quinn, P. J., Hernquist, L., & Fullagar, D. P. "Heating of galactic disks by mergers". *Astrophysical Journal*, **403**, 74–93, 1993.

Raha, N., Sellwood, J. A., James, R. A., & Kahn, F. D. "A dynamical instability of bars in disk galaxies". *Nature*, **352**, 411, 1991.

Randall, L. & Reece, M. "Dark Matter as a Trigger for Periodic Comet Impacts". *Physical Review Letters*, **112**(16), 161301, 2014.

Randich, S., Gilmore, G., & Gaia-ESO Consortium. "The Gaia-ESO Large Public Spectroscopic Survey". *The Messenger*, **154**, 47–49, 2013.

Read, J. I. "The local dark matter density". *Journal of Physics G Nuclear Physics*, **41**(6), 063101, 2014.

Recio-Blanco, A., de Laverny, P., Kordopatis, G., Helmi, A., Hill, V., et al. "The Gaia-ESO Survey: the Galactic thick to thin disc transition". *Astronomy and Astrophysics*, **567**, A5, 2014.

Rees, M. J. & Ostriker, J. P. "Cooling, dynamics and fragmentation of massive gas clouds - Clues to the masses and radii of galaxies and clusters". *Monthly Notices of the Royal Astronomical Society*, **179**, 541–559, 1977.

Reid, M. J., Menten, K. M., Brunthaler, A., Zheng, X. W., Dame, T. M., et al. "Trigonometric Parallaxes of High Mass Star Forming Regions: The Structure and Kinematics of the Milky Way". *Astrophysical Journal*, **783**, 130, 2014.

Reid, M. J., Menten, K. M., Zheng, X. W., Brunthaler, A., Moscadelli, L., et al. "Trigonometric Parallaxes of Massive Star-Forming Regions. VI. Galactic Structure, Fundamental Parameters, and Noncircular Motions". *Astrophysical Journal*, **700**, 137-148, 2009.

Riess, A. G., Filippenko, A. V., Challis, P., Clocchiatti, A., Diercks, A., et al. "Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant". *Astronomical Journal*, **116**, 1009–1038, 1998.

Rix, H.-W. & Bovy, J. "The Milky Way's stellar disk. Mapping and modeling the Galactic disk". *Astronomy and Astrophysics Reviews*, **21**, 61, 2013.

Rix, H.-W., de Zeeuw, P. T., Cretton, N., van der Marel, R. P., & Carollo, C. M. "Dynamical Modeling of Velocity Profiles: The Dark Halo around the Elliptical Galaxy NGC 2434". *Astrophysical Journal*, **488**, 702–719, 1997.

Robin, A. C., Luri, X., Reylé, C., Isasi, Y., Grux, E., et al. "Gaia Universe model snapshot. A statistical analysis of the expected contents of the Gaia catalogue". *Astronomy and Astrophysics*, **543**, A100, 2012.

Rocha-Pinto, H. J., Majewski, S. R., Skrutskie, M. F., & Crane, J. D. "Tracing the Galactic Anticenter Stellar Stream with 2MASS M Giants". *Astrophysical Journal, Letters*, **594**, L115–L118, 2003.

Roškar, R., Debattista, V. P., Quinn, T. R., Stinson, G. S., & Wadsley, J. "Riding the Spiral Waves: Implications of Stellar Migration for the Properties of Galactic Disks". *Astrophysical Journal, Letters*, **684**, L79, 2008.

Rubin, V. C., Ford, W. K., Jr., & Thonnard, N. "Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 /R = 4kpc/ to UGC 2885 /R = 122 kpc/". *Astrophysical Journal*, **238**, 471–487, 1980.

Rubin, V. C., Thonnard, N., & Ford, W. K., Jr. "Extended rotation curves of high-luminosity spiral galaxies. IV - Systematic dynamical properties, SA through SC". *Astrophysical Journal, Letters*, **225**, L107–L111, 1978.

Sackett, P. D. "Does the Milky Way Have a Maximal Disk?" *Astrophysical Journal*, **483**, 103–110, 1997.

Salaris, M. & Girardi, L. "Population effects on the red giant clump absolute magnitude: the K band". *Monthly Notices of the Royal Astronomical Society*, **337**, 332–340, 2002.

Sanders, J. "Angle-action estimation in a general axisymmetric potential". *Monthly Notices of the Royal Astronomical Society*, **426**, 128–139, 2012.

Sanders, J. L. & Binney, J. "Actions, angles and frequencies for numerically integrated orbits". *Monthly Notices of the Royal Astronomical Society*, **441**, 3284–3295, 2014.

Sanders, J. L. & Binney, J. "A fast algorithm for estimating actions in triaxial potentials". *Monthly Notices of the Royal Astronomical Society*, **447**, 2479–2496, 2015a.

Sanders, J. L. & Binney, J. "Extended distribution functions for our Galaxy". *Monthly Notices of the Royal Astronomical Society*, **449**, 3479–3502, 2015b.

Sanders, J. L. & Binney, J. "A review of action estimation methods for galactic dynamics". *Monthly Notices of the Royal Astronomical Society*, **457**, 2107–2121, 2016.

Sanderson, R. E., Helmi, A., & Hogg, D. W. "Action-space Clustering of Tidal Streams to Infer the Galactic Potential". *Astrophysical Journal*, **801**, 98, 2015.

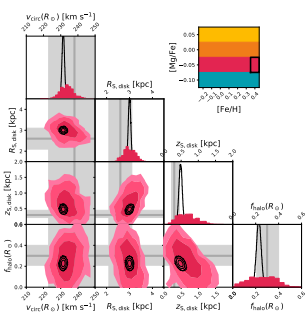
Sato, M., Hirota, T., Reid, M. J., Honma, M., Kobayashi, H., et al. "Distance to G14.33-0.64 in the Sagittarius Spiral Arm: H_2O Maser Trigonometric Parallax with VERA". *Publications of the Astronomical Society of Japan*, **62**, 287–299, 2010.

Schechter, P. "An analytic expression for the luminosity function for galaxies." *Astrophysical Journal*, **203**, 297–306, 1976.

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. "Maps of Dust Infrared Emission for Use in Estimation of Reddening and Cosmic Microwave Background Radiation Foregrounds". *Astrophysical Journal*, **500**, 525–553, 1998.

Schmidt, M. "The Rate of Star Formation." *Astrophysical Journal*, **129**, 243, 1959.

Schödel, R., Feldmeier, A., Kunneriath, D., Stolovy, S., Neumayer, N., et al. "Surface brightness profile of the Milky Way's nuclear star cluster". *Astronomy and Astrophysics*, **566**, A47, 2014.



Schönrich, R. & Aumer, M. "Assessing distances and consistency of kinematics in Gaia/TGAS". *ArXiv e-prints*, 2017.

Schönrich, R. & Binney, J. "Chemical evolution with radial mixing". *Monthly Notices of the Royal Astronomical Society*, **396**, 203–222, 2009a.

Schönrich, R. & Binney, J. "Origin and structure of the Galactic disc(s)". *Monthly Notices of the Royal Astronomical Society*, **399**, 1145–1156, 2009b.

Schönrich, R., Binney, J., & Dehnen, W. "Local kinematics and the local standard of rest". *Monthly Notices of the Royal Astronomical Society*, **403**, 1829–1833, 2010.

Schönrich, R. & McMillan, P. J. "Understanding inverse metallicity gradients in galactic discs as a consequence of inside-out formation". *Monthly Notices of the Royal Astronomical Society*, **467**, 1154–1174, 2017.

Schwarzschild, M. "A numerical model for a triaxial stellar system in dynamical equilibrium". *Astrophysical Journal*, **232**, 236–247, 1979.

SDSS Collaboration, Albareti, F. D., Allende Prieto, C., Almeida, A., Anders, F., et al. "The Thirteenth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-IV Survey MApping Nearby Galaxies at Apache Point Observatory". *ArXiv e-prints*, 2016.

Searle, L. & Zinn, R. "Compositions of halo clusters and the formation of the galactic halo". *Astrophysical Journal*, **225**, 357–379, 1978.

Sellwood, J. A. "A recent Lindblad resonance in the solar neighbourhood". *Monthly Notices of the Royal Astronomical Society*, **409**, 145–155, 2010.

Sellwood, J. A. "The lifetimes of spiral patterns in disc galaxies". *Monthly Notices of the Royal Astronomical Society*, **410**, 1637–1646, 2011.

Sellwood, J. A. "Secular evolution in disk galaxies". *Reviews of Modern Physics*, **86**, 1–46, 2014.

Sellwood, J. A. & Binney, J. J. "Radial mixing in galactic discs". *Monthly Notices of the Royal Astronomical Society*, **336**, 785–796, 2002.

Sellwood, J. A. & Carlberg, R. G. "Transient Spirals as Superposed Instabilities". *Astrophysical Journal*, **785**, 137, 2014.

Sesar, B., Bovy, J., Bernard, E. J., Caldwell, N., Cohen, J. G., et al. "The Nature and Orbit of the Ophiuchus Stream". *Astrophysical Journal*, **809**, 59, 2015.

Shen, J., Rich, R. M., Kormendy, J., Howard, C. D., De Propris, R., et al. "Our Milky Way as a Pure-disk Galaxy - A Challenge for Galaxy Formation". *Astrophysical Journal, Letters*, **720**, L72–L76, 2010.

Shu, F. H. "Models of Partially Relaxed Stellar Disks". *Astrophysical Journal*, **158**, 505, 1969.

Siebert, A., Bienaymé, O., & Soubiran, C. "Vertical distribution of Galactic disk stars. II. The surface mass density in the Galactic plane". *Astronomy and Astrophysics*, **399**, 531–541, 2003.

Siebert, A., Famaey, B., Binney, J., Burnett, B., Faure, C., et al. "The properties of the local spiral arms from RAVE data: two-dimensional density wave approach". *Monthly Notices of the Royal Astronomical Society*, **425**, 2335–2342, 2012.

Siebert, A., Williams, M. E. K., Siviero, A., Reid, W., Boeche, C., et al. "The RAdial Velocity Experiment (RAVE): Third Data Release". *Astronomical Journal*, **141**, 187, 2011.

Silk, J. "Cosmic Black-Body Radiation and Galaxy Formation". *Astrophysical Journal*, **151**, 459, 1968.

Skrutskie, M. F., Cutri, R. M., Stiening, R., Weinberg, M. D., Schneider, S., et al. "The Two Micron All Sky Survey (2MASS)". *Astronomical Journal*, **131**, 1163–1183, 2006.

Smith, R., Flynn, C., Candlish, G. N., Fellhauer, M., & Gibson, B. K. "Simple and accurate modelling of the gravitational potential produced by thick and thin exponential discs". *Monthly Notices of the Royal Astronomical Society*, **448**, 2934–2940, 2015.

Smoot, G. F., Bennett, C. L., Kogut, A., Wright, E. L., Aymon, J., et al. “Structure in the COBE differential microwave radiometer first-year maps”. *Astrophysical Journal, Letters*, **396**, L1–L5, 1992.

Soderblom, D. R. “The Ages of Stars”. *Annual Review of Astronomy and Astrophysics*, **48**, 581–629, 2010.

Solway, M., Sellwood, J. A., & Schönrich, R. “Radial migration in galactic thick discs”. *Monthly Notices of the Royal Astronomical Society*, **422**, 1363–1383, 2012.

Somerville, R. S. & Davé, R. “Physical Models of Galaxy Formation in a Cosmological Framework”. *Annual Review of Astronomy and Astrophysics*, **53**, 51–113, 2015.

Somerville, R. S., Hopkins, P. F., Cox, T. J., Robertson, B. E., & Hernquist, L. “A semi-analytic model for the co-evolution of galaxies, black holes and active galactic nuclei”. *Monthly Notices of the Royal Astronomical Society*, **391**, 481–506, 2008.

Spergel, D. N., Bean, R., Doré, O., Nolta, M. R., Bennett, C. L., et al. “Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology”. *Astrophysical Journal, Supplement*, **170**, 377–408, 2007.

Spitzer, L., Jr. “The Dynamics of the Interstellar Medium. III. Galactic Distribution.” *Astrophysical Journal*, **95**, 329, 1942.

Springel, V. “The cosmological simulation code GADGET-2”. *Monthly Notices of the Royal Astronomical Society*, **364**, 1105–1134, 2005.

Springel, V., Di Matteo, T., & Hernquist, L. “Modelling feedback from stars and black holes in galaxy mergers”. *Monthly Notices of the Royal Astronomical Society*, **361**, 776–794, 2005.

Springel, V., Frenk, C. S., & White, S. D. M. “The large-scale structure of the Universe”. *Nature*, **440**, 1137–1144, 2006.

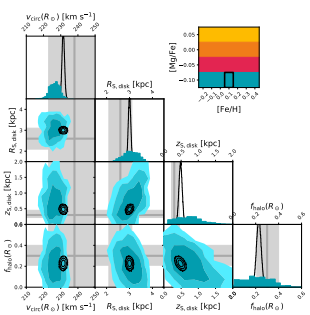
Steinmetz, M., Zwitter, T., Siebert, A., Watson, F. G., Freeman, K. C., et al. “The Radial Velocity Experiment (RAVE): First Data Release”. *Astronomical Journal*, **132**, 1645–1668, 2006.

Stello, D., Huber, D., Bedding, T. R., Benomar, O., Bildsten, L., et al. “Asteroseismic Classification of Stellar Populations among 13,000 Red Giants Observed by Kepler”. *Astrophysical Journal, Letters*, **765**, L41, 2013.

Stoehr, F., White, S. D. M., Tormen, G., & Springel, V. “The satellite population of the Milky Way in a ΛCDM universe”. *Monthly Notices of the Royal Astronomical Society*, **335**, L84–L88, 2002.

Strigari, L. E. “Galactic searches for dark matter”. *Physics Reports*, **531**, 1–88, 2013.

Su, M. & Finkbeiner, D. P. “Double Gamma-ray Lines from Unassociated Fermi-LAT Sources”. *ArXiv e-prints*, 2012.



Syer, D. & Tremaine, S. “Made-to-measure N-body systems”. *Monthly Notices of the Royal Astronomical Society*, **282**, 223–233, 1996.

The RAVE database. “RAVE âĂŞ the Radial Velocity Experiment”. `https://www.rave-survey.org/project/`, 2014. Accessed: 2017-05-15.

Ting, Y.-S., Conroy, C., & Rix, H.-W. “APOGEE Chemical Tagging Constraint on the Maximum Star Cluster Mass in the Alpha-enhanced Galactic Disk”. *Astrophysical Journal*, **816**, 10, 2016.

Ting, Y.-S., Rix, H.-W., Bovy, J., & van de Ven, G. "Constraining the Galactic potential via action-based distribution functions for mono-abundance stellar populations". *Monthly Notices of the Royal Astronomical Society*, **434**, 652–660, 2013.

Toomre, A. & Toomre, J. "Galactic Bridges and Tails". *Astrophysical Journal*, **178**, 623–666, 1972.

Tremaine, S., Gebhardt, K., Bender, R., Bower, G., Dressler, A., et al. "The Slope of the Black Hole Mass versus Velocity Dispersion Correlation". *Astrophysical Journal*, **574**, 740–753, 2002.

Trick, W. H., Bovy, J., D'Onghia, E., & Rix, H.-W. "Action-based Dynamical Modeling for the Milky Way Disk: The Influence of Spiral Arms". *Astrophysical Journal*, **839**, 61, 2017.

Trick, W. H., Bovy, J., & Rix, H.-W. "Action-Based Dynamical Modeling for the Milky Way Disk". *Astrophysical Journal*, **830**, 97, 2016a.

Trick, W. H., van de Ven, G., & Dutton, A. A. "A spiral galaxy's mass distribution uncovered through lensing and dynamics". *Monthly Notices of the Royal Astronomical Society*, **463**, 3151–3168, 2016b.

Tully, R. B. "Nearby groups of galaxies. II - an all-sky survey within 3000 kilometers per second". *Astrophysical Journal*, **321**, 280–304, 1987.

Tully, R. B. & Fisher, J. R. "A new method of determining distances to galaxies". *Astronomy and Astrophysics*, **54**, 661–673, 1977.

Turon, C., Primas, F., Binney, J., Chiappini, C., Drew, J., et al. "Galactic Populations, Chemistry and Dynamics". Technical report, 2008.

Tyson, J. A., Kochanski, G. P., & Dell'Antonio, I. P. "Detailed Mass Map of CL 0024+1654 from Strong Lensing". *Astrophysical Journal, Letters*, **498**, L107–L110, 1998.

Tyson, J. A., Valdes, F., & Wenk, R. A. "Detection of systematic gravitational lens galaxy image alignments - Mapping dark matter in galaxy clusters". *Astrophysical Journal, Letters*, **349**, L1–L4, 1990.

Urquhart, J. S., Figura, C. C., Moore, T. J. T., Hoare, M. G., Lumsden, S. L., et al. "The RMS survey: galactic distribution of massive star formation". *Monthly Notices of the Royal Astronomical Society*, **437**, 1791–1807, 2014.

Vallée, J. P. "New Velocimetry and Revised Cartography of the Spiral Arms in the Milky Way—A Consistent Symbiosis". *Astronomical Journal*, **135**, 1301-1310, 2008.

Vallée, J. P. "The Spiral Arms of the Milky Way: The Relative Location of Each Different Arm Tracer within a Typical Spiral Arm Width". *Astronomical Journal*, **148**, 5, 2014.

van Albada, T. S. & Sancisi, R. "Dark matter in spiral galaxies". *Philosophical Transactions of the Royal Society of London Series A*, **320**, 447–464, 1986.

van de Ven, G., de Zeeuw, P. T., & van den Bosch, R. C. E. "Recovery of the internal orbital structure of galaxies". *Monthly Notices of the Royal Astronomical Society*, **385**, 614–646, 2008.

van der Kruit, P. C. & Freeman, K. C. "Galaxy Disks". *Annual Review of Astronomy and Astrophysics*, **49**, 301–371, 2011.

van Helshoecht, V. & Groenewegen, M. A. T. "K-band magnitude of the red clump as a distance indicator". *Astronomy and Astrophysics*, **463**, 559–565, 2007.

van Leeuwen, F. "Validation of the new Hipparcos reduction". *Astronomy and Astrophysics*, **474**, 653–664, 2007.

van Leeuwen, F., Evans, D. W., De Angeli, F., Jordi, C., Busso, G., et al. "Gaia Data Release 1. The photometric data". *Astronomy and Astrophysics*, **599**, A32, 2017.

Vera-Ciro, C. & D'Onghia, E. "On the Conservation of the Vertical Action in Galactic Disks". *Astrophysical Journal*, **824**, 39, 2016.

Villalobos, Á. & Helmi, A. "Simulations of minor mergers - I. General properties of thick discs". *Monthly Notices of the Royal Astronomical Society*, **391**, 1806–1827, 2008.

Wagoner, R. V., Fowler, W. A., & Hoyle, F. "On the Synthesis of Elements at Very High Temperatures". *Astrophysical Journal*, **148**, 3, 1967.

Walker, I. R., Mihos, J. C., & Hernquist, L. "Quantifying the Fragility of Galactic Disks in Minor Mergers". *Astrophysical Journal*, **460**, 121, 1996.

Walker, T. P., Steigman, G., Kang, H.-S., Schramm, D. M., & Olive, K. A. "Primordial nucleosynthesis redux". *Astrophysical Journal*, **376**, 51–69, 1991.

Wallerstein, G. "Abundances in G. Dwarfs.VI. a Survey of Field Stars." *Astrophysical Journal, Supplement*, **6**, 407, 1962.

Wegg, C. & Gerhard, O. "Mapping the three-dimensional density of the Galactic bulge with VVV red clump stars". *Monthly Notices of the Royal Astronomical Society*, **435**, 1874–1887, 2013.

Wegg, C., Gerhard, O., & Portail, M. "The structure of the Milky Way's bar outside the bulge". *Monthly Notices of the Royal Astronomical Society*, **450**, 4050–4069, 2015.

Weiland, J. L., Arendt, R. G., Berriman, G. B., Dwek, E., Freudenreich, H. T., et al. "COBE diffuse infrared background experiment observations of the galactic bulge". *Astrophysical Journal, Letters*, **425**, L81–L84, 1994.

Weinberg, D. H., Bullock, J. S., Governato, F., Kuzio de Naray, R., & Peter, A. H. G. "Cold dark matter: Controversies on small scales". *Proceedings of the National Academy of Science*, **112**, 12249–12255, 2015.

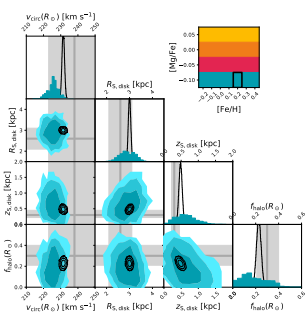
White, S. D. M., Davis, M., & Frenk, C. S. "The size of clusters in a neutrino-dominated universe". *Monthly Notices of the Royal Astronomical Society*, **209**, 27P–31P, 1984.

White, S. D. M. & Frenk, C. S. "Galaxy formation through hierarchical clustering". *Astrophysical Journal*, **379**, 52–79, 1991.

White, S. D. M. & Rees, M. J. "Core condensation in heavy halos - A two-stage theory for galaxy formation and clustering". *Monthly Notices of the Royal Astronomical Society*, **183**, 341–358, 1978.

Wojno, J., Kordopatis, G., Piffl, T., Binney, J., Steinmetz, M., et al. "The selection function of the RAVE survey". *Monthly Notices of the Royal Astronomical Society*, **468**, 3368–3380, 2017.

Wojno, J., Kordopatis, G., Steinmetz, M., McMillan, P., Matijevič, G., et al. "Chemical separation of disc components using RAVE". *Monthly Notices of the Royal Astronomical Society*, **461**, 4246–4255, 2016.



Woosley, S. E. & Weaver, T. A. "The Evolution and Explosion of Massive Stars. II. Explosive Hydrodynamics and Nucleosynthesis". *Astrophysical Journal, Supplement*, **101**, 181, 1995.

Wyse, R. "Galactic Surveys in the Gaia Era", 2017. IAU Symposium 330: Astrometry and Astrophysics in the Gaia Sky.

Wyse, R. F. G. & Gilmore, G. "Formation and evolution of the Galactic bulge and spheroid - Where did the spheroid gas go?" *Astronomical Journal*, **104**, 144–153, 1992.

Wyse, R. F. G. & Gilmore, G. "Chemical and Dynamical Evolution of the Galaxy". In "The Globular Cluster-Galaxy Connection", , edited by Smith, G. H. & Brodie, J. P., volume 48 of *Astronomical Society of the Pacific Conference Series*, p. 727, 1993.

Wyse, R. F. G., Gilmore, G., & Franx, M. "Galactic Bulges". *Annual Review of Astronomy and Astrophysics*, **35**, 637–675, 1997.

Xu, Y., Li, J. J., Reid, M. J., Menten, K. M., Zheng, X. W., et al. "On the Nature of the Local Spiral Arm of the Milky Way". *Astrophysical Journal*, **769**, 15, 2013.

Xu, Y., Newberg, H. J., Carlin, J. L., Liu, C., Deng, L., et al. "Rings and Radial Waves in the Disk of the Milky Way". *Astrophysical Journal*, **801**, 105, 2015.

Xu, Y., Reid, M. J., Zheng, X. W., & Menten, K. M. "The Distance to the Perseus Spiral Arm in the Milky Way". *Science*, **311**, 54–57, 2006.

Yanny, B., Rockosi, C., Newberg, H. J., Knapp, G. R., Adelman-McCarthy, J. K., et al. "SEGUE: A Spectroscopic Survey of 240,000 Stars with g = 14-20". *Astronomical Journal*, **137**, 4377-4399, 2009.

Yoachim, P. & Dalcanton, J. J. "Structural Parameters of Thin and Thick Disks in Edge-on Disk Galaxies". *Astronomical Journal*, **131**, 226–249, 2006.

Yoachim, P. & Dalcanton, J. J. "The Kinematics of Thick Disks in Nine External Galaxies". *Astrophysical Journal*, **682**, 1004-1019, 2008.

Zasowski, G., Johnson, J. A., Frinchaboy, P. M., Majewski, S. R., Nidever, D. L., et al. "Target Selection for the Apache Point Observatory Galactic Evolution Experiment (APOGEE)". *Astronomical Journal*, **146**, 81, 2013.

Zel'dovich, Y. B. "Gravitational instability: An approximate theory for large density perturbations." *Astronomy and Astrophysics*, **5**, 84–89, 1970.

Zhang, B., Reid, M. J., Menten, K. M., Zheng, X. W., Brunthaler, A., et al. "Parallaxes for W49N and G048.60+0.02: Distant Star Forming Regions in the Perseus Spiral Arm". *Astrophysical Journal*, **775**, 79, 2013a.

Zhang, L., Rix, H.-W., van de Ven, G., Bovy, J., Liu, C., et al. "The Gravitational Potential near the Sun from SEGUE K-dwarf Kinematics". *Astrophysical Journal*, **772**, 108, 2013b.

Zhu, L., van den Bosch, R. C. E., van de Ven, G., Falcón-Barroso, J., Lyubenova, M., et al. "Dynamical decomposition of galaxies across the Hubble sequence". In "The Interplay between Local and Global Processes in Galaxies,", , 2016.

Zibetti, S., White, S. D. M., & Brinkmann, J. "Haloes around edge-on disc galaxies in the Sloan Digital Sky Survey". *Monthly Notices of the Royal Astronomical Society*, **347**, 556–568, 2004.

Zoccali, M., Renzini, A., Ortolani, S., Greggio, L., Saviane, I., et al. "Age and metallicity distribution of the Galactic bulge from extensive optical and near-IR stellar photometry". *Astronomy and Astrophysics*, **399**, 931–956, 2003.

Zwicky, F. "Die Rotverschiebung von extragalaktischen Nebeln". *Helvetica Physica Acta*, **6**, 110–127, 1933.

Zwitter, T., Siebert, A., Munari, U., Freeman, K. C., Siviero, A., et al. "The Radial Velocity Experiment (RAVE): Second Data Release". *Astronomical Journal*, **136**, 421–451, 2008.

Statement

This thesis is my own work and I have only used the sources indicated. Where the work of others has been quoted or reproduced, the source is always given.

Heidelberg, September 18, 2017 .....