Galactic substructures as tracers of dark matter and stellar evolution

Reino, S.

Chapter 1

Introduction

Astronomy is unique among most other scientific fields due to the route which we must take to unearth the principles that govern our objects of study. The timelines over which most astronomical systems evolve are too great to allow for direct observation of any changes. Observations of single objects offer us a snapshot of the state of the system at a unique time, but are not sufficient to reveal any past or future evolution and, therefore, deep understanding of the object as a whole. Progress is instead made by observing a multitude of similar objects that are currently at a variety of evolutionary stages, and using the full inventory of these snapshots to reconstruct the story of their origin and evolution.

For example, to reveal how galaxies form and evolve it is necessary to observe a whole ensemble of galaxies at different distances, and therefore look-back times, and map their relation to each other. Piecing together the evolutionary sequence requires developing models and numerical simulations that can be tested against the whole compilation of observed snapshots, and thus theoretical insight is created without the need of witnessing changes in any individual galaxies.

Two further examples are especially relevant in the context of my work. Our quest to measure the gravitational potential of our Galaxy would be much simplified if we could follow stars over long sections of their orbits around the Galaxy and witness the changes in their velocities over this path. These accelerations directly relate to the gravitational field in which the stars move. However, the timescales for orbital periods are extremely long and typical accelerations of stars are extremely small, estimated to be of the order of a few cm s^{-1} velocity change over a decade (Silverwood & Easther 2019; Chakrabarti et al. 2020). Fortunately, there exist structures which allow us to deduce the accelerations that specific stars undergo, forgoing the need for direct measurement. These structures are called stellar streams (Section 1.1.2) and they form when globular clusters or dwarf galaxies are torn apart by Galactic tidal forces, resulting in the formation of long thin streams of stars which follow a similar orbit. In an idealized way, if all their energies and angular momenta were identical, then the stars in the stream could be viewed as snapshots of a single star observed at different stages of their orbit, which allows us to derive the acceleration the star must be experiencing and, therefore, the Galactic gravitational potential. So, even though we cannot follow the path of individual stars over a significant length of their orbit, we can use stream stars to trace out an implicit orbit.

In a similar manner, we cannot study the evolutionary path of a star by noting the changes it undergoes throughout its life. Instead we must rely on observing a collection of stars at distinct evolutionary stages to put together a coherent theory of stellar evolution. However, these theories are complex to test using the general population of stars because different stellar properties and observables are difficult to control for. For example, the main predictor of a star's evolution is its mass, which is not straightforward to infer without knowing the distance with which to calibrate its measured brightness. Furthermore, chemical composition and age of the star can affect the comparisons. Open clusters (Section 1.1.1) provide an excellent setting for stellar evolution studies since many of its stellar properties are fixed: the cluster stars formed at the same time from the same molecular cloud, therefore all share the same age and chemical composition. In addition, the distance to each star in the cluster is constant, meaning their relative brightness is caused solely by their differences in intrinsic luminosity, and thus, mass. Open clusters therefore offer us a snapshot of a collection of stars that is ideal for developing models of stellar evolution and stellar physics.

Knowledge about topics such as galaxy formation and evolution, galactic mass distribution, and stellar evolution and physics is fundamental in understanding the universe around us. Studies of our own Galaxy, the Milky Way, are crucial for developing this knowledge for it is the Milky Way that offers us a detailed three-dimensional internal view of a galaxy, its rich information content being at our fingertips. Furthermore, Milky Way is the only galaxy whose stellar content we can observe and measure with a precision that prompts the greatest advances in these, and many other, fields. The implications of these local measurements, however, reach beyond just the Milky Way, helping us place constraints on the larger Universe and physics. Since the Milky Way forms the laboratory in which we have conducted our research, we will begin by giving an overview of the Galaxy and its components.

1.1 The structure of the Galaxy

The Milky Way is a typical size galaxy in the local Universe with a stellar mass $\sim 10^{11} M_{\odot}$ and a total mass roughly an order of magnitude larger (for more discussion on this point see Section 1.2. It consists of several distinct components, which can be distinguished based on their position in the Galaxy, spatial distribution, kinematic signature, ages and chemical composition. Here, we will give a brief overview of each constituent and describe their most important characteristics. A conceptual picture of the structure of the Milky Way is shown in Figure 1.1.

Thin disc. The disc is the most distinctive feature of the Milky Way, containing the bulk of its stellar content. Two disc-shaped components are usually distinguished (Gilmore & Reid 1983): the thin and the thick disc. Situated on the mid-plane of the Galaxy, the thin disc is a rotationally supported structure with most of its stars on approximately circular orbits. Although the structure extends at least 20 - 25 kpc from the Galactic centre in the mid-plane (Carraro et al. 2010; Liu et al. 2017; López-Corredoira et al. 2018), it is extremely flat in the vertical direction: its shape is typically described with an exponential distribution with a scale height of about 300 pc and a scale length of about 2.6 kpc (Jurić et al. 2008). Although, the majority of the disc is flat, the outer edges of the disc are warped with a twist downwards on one side of the Galaxy and upwards on the other (Kerr 1957). Since, alongside stars, the thin disc also contains an abundant supply of gas and dust, it is the most active region of star formation in the Milky Way.

Thick disc. There are several prominent features that are commonly used to distinguish the thick disc from the thin disc. The thick disc, as the name suggests, reaches further from the mid-plane with a scale height of around 900 pc (Jurić et al. 2008). Its stellar population is older, more metal-poor and α-element rich compared to the thin disc (e.g. Haywood et al. 2013; Hayden et al. 2015). Although the thick disc is kinematically hotter than the thin disc, the thick disc stars trail behind those of the thin disc in rotation speeds (Haywood et al. 2013; Allende Prieto et al. 2016). Evidence currently points to the thick disc having formed via a merger with the Gaia-Enceladus dwarf galaxy (Helmi et al. 2018) which dynamically heated the stars from the already existing proto-disc and, due to the infalling gas from the merger, triggered a burst of star formation that would further populate the thick disc (Gallart et al. 2019; Helmi 2020).

Spiral arms. Embedded in the thin disc of the Milky Way are the spiral arms, principally consisting of molecular clouds and young stars. Although the specific structure of the spiral arms has long been a puzzle, recent mea-

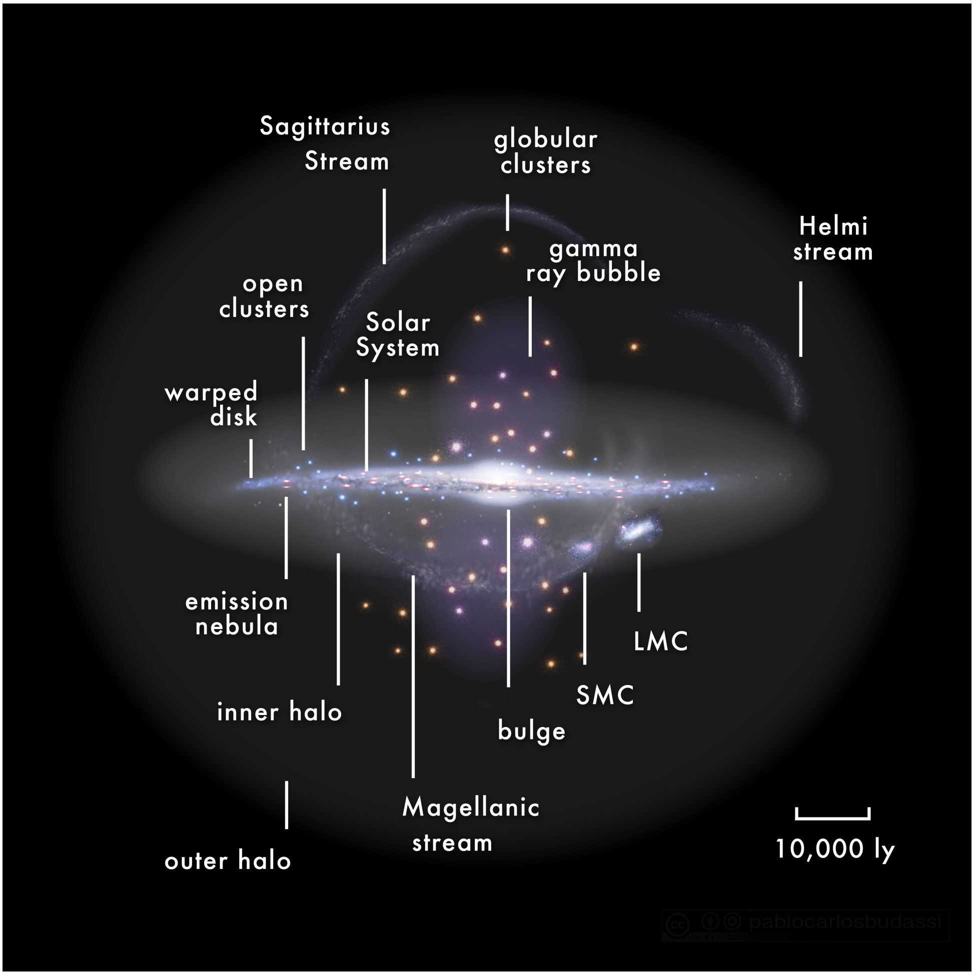

Figure 1.1: Artist impression of the structure of the Milky Way. The positions of the components are not to scale. Image credit: Pablo Carlos Budassi.

surements reveal evidence that the Milky Way contains four spiral arms, some additional segments (such as the Local arm, closest to the Sun) and a 3-kpc arms surrounding the central bar in a ring-shape (Reid et al. 2019). Similarly, the nature and formation of the spiral arms have been the source of a long-standing debate. The main questions relate to whether the spiral arms are material – corotating with the disc at all radii – or density waves – regions of higher density which rotate at a speed that is different from the disc; whether they are long-lived or transient and recurring; and whether they form due to small scale instabilities or tidal interactions (e.g. Lin &

Shu 1964; Toomre 1969; Sellwood & Carlberg 1984 and a review by Dobbs & Baba 2014).

Bulge. The Galactic bulge contains some of the oldest stars in the Milky Way (> 10 Gyr, Zoccali et al. 2003) and is concentrated within 2-3 kpc of the Galactic centre. The "classical" bulge is a spheroidal, pressure supported structure reminiscent of elliptical galaxies considered to be formed though galaxy mergers (Kormendy & Kennicutt 2004). However, the existence of this classical bulge component in the Milky Way is debated since the majority of the bulge stars unambiguously exhibit a rotating boxy/peanut shape instead (e.g. McWilliam & Zoccali 2010; Shen et al. 2010; Wegg & Gerhard 2013). Compared to the classical bulge, the stars in the box/peanut bulge are supported by their near-cylindrical rotation (Howard et al. 2009; Ness et al. 2013). The box/peanut has been shown to be a manifestation of the longer, more planar bar which in the inner regions has grown vertically thick due to buckling instabilities (e.g. Athanassoula 2005; Shen et al. 2010).

Bar. An elongated bar, which rotates as a solid body, can develop near the centres of galaxies naturally due to instabilities in the thin disc (Athanassoula 2013) or interactions with external perturbers (e.g. Peschken & Łokas 2019). Although, as discussed above, the bulge and the bar of the Milky Way are now known not to be physically distinct, we can still distinguish the long bar: a thin structure that reaches out of the box/peanut bulge and has a half length of about 5 kpc (Wegg et al. 2015). The current estimates of the angle between the long axis of the bar and the Sun-Galactic center line vary between $25-30°$, while the pattern speed of the bar has been established as between $35-40$ km s^{-1} kpc^{-1} (Shen & Zheng 2020, and the references therein).

Stellar halo. Stellar halo, although centrally concentrated, is the most extended stellar component of the Galaxy (> 100 kpc). It is a spheroidal structure (the specific shape of which varies with distance from the Galactic centre, Kinman et al. 1966) supported mainly by velocity dispersion (Chiba & Beers 2000, e.g.), its content dominated by metal-poor and very old stars. It contains only about 1% of the stellar mass of the Milky Way and is therefore much less dense and luminous compared to the disc and the other central parts of the Galaxy (Freeman & Bland-Hawthorn 2002). It is currently understood that the stellar halo forms through repeated accretion of smaller dwarf galaxies (Searle & Zinn 1978), a process which embeds the halo with a large amount of substructure (see Section 1.1.2), although a fraction of stars could also have an in situ origin.

Dark halo. The stellar halo, and the rest of the stellar content of the Galaxy, is embedded in a dark matter halo. Although the dark halo accounts for the largest fraction of the total Galactic mass, our knowledge about it is incomplete. The determination of the mass, profile and shape of the dark halo is currently under intense research due to its importance not only in understanding our own Galaxy but also providing insight into more general cosmological questions (see Section 1.2).

Within these larger structures there also exists a significant amount of substructure such as star clusters, dwarf galaxies, stellar streams and other tidal debris. Star clusters are gravitationally bound stellar systems whose stars were likely born concurrently in the same gas cloud and can be divided into two distinct categories: globular clusters and open clusters. **Globular clusters** are massive, spherical systems with old populations of stars (Gratton et al. 2019; Beasley 2020). There are currently approximately 150 known globular clusters in the Milky Way. They are nearly spherically distributed in the halo region of the Galaxy and, although their origin remains somewhat of a puzzle, many of them are believed to be of extragalactic origin having been accreted into the Galactic halo (Massari et al. 2019). **Open clusters**, instead, are found mainly in the thin disc of the Galaxy, contain very young stars and are, compared to globular clusters, more irregularly shaped and less massive. **Dwarf galaxies**, as their name suggests, are small galaxies with a stellar mass of $\lesssim 10^9\ M_\odot$ (Bullock & Boylan-Kolchin 2017). Although they are fainter and harder to detect than their more luminous counterparts, they constitute the majority of galaxies in the Universe. In the context of the Milky Way, they are often synonymous with “satellite galaxies” which are in the process of being accreted by the Milky Way. There are approximately 50 known dwarf galaxies in the Milky Way (McConnachie 2012; Newton et al. 2018; Putman et al. 2021), the most prominent of these are the Large and Small Magellanic Clouds (LMC and SMC). **Stellar streams** are the remnants of globular clusters and dwarf galaxies which were disrupted by the tidal forces of the Galactic potential. A more detailed discussion of both open clusters and stellar streams is given in the next two sections.

1.1.1 Open clusters

Open clusters are loosely bound stellar groups comprising of young stars mostly located on the plane of the Galaxy. Currently, over 2000 open clusters have been identified in the Galaxy (Kharchenko et al. 2013; Cantat-Gaudin et al. 2018), with a handful of them, such as the Hyades and Pleiades clusters, also visible to the naked eye. New open clusters are continuously forming in the thin disc, but since these structures are easily disrupted by external forces, before long they will have dissolved and fed their stellar content to the bulk of the disc. In fact, it is generally thought that most of the stars in the Galactic disc likely formed in open clusters that have dispersed (Binney & Tremaine 2008).

The stars in open clusters were formed at the same time from the same molecular cloud, thus all members have identical ages, chemical composition, distances and space velocities, their only differences stemming from their varied initial masses. The homogeneity of these properties makes open clusters invaluable in developing and testing stellar evolution models and understanding stellar interiors and atmospheres.

Furthermore, the properties of open clusters can be determined with great precision: their colour–magnitude diagrams can be fitted with isochrone models to simultaneously constrain their distance, reddening, age and metallicity beyond the precision attainable for the general population of stars. Due to this, open clusters are widely considered as a key class of objects for exploring the structure and dynamics of the Galaxy. For example, tracing the galactic spiral structure (Castro-Ginard et al. 2021), measuring the pattern speed of the spiral pattern (Dias et al. 2019), investigating radial migration in the disc (Chen & Zhao 2020), tracing the Galactic warp (Cantat-Gaudin et al. 2020), etc.

A prerequisite for all these studies is the accurate membership determination for the cluster. The remarkable wealth of data from ESA's *Gaia* mission (Gaia Collaboration et al. 2016a), delivering proper motions and parallaxes for more than 1 billion stars in the Milky Way, enables not only the identification of many new open clusters (Castro-Ginard et al. 2020), but also unprecedented precision in both the membership determination and parameter derivation (Cantat-Gaudin & Anders 2020). The release of this new data set also prompted the work presented in Chapter 5, where we derive the membership of and, subsequently, characterize the Hyades open cluster.

1.1.2 Stellar streams and halo substructure

The currently accepted Λ cold dark matter (ΛCDM) model prescribes that structure formation proceeds in a hierarchical manner whereby galaxies grow via mergers with smaller satellite galaxies (e.g. Press & Schechter 1974; White & Rees 1978). As dwarf galaxies fall into the potential well of the Milky Way, they experience Galactic tidal forces which can lead to their disruption. If, for a star in the progenitor structure, the potential of the Milky Way begins to dominate over the self-gravity of the progenitor, the star will be stripped from the original structure, the most likely escape route being through the two Lagrange points (Küpper et al. 2012). As the stars leave the progenitor they form two nearly symmetrical tails around the cluster: the trailing tail is formed from stars with slightly higher energy and longer orbital periods compared to the progenitor, while the leading tail is formed from lower energy stars with shorter orbital periods (Binney & Tremaine 2008). Other than these small offsets, the orbits of these stars are closely aligned with that of the progenitor (Johnston et al. 1996; Binney 2008; Eyre & Binney 2011). Their differential orbital periods serve to gradually stretch the stripped stars out on their neighbouring orbits forming an elongated filament called the stellar stream. An illustration of the stream shape is presented in Figure 1.2. The trajectories of stream stars are mostly determined by the Galactic potential, allowing us to view them as test masses and making them valuable probes of the underlying Galactic mass profile (Johnston et al. 1999).

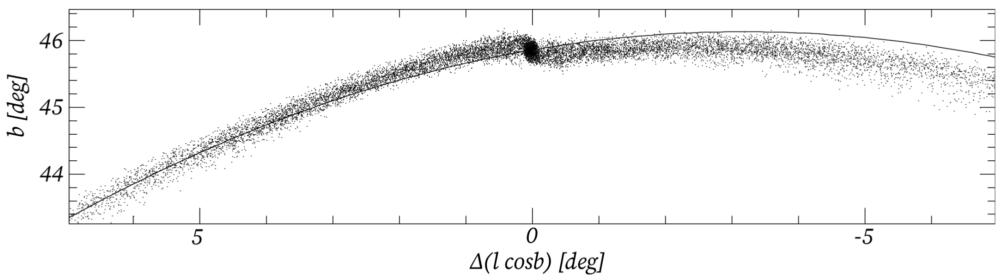

Figure 1.2: A model of a stream in Galactic coordinates, with the orbit of the progenitor shown as a black line. Image credit: Dehnen et al. (2004).

The stream formation process doesn't exclusively impact accreted dwarf galaxies but also, as mentioned above, globular clusters. The main difference between globular cluster and dwarf galaxy streams is their relative widths and velocity dispersions: globular clusters occupy a smaller phase space volume compared to dwarf galaxies and thus their streams are thinner and dynamically colder, while dwarf galaxies streams are broader and hotter.

After some time, the progenitor of the stream is completely depleted of material and ceases to exist. The stellar stream, however, carries on evolving. At the earlier stages, when the stream is still a spatially coherent structure, bearing its characteristic appearance of a thin arc of stars, we refer to it as a **coherent stream**. As the stream stars keep stretching out along their orbits, lengthening and diffusing the stream, in due course the stream will have several "wraps" around the Galaxy, its stars occupying the same phase multiple times. Eventually, the stream reaches a **phase-mixed** state, where it is no longer spatially recognisable as a distinct structure (Helmi et al. 1999;

Helmi 2008). However, due to long dynamical timescales in the halo region, the stars in these phase-mixed streams nevertheless retain memory of their original orbit and can still be distinguished in phase space or with integrals of motion.

The majority of the stellar halo is thought to have been built up over time through the processes of accretion and subsequent tidal disruption. The stellar halo can therefore be viewed as superposition of a large amount of stellar streams, both coherent and phase mixed, of globular cluster and dwarf galaxy origin.

The first detections of stellar streams came when Ibata et al. (1994) reported the discovery of the tidally disrupted Sagittarius dwarf galaxy and Helmi et al. (1999) identified the debris of the Helmi streams in angular momentum space. Since then, nearly a hundred coherent stellar streams have been discovered in the Milky Way – a record of these is kept in the galstreams package[1] (although note that phase-mixed streams are not included, see Mateu 2022 for more information) – due to surveys such as SDSS (The "Field of Streams", Belokurov et al. 2006), PAndAS (Martin et al. 2014), Pan-STARRS1 (Bernard et al. 2016), DES (Shipp et al. 2018), *Gaia* (Ibata et al. 2021), S^5 (Li et al. 2022) etc. Besides the Sagittarius and Helmi streams, some of the other prominent earlier discoveries included the GD-1 stream (Grillmair & Dionatos 2006b), Pal 5 stream (Odenkirchen et al. 2001), NGC 5466 stream (Grillmair 2006), Orphan stream (Grillmair 2006; Belokurov et al. 2006), Atlas stream (Koposov et al. 2014), Ophiuchus stream (Bernard et al. 2014) and the Cetus stream (Newberg et al. 2009). Figure 1.3 presents the overlapping tracks of all currently known coherent streams.

Although many of these early discoveries were related to coherent stellar streams – these can be detected easier using star counts relative to field stars – as the six-dimensional phase space maps and chemical abundances have become available for a larger fraction of stars, more and more phase-mixed streams and substructures have been identified through kinematics and chemistry. One of the most important recent discoveries, made possible by the unprecedented data from the *Gaia* satellite (Gaia Collaboration et al. 2018), was the presence of a significant amount of debris from what appears to be Milky Way's last major merger with a satellite dubbed Gaia-Enceladus (Helmi et al. 2018) or Gaia-Sausage (Belokurov et al. 2018), and often referred to simply as GSE. This merger was likely a head-on collision between the Milky Way and the GSE about 10 Gyr ago, with mass ratios of 4:1, and has been hypothesized to be responsible for the formation of the

[1]https://github.com/cmateu/galstreams

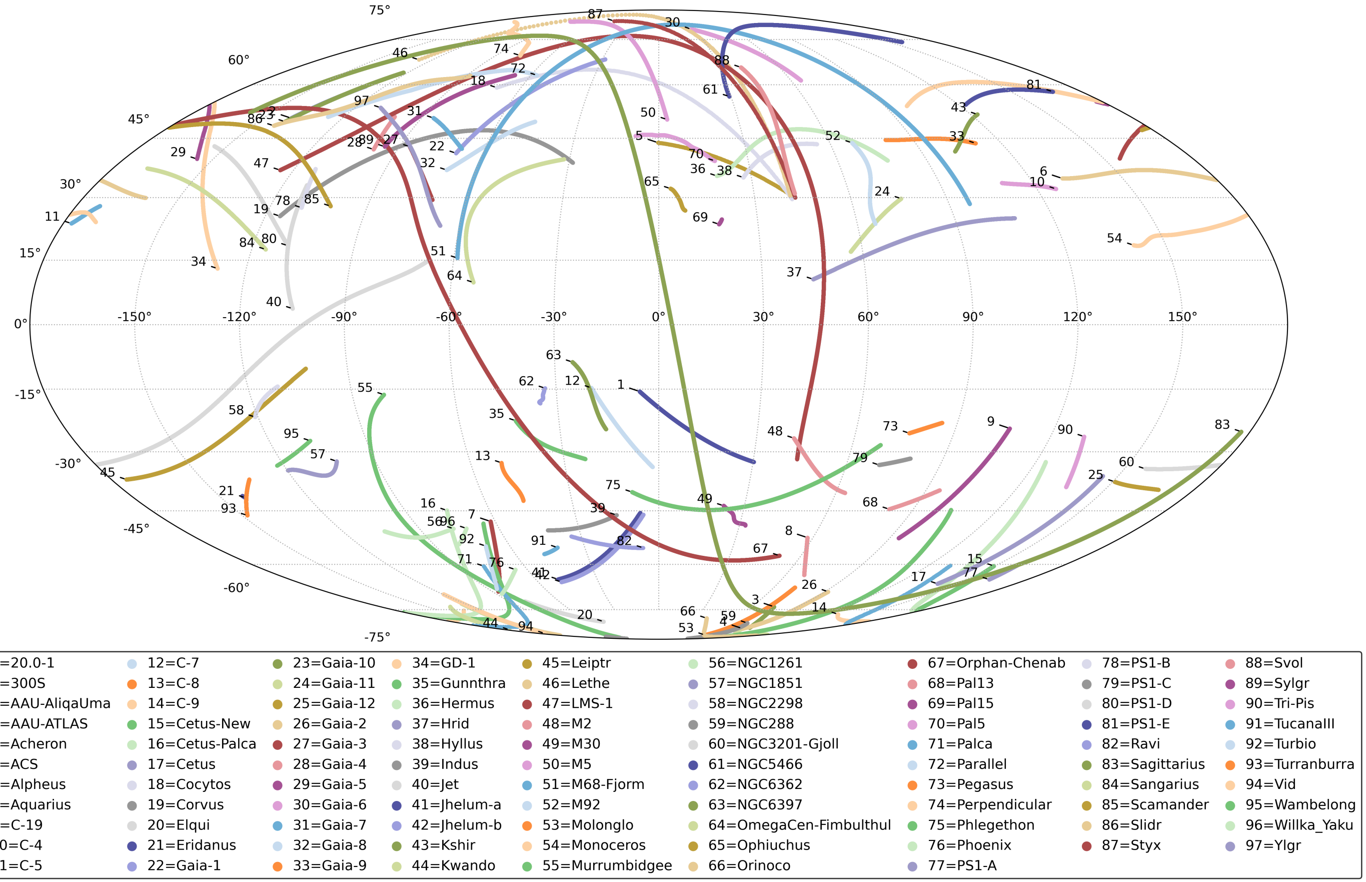

Figure 1.3: Map of the tracks of the 97 currently known coherent stellar streams cataloged in the galstreams package in Galactic coordinates. Image credit: Mateu (2022)

thick disc from the proto-thin disc through dynamical heating (Helmi et al. 2018). Other recent discoveries of substructure in the halo include the Thamnos structure (Koppelman et al. 2019b), Sequoia dwarf (Myeong et al. 2019), Aleph, Arjuna, I'itoi and Wukong (Naidu et al. 2020), and Pontus (Malhan et al. 2022).

1.2 Galactic potential

The knowledge of mass, shape and mass profile of the Milky Way is crucial for answering questions relevant to many fields of astronomy and physics, such as:

- What is the nature of dark matter?
- How do galaxies form and evolve?
- What is the accretion history of the Milky Way?
- How does the Milky Way compare to galaxies in cosmological simulations?

The total mass of the Milky Way is necessary for addressing the small scale challenges of the ΛCDM model (see reviews by Del Popolo & Le Delliou 2017; Bullock & Boylan-Kolchin 2017). One of the earliest issues pointed out when comparing the Milky Way to galaxies from cosmological simulations was the Missing Satellites Problem (Klypin et al. 1999; Moore et al. 1999). The essence of this problem lies in the discrepancy between the number of satellite galaxies found within the halo of the Milky Way and the halos of Milky-Way-like galaxies in simulations. Simulations suggest that thousands of satellites should be expected (e.g. Springel et al. 2008; Griffen et al. 2016), while ~ 50 have currently been observed in the Milky Way (Newton et al. 2018). Various solutions to this problem have been suggested within the ΛCDM paradigm, for example the inclusion of the recently discovered ultra faint dwarf satellites in the satellite budget (Drlica-Wagner et al. 2015) and different baryonic feedback effects that suppress the formation of subhalos or visible stellar populations within subhalos (Wetzel et al. 2016, see e.g.). But, since the predicted number of satellites depends on the host galaxy's total mass (Fattahi et al. 2020, see also Mao et al. 2015) this problem could also be alleviated by a lower mass Milky Way.

A closely related Too Big To Fail problem (Boylan-Kolchin et al. 2011), which posits that the predicted amount of massive and dense subhalos in ΛCDM simulations do not have an observed counterpart in the Milky Way despite these subhalos being too big to have failed at forming stars, could be solved in the same manner. As the number of massive subhalos is dependent on the Milky Way mass, a lighter Milky Way would settle this problem (Wang et al. 2012). These problems can additionally be explained by modifications to the standard ΛCDM model, for example, with other forms of dark matter. Acquiring accurate measurements of the Milky Way mass is therefore a crucial step towards appraising the validity of the ΛCDM model.

Alongside the total mass, the radial mass profile is likewise of considerable interest, because the mass profile or, equivalently, the gravitational potential is a prerequisite for the derivation of orbits and, therefore, vital for all studies of orbital evolution. For example, Galactic archaeology relies upon the knowledge of the gravitational potential both for identifying as yet undiscovered substructures in integrals of motion space (Naidu et al. 2020; Horta et al. 2021; Malhan et al. 2022) and for tracing these individual structures backwards in time to reconstruct the assembly history of the Galaxy. However, the orbital properties of substructure do not only reveal much about our own Galaxy and its past, but also have connotations further afield: orbital modelling is instrumental in constraining general galaxy formation scenarios, such as satellite quenching mechanisms (Fillingham et al. 2019), and for providing tests of the ΛCDM paradigm by e.g. revealing the nature of the vast polar structure of satellite galaxies in the Milky Way, a configuration that would be very uncommon in ΛCDM (Pawlowski 2018; Banik et al. 2022, although see Samuel et al. 2021).

The predicted shapes of galactic dark matter halos differ between different models of dark matter. For example, in the standard collisionless cold dark matter model the halos of galaxies acquire a triaxial shape (Dubinski & Carlberg 1991) whereas self-interacting dark matter would create halos that are more spherical (Tulin & Yu 2018, see also Vargya et al. 2021; Sameie et al. 2020, 2021). The shape of the Milky Way halo can therefore offer constraints on the dark matter models and, moreover, potentially rule out modified gravity models (Khoury 2015).

Despite the fundamental importance of these Galactic properties, there is considerable scatter, or even outright contradictions, in their estimates. The total mass of the Milky Way is uncertain to within a factor of two, with estimates ranging from $0.5 - 2.0 \times 10^{12} M_{\odot}$ (see detailed overview by Wang et al. 2020). Likewise, there is a high degree of inconsistency between different halo shape measurements, from prolate (Posti & Helmi 2019) to oblate (Loebman et al. 2014) to nearly spherical (Bovy et al. 2016) (excellent overview of different results is given in the Introduction of Gallo et al. 2021 and Discussion of Hattori et al. 2021). This picture is further complicated by the fact that the infalling LMC has an effect on the shape and orientation of the Galactic halo (Vasiliev et al. 2021).

Although these results were obtained with a variety of different dynamical tracers and techniques, one of the main causes of the variation between these estimates is due to unaccounted for systematics in the methods (Wang et al. 2020). This strongly emphasizes on the one hand, the necessity of serious scrutiny into the employed methods and models, and, on the other hand, the value of having a variety of dynamical tracers and analysis methods as a way of verifying the results.

Stellar streams, as discussed above in Section 1.1.2, are sensitive probes of the Galactic mass distribution. Yet, only relatively few individual streams have thus far been used to constrain the Galactic potential. Two factors have, until recently, hindered the progress: the challenge of developing methods that consider multiple streams simultaneously without compromising on computational cost or accuracy, and the scarcity of streams with full 6-dimensional phase space information.

Within the past few years we have entered a golden age of Galactic astronomy, brought on largely by ESA's *Gaia* satellite (Gaia Collaboration et al. 2016a). In 2018, *Gaia* data release 2 (DR2) provided the 5-dimensional phase space information for 1.3 billion sources and the full 6-dimensional map for 7.2 million bright stars (Gaia Collaboration et al. 2018). In each subsequent data release, both the number of sources and the precision of the astrometry and spectroscopy is forecast to increase. This revolutionary data has already facilitated great leaps in our understanding of the Milky Way, for instance the discovery of Gaia-Enceldaus merger debris (Helmi et al. 2018), the detection of the "phase spiral" in the disc indicating an out-of-equilibrium disc (Antoja et al. 2018) and the indication that major star formation episodes have been triggered by the pericentric passages of the Sagittarius dwarf (Ruiz-Lara et al. 2020). However, since stellar streams are faint and orbit the galaxy at large distances – the more distant ones being also of higher interest since they probe the further reaches of the halo – even with *Gaia* data on hand, attaining the full 6-dimensional phase space for stream stars requires complementary sources of data for distances and radial velocities. For halo studies standard candles, such as RR Lyrae, or asteroseismology (Auge et al. 2020) are great options for mapping distances, while current or upcoming ground-based spectroscopic surveys, such as WEAVE (Dalton et al. 2012), 4MOST de Jong et al. (2019a), DESI (Levi et al. 2019), H3 (Conroy et al. 2019a), S^5 (Li et al. 2019), deliver precision radial velocities.

1.2.1 Extracting the Galactic potential with stellar streams

A variety of methods have been developed for extracting the Galactic potential with the use of stellar streams. The general approach is to produce a model of a stellar stream under the assumption of a particular Galactic potential which is then compared to the stream data. Then, the potential is adjusted and a new comparison is made, until a certain figure of merit is optimized. However, the currently available methods vary greatly in terms of the complexity with which they model the streams and the disruption process, and the space in which they make comparisons with data.

Orbit-fitting technique. The simplest way to model a stream is to assume that stream stars perfectly delineate a single orbit. The progenitor's position is integrated backwards and forwards in time in different potentials, resulting in a series of trajectories that can be compared with the mean tracks of observed stream stars. This method has been used with success, for example, by Koposov et al. (2010); Willett et al. (2009) and Malhan & Ibata (2019) on GD-1 data and by Newberg et al. (2010) on Orphan stream data. However, in reality streams do not perfectly follow the progenitor's orbit and making this assumption can lead to systematic biases in Galactic potential constraints (Eyre & Binney 2011; Sanders & Binney 2013a).

N-body simulations. The most accurate model of a stream can be achieved by performing detailed N-body simulations. Typically, a single particle representing the present day phase-space position of the progenitor is integrated backwards in time in a test potential, whereupon this past position is used to launch the N-body simulation again forwards in time (in the same potential) to present day. The present day N-body particles are required to fulfill a series of observational constraints which helps distinguish between different test potentials. This approach is computationally expensive and thus restricts the amount of potential parameters (as well as parameters related to the progenitor's orbit) that can reasonably be explored. However, the parameter space can be reduced by applying the N-body approach only to those regions of parameter space which have previously been deemed relevant through less intensive methods, as was done by Law & Majewski (2010) when modelling the Sagittarius stream.

Lagrange point stripping methods. The methods in this class were developed to address the limitation of the N-body simulations. Although simpler than full N-body simulations, they offer vastly improved speeds while also successfully reproducing the tracks of simulated streams. Multiple approaches have been proposed by different authors such as Küpper et al. (2012); Gibbons et al. (2014); Bowden et al. (2015), but they all follow similar principles with only subtle differences. Ejected stars are generated, with some time intervals, at the two Lagrange points near the progenitor and assigned velocities which are slightly offset from that of the progenitor. After the particles are released, they are evolved forwards in time in a Galactic potential (the influence of the progenitor's potential can also be included). As the present day is reached, the track of the stream is calculated from the star particles and compared to the observed data. These methods have been successfully applied to multiple streams such as GD-1 (Bowden et al. 2015), Pal 5 (Küpper et al. 2015), Sagittarius (Gibbons et al. 2014) and Orphan (Erkal et al. 2019).

Angle-frequency correlation methods. Motivated by the stream-orbit misalignment that renders the orbit-fitting method inaccurate, Sanders & Binney (2013b) developed a method using action-angle coordinates (see Section 2.2.2 of this Thesis for a brief overview and Binney & Tremaine 2008 for details) which bypasses the assumption that streams delineate a single orbit. Their method relies on the notion that angles, representing the orbital phase, and orbital frequencies of stream stars lie on straight lines with equal slopes if evaluated with the true Galactic potential. The Galactic potential can therefore be constrained by searching for a trial potential that maximizes the correlation between angles and frequencies. Although demonstrated on simulated streams, the method has yet to be applied to observational data. Sanders (2014) and Bovy (2014) further exploit the linear time-evolution that stellar streams exhibit in angle-frequency space to produce models of streams that, after transformation to position-velocity space with an assumed Galactic potential, can be compared to real stellar stream data. This method has been applied to observational data of GD-1, Pal 5 and their combination by Bovy et al. (2016).

Rewinder method. The Rewinder method proposed by Price-Whelan et al. (2014) avoids explicitly modelling the streams altogether. Starting from the current day observed positions and velocities, their method integrates the stream stars and the progenitor backwards in time and poses that the true Galactic potential is one in which all the stream stars' orbits converge on the progenitor orbit, within some range defined by the the tidal radius and escape velocity of the progenitor. One of the advantages of this approach is that constraints on the Galactic potential can be placed with a very small number of observed stream stars.

Although these methods approach the problem in different ways and come with different advantages, they all have a few limiting aspects in common. Firstly, they typically require knowledge on the progenitor properties, such as position, velocity, mass, etc. Most observed stellar streams, however, have not been associated with a known progenitor. To analyse these streams one would therefore be compelled to include the progenitor's properties as free parameters if the model allows it. Secondly, the majority of these methods have been developed and demonstrated to work on single streams. Individual stream fits have, however, been shown to lead to systematic biases in Galactic potential constraints (Bonaca et al. 2014, Chapter 3 of this Thesis). Furthermore, as the number of streams with well-measured phase-space maps is rapidly growing, it is imperative to take maximal advantage of this data and aim for a joint fit from a collection of streams. Yet, a fit from multiple streams can be challenging to obtain with the methods discussed above as each added stream would require their progenitor's properties to be added as free parameters, leading to a explosive growth in the considered parameter space. These approaches therefore reach the limit of their practicability quickly. Thirdly, the methods all depend on individual stars having been identified as stream members. This membership information is not always readily available or sufficiently certain, especially if the intention is to shift from analysing single streams to taking advantage of the highly structured nature of the stellar halo as a whole. Finally, these methods have been tailored for use with coherent streams and, thus far, disregarded phase-mixed streams as a source of information. Procedures that compare the tracks of model streams to observed streams in position and/or velocity space are inapplicable to phase-mixed streams, since phase-mixed streams no longer exhibit a distinguishable track in these spaces.

1.2.2 The action-clustering method

To overcome the issues discussed in the previous section, and with an eye on future data sets, Sanderson et al. (2015) introduced a new method for constraining the Galactic potential based on action-space clustering of stream stars (see also Peñarrubia et al. 2012 and Magorrian 2014).

Dynamics of stellar streams has a simple and natural description in action-angle coordinates. Here, we will give only a brief overview of this coordinate system and refer the readers to Binney & Tremaine (2008) for a more rigorous derivation and discussion. The actions, $\boldsymbol{J}$, are integrals of motion that uniquely define an orbit and the angles, $\boldsymbol{\theta}$, are periodic coordinates that specify the orbital phase in a coordinate system $(\boldsymbol{q}, \boldsymbol{p})$ where the Hamiltonian is separable. For a star on a bound orbit in a static or adiabatically evolving potential, the actions are conserved while angles increase linearly with time

$$\boldsymbol{J} = const., \qquad \boldsymbol{\theta}(t) = \boldsymbol{\theta}(0) + \boldsymbol{\Omega} t, \tag{1.1}$$

where $\boldsymbol{\Omega}$ are the orbital frequencies.

To calculate actions, one needs (i) the 6-dimensional phase-space position of the star, $(\boldsymbol{x}, \boldsymbol{v})$, and (ii) a model for the Galactic potential, Φ. The transformation from phase space to action space can then be done following

$$J_i = \frac{1}{2\pi} \oint p_i dq_i \ , \tag{1.2}$$

where q_i are the coordinates and p_i the momenta of the star, and the integration is over a full oscillation in q_i. This problem is often expressed in cylindrical coordinates, where $q_1 = R$, $q_2 = z$ and $q_3 = \phi$, since the Galactic potential is approximately axisymmetric and hence separable in these coordinates. The actions would then be specified as J_R, J_z and J_ϕ, representing the extent of the orbit in the corresponding coordinate directions. This formalism therefore very conveniently allows us to express an entire stellar orbit – as long as the star's current position along the orbit is not of interest – with just three action values, whereas the description of an orbit in position-velocity space is complex and requires the six-dimensional $(\boldsymbol{x}(t), \boldsymbol{v}(t))$ at multiple time steps to specify the star's trajectory.

The process of stream disruption and formation in the action-angle coordinates can be described as follows. When the stars are stripped from the progenitor, they acquire a small initial offset both in actions and angles. Assuming that the progenitor's influence on the star is negligible and can be ignored, this initial offset in actions remains constant, while the differences in angles linearly increase – these ever-increasing angle differences manifest as stars stretching out along their orbits over time and are therefore responsible for stream formation (Helmi et al. 1999; Binney & Tremaine 2008; Bovy 2014).

As stream stars move on very similar orbits, their actions form tight clusters in action space. However, this compact configuration weakens or disappears when the actions are calculated with a potential model that is far from the potential in which the stream actually evolved: with the incorrect potential the stars would no longer be placed on neighbouring orbits. This principle is illustrated in Figure 1.4 where we compare the action space of mock stellar streams calculated with the correct potential (left panel) and an incorrect potential (right panel). This is the property that action-clustering method capitalizes on. Specifically, the action-clustering method searches, among a set of trial potentials, for the potential that maximises the clustering of stream stars in action space. This is equivalent to seeking for a series of orbits, one for each stream star, that are as similar as possible.

The action-clustering method is the technique of choice for constraining the Galactic potential with stellar steams in this Thesis, forming an integral part of Chapters 2-4. Although the amount of clustering found in action space can be measured with various techniques, in this work we make use of a method that utilizes the Kullback-Liebler Divergence (Kullback & Leibler 1951), explained in detail in Section 2.2.3 and 2.2.4.

The action-clustering method has multiple crucial advantages:

(i) The method avoids explicitly modelling the streams, and therefore is not sensitive to the specific assumptions of stream disruption process.

(ii) The method does not require knowledge of the properties of the stream's progenitor, nor does it add these properties as free parameters. Both streams with and without known progenitors can be analysed and will be treated equivalently, except that stars still within the progenitor are not of interest.

(iii) Similarly, membership information, that is the assignment of individual stars to specific streams, is not necessary for the method to work. Including stars that do not actually belong to the stream in the analysis would serve to decrease the degree of clustering for all trial potentials – the interlopers are unlikely to cluster with the true stream in any potential and only add random noise to the action space – and thus does not impact the results.

(iv) The method is applicable both to coherent and phase-mixed streams.

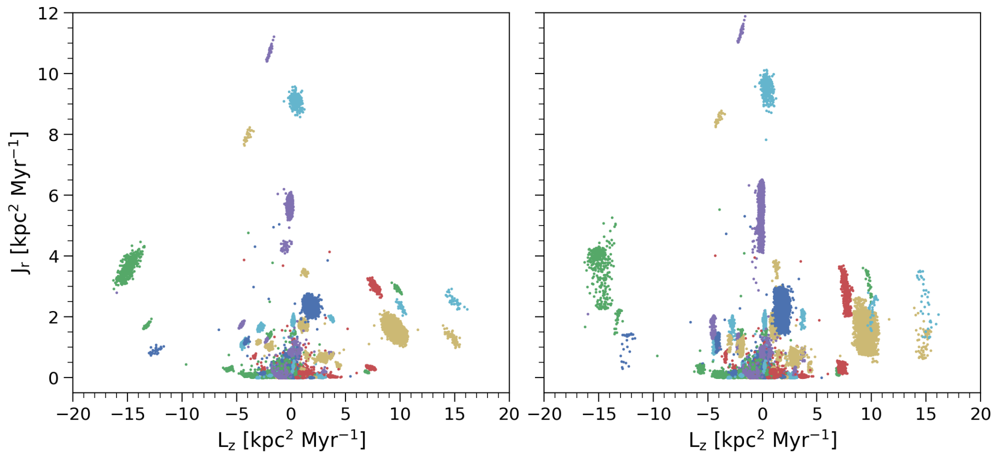

Figure 1.4: Action space of mock stellar streams when calculated with the true galactic potential (left) and with an incorrect potential (right). Image credit: adapted using data from Sanderson et al. (2015).

Although the members of phase-mixed streams have evolved to such a point that they cover the angle-space completely, if their actions have nevertheless been conserved, i.e. if the Galactic potential has evolved adiabatically the entire time, they still appear as a compact cluster in action-space.

(v) Finally, the method can be applied to multiple stellar streams simultaneously without inflating the considered parameter space. The only free parameters are those relating to the potential model, consequently the inclusion of additional streams has no effect on the size of the parameter space. Even with multiple streams in the analysed sample, the method is still searching for the maximally clustered total action space, without needing to specify the expected location of each cluster or the expected number of clusters.

The combination of these points means that besides applying the method on known streams and their likely members, we can also apply it to the stellar halo as a whole, since the latter is essentially a collection of overlapping streams. These advantages, however, come at a cost. Firstly, the full six-dimensional phase space is needed to calculate the action coordinates. There are, at present, no means to handle stars with missing phase space information within this method, rendering stars with incomplete data unusable. Secondly, exact actions exist only in a select few potentials, so we are either limited to using approximate action calculation techniques or making do with this restrictive choice of potential models. In this Thesis, we have taken latter route, and work with potentials of Stäckel type (see Section 2.2.1). Finally, currently the method does not account for the phase space measurement errors when transforming from position-velocity coordinates to action coordinates, and thus, when deriving the constraints on the Galactic potential. However, this feature can be added in the future.

The action-clustering method has previously been demonstrated on mock streams evolved in a toy potential (Sanderson et al. 2015) and on streams from Aquarius cosmological N-body simulations (Sanderson et al. 2017). Chapters 2 and 4 of this Thesis mark the first time the method has been applied to observational data.

1.3 Thesis outline

This Thesis can be divided into two branches, the first aiming to provide constraints on the Galactic potential using stellar streams and explore the details of the action-clustering method (Chapters 2 - 4), while the second deals with the kinematic study of the Hyades open cluster (Chapter 5).

In Chapter 2, we apply the action-clustering algorithm to observed stream data for the first time. Specifically, we analyse a collection of four streams – GD-1, Pal 5, Orphan and Helmi streams – both independently and jointly, and derive constraints on the Milky Way's mass profile. To compile the full six-dimensional phase space information for the members of these streams, we cross-match the proper motions from the *Gaia* survey with radial velocity and distance measurements from other surveys and literature sources. We find that different streams yield different estimates of the mass profile with hints that some streams might be more susceptible to biases than others. We also learn that these biases can be counteracted when a combination of streams is simultaneously analysed.

In Chapter 3, we delve deeper into the systematic biases at play with the help of simulated streams found in FIRE cosmological-hydrodynamical simulations. We confirm that our choice of potential model, the Stäckel model, is itself not a source of bias. The rest of the Chapter is then dedicated to analysing the connection between different stream properties and the systematic biases in their results. To allow this kind of a comparison, while keeping as many properties as possible constant, we spilt two long multi wrap streams into sections and analyse them separately as if they were individual streams. This set-up reveals that the there is a clear connection between the quality of the results and the phase of the stream from which the constraints are drawn: the pericentre streams consistently returned worse constraints. We also hypothesize that this pericentre bias is fundamental, since it is linked to the amount of correlation present between positions and momenta in the stream's data.

In Chapter 4, we return to constraining the Galactic potential with observed data while, in parallel, also working with simulations, to interpret our results and further examine the impact of our choice of method and data. Instead of considering separate streams, like in Chapter 2, this time we analyse the general population of halo stars with only approximate membership information. Our sample for this work originates from the stellar halo observations of Hectochelle in the Halo at High Resolution (H3) Survey and *Gaia* DR2. Concurrently, we apply the same pipeline to a population of streams from FIRE simulations to investigate the effects that a significant fraction of added interlopers and/or realistic phase-space measurement errors would have on our results. These tests help us uncover a limitation in our error estimation method and to propose an alternative one.

In Chapter 5, we quit the topic of stellar streams to move to the realm of open clusters and present the study of the Hyades open cluster. We use the positions, proper motions and parallaxes of stars in the Hyades region from *Gaia* DR1 Tycho-*Gaia* Astrometric Solution (TGAS) and Hipparcos-2. When possible, this 5-dimensional phase-space data is supplemented with radial velocities from multiple literature sources to obtain the full six-dimensional phase-space for a subset of the stars. Using this data set, we identify new candidate members for the cluster and, subsequently, derive improved parallaxes through a maximum-likelihood kinematical modelling method, which supports the simultaneous analysis of stars both with and without radial velocities. These kinematically-modelled parallaxes are more precise than the trigonometric parallaxes from either TGAS or Hipparcos-2 and, therefore, allow us to construct an extremely sharp main sequence for the Hyades.