Galactic Stellar Populations in the Era of SDSS and Other Large Surveys

Željko Ivezić

Department of Astronomy, University of Washington, Box 351580, Seattle, WA 98195

Timothy C. Beers

Department of Physics & Astronomy and JINA: Joint Institute for Nuclear Astrophysics, Michigan State University, East Lansing, MI 48824

Mario Jurić

Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138

Abstract Studies of stellar populations, collections of stars with common spatial, kinematic, chemical, or age distributions, have been reinvigorated during the last decade by the advent of large-area sky surveys such as SDSS, 2MASS, RAVE, and others. We review recent analyses of

these data that, together with theoretical and modeling advances, are revolutionizing our understanding of the nature of the Milky Way, and galaxy formation and evolution in general. The formation of galaxies like the Milky Way was long thought to be a steady process leading to a smooth distribution of stars. However, the abundance of substructure in the multi-dimensional space of various observables, such as position, kinematics, and metallicity, is by now proven beyond doubt, and demonstrates the importance of mergers in the growth of galaxies. Unlike smooth models that involve simple components, the new data reviewed here clearly show many irregular structures, such as the Sagittarius dwarf tidal stream and the Virgo and Pisces overdensities in the halo, and the Monoceros stream closer to the Galactic plane. These recent developments have made it abundantly clear that the Milky Way is a complex and dynamical structure that is still being shaped by the merging of neighboring smaller galaxies. We also briefly discuss the next generation of wide-field sky surveys, such as SkyMapper, Pan-STARRS, Gaia and LSST, which will improve measurement precision manyfold, and comprise billions of individual stars. The ultimate goal, development of a coherent and detailed story of the assembly and evolutionary history of the Milky Way and other large spirals like it, now appears well within reach.

CONTENTS

1   INTRODUCTION

1.1   The Big Picture: Structure Formation and Near-field Cosmology

The current cosmological paradigm states that the Universe had its beginning in the Big Bang. Galaxies, the fundamental building blocks of the Universe, formed soon after the Big Bang. A major objective of modern astrophysics is to understand when and how galaxies formed, and how they have evolved since then. Our own galaxy, the Milky Way, provides a unique opportunity to study a galaxy in great detail by measuring and analyzing the properties of a large number of individual stars. Since the individual stars that make up the stellar populations in the Milky Way can be studied in great detail, their characterization provides *clues about galaxy formation and evolution that cannot be extracted from*

observations of distant galaxies.

In the canonical model of Milky Way formation (Eggen, Lynden-Bell & Sandage 1962) the Galaxy began with a relatively rapid ($\sim 10^8$yr) radial collapse of the initial protogalactic cloud, followed by an equally rapid settling of gas into a rotating disk. The ELS scenario readily explained the origin and general structural, kinematic and metallicity correlations of observationally identified populations of field stars, and implied a smooth distribution of stars observable today. The predictions of the ELS scenario were quantified by the Bahcall & Soneira (1980) and Gilmore, Wyse & Kuijken (1989) models, and reviewed in detail by, e.g., Majewski (1993). In these smooth models, the Milky Way is usually modeled by three discrete components described by relatively simple analytic expressions: the thin disk, the thick disk, and the halo.

However, for some time, starting with the pioneering work of Searle & Zinn (1978) and culminating with recent discoveries of complex substructure in the distribution of the Milky Way's stars, this standard view has experienced difficulties. Unlike those smooth models that involve simple components, the new data indicate much more irregular structures, such as the Sgr dwarf tidal stream and the Virgo and Pisces overdensities in the halo, and the Monoceros stream closer to the Galactic plane. The recent observational developments, based on accurate large-area sky surveys, have made it abundantly clear that the Milky Way is a complex and dynamical structure that is still being shaped by the infall (merging) of neighboring smaller galaxies. Numerical simulations suggest that this merger process plays a crucial role in setting the structure and motions of stars within galaxies, and is a generic feature of current cosmological models (Brook et al. 2005; Bullock & Johnston 2005; Governato et al. 2004, 2007; John-

ston et al. 2008; Sommer-Larsen, Götz & Portinari 2003; Steinmetz & Navarro 2002).

The main purpose of this review is to summarize some of the recent observational progress in Milky Way studies, and the paradigm shifts in our understanding of galaxy formation and evolution resulting from this progress. The review is focused on only a few studies based mostly on data collected by the Sloan Digital Sky Survey York et al. (2000, hereafter SDSS), and does not represent an exhaustive overview of all the progress during the last decade. One of our main goals is to illustrate novel analysis methods enabled by new datasets. We begin with a brief overview of methodology, and of a few major datasets, and then describe the main observational results. We conclude by discussing some of the unanswered questions, and observational prospects for the immediate future.

1.2   Stellar Populations: Definition and Role

In astronomy, the term *stellar populations* is often associated with *Populations I, II and III*. These stellar classes represent a sequence of decreasing metallicity and increasing age. Here, we will use the term "stellar population" to mean any collection of stars with common spatial, kinematic, chemical, luminosity, or age distributions. For example, a sample of red giant stars selected using appropriate observables and selection criteria is considered a population, although such a sample can include both Population I and Population II stars. Similarly, we will often consider populations of "disk" and "halo" stars, or samples selected from a narrow color range. In summary, any sample of stars that share some common property that is appropriate for mapping the Galaxy in the space of various observables is hereafter considered to be a "population".

Most studies of the Milky Way can be described as investigations of the stellar distribution, or statistical behavior of various stellar populations, in the seven-dimensional (7-D) phase space spanned by the three spatial coordinates, three velocity components, and metallicity (of course, the abundances of individual chemical elements can be treated as additional coordinates). Depending on the quality, diversity and quantity of data, such studies typically concentrate on only a limited region of this 7-D space (e.g. the nearby solar neighborhood, pencil beam surveys, kinematically biased surveys), or consider only marginal distributions (e.g., number density of stars irrespective of their metallicity or kinematics, proper motion surveys without metallicity or radial velocity information). *The main reason for the substantial progress in our knowledge of the Milky Way structure over the last decade is the ability of modern sky surveys to deliver the necessary data for determining phase-space coordinates of a star for unprecedented numbers of faint stars detected over a large sky area.* For example, in less than two decades the observational material for kinematic mapping has progressed from first pioneering studies based on only a few hundred objects (Majewski 1992), to over a thousand objects (Chiba & Beers 2000), to the massive datasets including millions of stars reviewed here.

The large number of stars enables detailed studies of various distributions, including determination of the distributions' *shape*, rather than considering only low-order statistics as done with small samples. Deviations from Gaussian shapes often encode more information about the history of galaxy assembly than the distribution's mean and dispersion. The large samples are especially important with considering multi-variate distributions (as opposed to one-dimensional marginal distributions), when the so-called "curse of dimensionality" prevents their accu-

rate determination with small samples.

In addition to increasing the sample size, the ability to detect faint stars is crucial for extending the sample distance limit. With SDSS, it has become possible to detect even main sequence (dwarf) stars to a distance limit exceeding 10 kpc and thus to probe both disk and halo with the same dataset. For comparison, the Hipparcos sample (Perryman et al. 1997) contains only main sequence stars within ∼100 pc. The main advantage of main sequence stars over probes such as RR Lyrae stars, blue horizontal branch (BHB) stars and red giant stars for studying Galaxy is that they are much more numerous (of the order thousand times), and thus enable a much higher spatial resolution of the resulting phase-space maps (assuming a fixed number of stars per multi-dimensional pixel in phase space). Of course, those other probes are still valuable because they can be used to explore the Galaxy to a larger distance limit than obtainable with main sequence stars.

A common theme to most studies reviewed here is the use of photometric parallax relations to estimate stellar distances, and subsequent *direct mapping* of various distributions using large samples of stars. This mapping approach does not require *a-priori* model assumptions, and instead *constructs multi-dimensional distribution maps first, and only then looks for structure in the maps and compares them to Galactic models.* A key observational breakthrough that made this approach possible was the availability of accurate multi-band optical photometry to a faint flux limit and over a large sky area delivered by SDSS, as discussed below.

1.3 Observations: Photometry, Spectroscopy, Astrometry

In order to determine coordinates of a star in the 7-D phase space, a variety of astronomical techniques must be used. As always, the most crucial quantity to measure is stellar distance. The largest sample of stars with trigonometric distances, obtained by the Hipparcos survey, is too shallow (and small) to complement deep surveys such as SDSS and 2MASS (for overview of these surveys see below). Until the all-sky Gaia survey measures trigonometric distances for about a billion stars brighter than $V = 20$ (see the last section), various photometric methods need be employed to estimate distances to stars. A common aspect of these methods is that luminosity (i.e., absolute magnitude) of a star is determined from its color measurements, and then its distance is determined from the difference between absolute and apparent magnitudes. For certain populations, for example RR Lyrae stars, a good estimate of their absolute magnitude is a simple constant; for other populations, such as main sequence stars, absolute magnitude depends on both effective temperature and metallicity, and sometimes on age (or surface gravity), too. A photometric parallax method for main sequence stars is summarized below.

The most accurate measurements of stellar metallicity are based on spectroscopic observations (but see below for a method for estimating metallicity using photometric data). The spectroscopic measurements are especially important when studying the extremely low end of the metallicity distribution where photometric methods become insensitive. In addition to measuring chemical composition, spectroscopic observations enable radial velocity measurements. The two largest existing stellar spectroscopic surveys are SDSS and RAVE (see the next section).

To measure all three components of the velocity vector, precise astrometric observations are also needed. The projection of the velocity vector into the tangent plane (i.e., perpendicular to the radial velocity component) is measured using proper motion (astrometric shift per unit time), which can be combined with distance estimate to yield velocity. The proper motion measurements place an additional constraint on observations that at least two astrometric epochs must be available.

Therefore, both multi-color imaging, multi-epoch astrometry, and spectroscopy are required for measuring coordinates of a star in the 7-D position-velocity-metallicity phase space. It is the advent of massive and accurate imaging and spectroscopic surveys that delivered such measurements for large relatively unbiased samples of stars, and thus enabled the major progress in the Milky Way mapping during the last decade.

1.3.1   PHOTOMETRIC PARALLAX METHOD FOR MAIN SEQUENCE STARS   In order to estimate distances to main sequence stars with an accuracy of 10-20% using photometric parallax relation, multi-band optical photometry accurate to several percent (i.e., to several hundredths of a magnitude) is required. This stringent requirement comes from the steepness of the color-luminosity relation (derivative of the absolute magnitude in the SDSS r band with respect to the $g - i$ color reaches ~ 10 mag/mag at the blue end), and is the main reason why it was not possible to use this method with large sky surveys prior to SDSS.

Using globular cluster data obtained in SDSS photometric system, Ivezić et al. (2008b, hereafter I08) derived a polynomial expression for the absolute magnitude of main sequence stars in the r band as a function of their $g - i$ color and metallicity (see their eqs. A2 and A7). The accuracy of the resulting magnitudes

is in the range 0.1-0.2 mag (I08; SIJ08), and the method enables studies of the ~100 pc to ~10 kpc distance range when used with SDSS data. *The ability to estimate distances to main sequence stars with sufficient accuracy using only SDSS photometry was crucial for wide-angle panoramic mapping of the Galaxy to a distance limit 100 times further than possible with the Hipparcos data.*

1.3.2 Photometric Metallicity Method for Main Sequence Stars

Stellar metallicity, together with effective temperature and surface gravity, is one of the three main parameters that affect observed spectral energy distribution of most stars. In addition to being an informative observable when deciphering the Milky Way history (e.g., Majewski 1993; Freeman & Bland-Hawthorn 2002a; Helmi 2008; Majewski 2010; and references therein), the knowledge of stellar metallicity is crucial for accurate estimates of distances using photometric parallax relation.

The most accurate measurements of stellar metallicity are based on spectroscopic observations. However, despite the recent progress in the availability of stellar spectra (approaching a million!), the number of stars detected in imaging surveys is still vastly larger. In addition to generally providing better sky and depth coverage than spectroscopic surveys, imaging surveys obtain essentially complete flux-limited samples of stars. These simple selection criteria are advantageous when studying Galactic structure, compared to the complex targeting criteria that are often used for spectroscopic samples.

As first suggested by Schwarzschild, Searle, & Howard (1955), the depletion of metals in a stellar atmosphere has a detectable effect on the emergent flux, in particular in the blue region where the density of metallicity absorption lines is highest (Beers & Christlieb 2005, and references therein). Recent analysis of

SDSS data demonstrated that for blue F and G main sequence stars, a reasonable metallicity estimate can be derived from the $u-g$ color (I08, B10). The expression A1 from B10, applicable to stars with $0.2 < g - r < 0.6$, was calibrated using $\sim$100,000 stars with spectroscopic metallicity, and has errors in the range 0.2-0.3 dex when used with SDSS data (for stars in the range $-2 < [Fe/H] < 0$). Although applicable only within a restricted color range, this calibration enabled the construction of metallicity maps using millions of stars, as discussed further below.

2   THE ADVENT OF LARGE-AREA DIGITAL SURVEYS

Major advances in our understanding of the Milky Way have historically arisen from dramatic improvements in our ability to "see", as vividly exemplified by Galileo resolving the Milky Way disk into individual stars. Progressively larger telescopes have been developed over the past century, but until recently most astronomical investigations have focused on small samples of objects because largest telescope facilities typically had rather small fields of view, and those with large fields of view could not detect very faint sources. Over the past two decades, however, astronomy moved beyond the traditional observational paradigm and undertook large-scale sky surveys, such as SDSS and the Two Micron All Sky Survey (Skrutskie et al. 2006, hereafter 2MASS). This observational progress, based on advances in telescope construction, detectors, and above all, information technology, has had a dramatic impact on nearly all fields of astronomy, including studies of the Milky Way structure. Here we briefly overview the characteristics of the most massive recent datasets.

2.1   SDSS Imaging and Spectroscopic Surveys

The SDSS is a digital photometric and spectroscopic survey which covered over one quarter of the Celestial Sphere in the North Galactic cap, and produced a smaller area ($\sim$300 deg^2) but much deeper survey in the Southern Galactic hemisphere (Abazajian et al. 2009, and references therein). The recent Data Release 7 has a sky coverage of about 12,000 deg, and includes photometric measurements for 357 million unique objects (approximately half are stars). The completeness of SDSS catalogs for point sources is $\sim$99% at the bright end and drops to 95% at the r band magnitude of $\sim$22. The wavelength coverage of the SDSS photometric system ($ugriz$, with effective wavelengths from 3540 Å to 9250 Å) and photometry accurate to $\sim$0.02 mag have enabled photometric parallax and metallicity estimates for many millions of stars. For comparison, the best large-area optical sky survey prior to SDSS, the photographic Palomar Observatory Sky Survey, had only two photometric bands and several times larger photometric errors (Sesar et al. 2006).

In addition to its imaging survey data, SDSS has obtained well over half a million stellar spectra (Yanny et al. 2009, $\sim$660,000). These spectra have wavelength coverage 3800–9200 Å and spectral resolution of $\sim$2000, with a signal-to-noise ratio per pixel of 5 at $r \sim 20$. SDSS stellar spectra are of sufficient quality to provide robust and accurate stellar parameters, such as effective temperature, surface gravity, and metallicity (parameterized as $[Fe/H]$). These publicly available parameters are estimated using a variety of methods implemented in an automated pipeline (Beers et al. 2006, the SEGUE Stellar Parameters Pipeline, SSPP). A detailed discussion of these methods and their performance can be found in Allende Prieto et al. (2008) and Lee et al. (2008a,b). Based on a com-

parison with high-resolution abundance determinations, they demonstrate that the combination of spectroscopy and photometry from SDSS is capable of delivering estimates of T_{eff}, $log(g)$, and $[Fe/H]$ accurate to 200 K (3%), 0.28 dex, and 0.17 dex, respectively. Random errors for the radial velocity measurements are a function of spectral type, but are usually < 5 km s^{-1} for stars brighter than $r \sim 18$, rising to $\sim$20 km s^{-1} for stars with $r \sim 20$ (Pourbaix et al. 2005, Yanny et al. 2009). Lee et al. (2011a) demonstrate that SDSS spectra are of sufficient quality to also determine $[\alpha/Fe]$ with errors below 0.1 dex (for stars with temperatures in the range 4500-7000 K and sufficient signal-to-noise ratios). The distribution of SDSS stars with spectra in the log(g) vs. color plane is shown in Figure 1.

2.2 SDSS-POSS Proper Motion Survey

The time difference of about half a century between the first Palomar Observatory Sky Survey (POSS) and SDSS imaging observations provides an excellent baseline to measure proper motions for tens of millions of stars to faint brightness levels. Munn et al. (2004) addressed the problem of large systematic astrometric errors in POSS catalogs by recalibrating the USNO-B catalog (Monet et al. 2003) using the positions of galaxies measured by SDSS. As a result of this calibration, the SDSS-POSS proper motion measurements are now available for about 100 million unresolved sources, mostly stars. This catalog also includes about 70,000 spectroscopically confirmed SDSS quasars that were used to robustly estimate the proper motion errors (Bond et al. 2010). The random errors increase from $\sim$3 mas yr^{-1} at the bright end to $\sim$6 mas yr^{-1} at $r \sim 20$ (the sample completeness limit), with systematic errors typically an order of magnitude smaller and

with very small variation across the sky. Even for stars at 1 kpc, the implied tangential velocity errors are as small as 10-20 km s^{-1}, and well matched to the SDSS radial velocity accuracy. This catalog represents a major improvement over previously available data sets *both in size and accuracy.*

2.3   2MASS Imaging Survey

The 2MASS is an all-sky near-IR survey with limiting (Vega-based, 10σ) magnitudes of J=15.8, H=15.1, and K=14.3. The 2MASS point source catalog contains positional and photometric information for 471 million sources (mostly stars). The near-IR 2MASS colors are not as good as optical SDSS colors for estimating photometric parallax and metallicity because they only probe the Rayleigh-Jeans tail of the stellar spectral energy distribution. On the other hand, a major advantage of 2MASS over SDSS is the full sky coverage, and its ability to penetrate deeper through the interstellar dust in the Galactic plane. In addition, it is much easier to photometrically identify certain stellar populations using near-IR data than with optical data. For example, Majewski et al. (2003) have demonstrated that M giant candidates color-selected from 2MASS database are extremely powerful probe for finding and tracing halo substructure out to ~100 kpc and across the whole sky (these stars are practically impossible to robustly identify using SDSS photometry). For an analysis of the joint SDSS-2MASS dataset for stars, we refer the reader to Covey et al. (2007).

2.4   RAVE Spectroscopic Survey

RAVE is a major new spectroscopic survey aiming to measure radial velocities and stellar atmosphere parameters (temperature, metallicity, and surface gravity)

of up to one million stars using the Six Degree Field multiobject spectrograph on the 1.2 m UK Schmidt Telescope of the Anglo-Australian Observatory (Steinmetz et al. 2006). RAVE stars are selected from the magnitude range $9 < I < 12$ and represent a bright complement to the SDSS spectroscopic sample (Siebert et al. 2011). The wavelength range for the RAVE spectra (8410–8795 Å with a spectral resolution of ~ 8000) includes a number of lines in addition to iron, and detailed data analyses should eventually provide an estimate of $[\alpha/Fe]$, in addition to overall metallicity ($[Fe/H]$).

The latest third data release includes radial velocity data for $\sim 77,000$ stars and stellar parameters for $\sim 40,000$ stars (Siebert et al. 2011), but spectra are already collected for over 300,000 stars (Zwitter et al. 2010). With a radial velocity error of about 2 km s^{-1}, the RAVE velocities are more accurate than those delivered by SDSS, and well matched to detailed kinematic studies of nearby disk stars. Proper motions (of varying accuracy) are available for most of the RAVE stars from other surveys, and model-based distance determinations accurate to $\sim 20\%$ are also available (Burnett et al. 2011, Zwitter et al. 2010).

The distance range probed by RAVE stars ranges from $\sim$300 pc (dwarfs) to $\sim$1-2 kpc (giants), and thus the RAVE dataset "connects" the nearby Hipparcos sample and the more distant SDSS sample. Due to these distance limits, RAVE data are more relevant for disk than for halo investigations. The large-area nearly-contiguous coverage of RAVE survey (see Figure 2) is very important for panoramic Galactic mapping.

3   OVERVIEW OF STATE-OF-THE-ART A DECADE AGO

UNFINISHED: here comes Tim's contribution.

Before we present results obtained during the last decade, we briefly overview the state of related knowledge a decade ago.

4   WHAT DID WE LEARN DURING THE LAST DECADE?

Until recently, our global view of the Milky Way was hampered by the fact that most detected stars had no reliable distance estimates. Those stars that had usable estimates were either limited to the solar neighborhood (e.g., for main sequence stars in the Hipparcos sample to within ∼100 pc, or only ∼1% of our distance to the Galactic center), or to small pencil-beam surveys. Our knowledge of the basic structural components was thus limited to indirect inferences based on stellar population models motivated by other spiral galaxies (e.g., Bahcall & Soneira 1980, Robin et al. 2003). This limitation was alleviated recently by the advent of SDSS which provided accurate digital multi-band optical photometry across a quarter of the sky. The SDSS photometry enabled development and application of photometric parallax methods, which in turn led to direct mapping of stellar distributions in the multi-dimensional space spanned by spatial coordinates, velocity components, and chemical abundance measurements. The resulting maps provided quantitative basis for separating the main structural Galaxy components and for their phenomenological description, and also enabled efficient searches for substructure and a robust comparison with various model predictions.

We first describe how these new data clearly reveal disk and halo as two distinct Galaxy components, and then describe each of them in more detail.

4.1   Separation of the Main Structural Components

Using photometric data for ~50 million stars, J08 have constructed 3-dimensional maps (data cubes) of the stellar number density distribution for 19 narrow color bins that span spectral types from mid-F to early M. When the bin color is varied from the reddest to the bluest one, the maps are "zoomed out", with subsamples covering distances ranging from 100 pc to 15 kpc. Distance to each star was estimated using a maximum likelihood implementation of photometric parallax method, and stars are binned and counted in small 3-dimensional pixels whose size depends on dynamical range provided by each color bin and Poisson noise limits (typically there are 250,000 pixels per map). Examples of two-dimensional projections of the resulting maps are shown in Figure 3.

These maps are a powerful tool for studying the Milky Way's stellar number density distribution. Traditional methods for modeling stellar counts in the magnitude-color space need to adopt a large number of poorly known functions such as the initial mass function, the mass-luminosity relationship, the luminosity function, and geometric description of the postulated components such as disks, bulge and halo. Instead, with these number density maps the Milky Way's structure can be studied without any a priori assumptions about its components. *With these maps, analysis of the Milky Way's structure is now akin to studies of external galaxies.*

The quantitative description of these maps is not a trivial task because of the rich substructure. While halo substructure has been known for a while (Belokurov et al. 2006, Ivezić et al. 2000, Majewski et al. 2003, Vivas & Zinn 2006, Yanny et al. 2000), these new maps demonstrate that the disk substructure is also complex. Nevertheless, the gross behavior can be captured by assuming standard Galaxy

models based on two exponential disks and a power-law halo. J08 determined the best-fit parameter values for full two-dimensional smooth models and further refined them using residual minimization algorithms.

A cross section of the maps from Figure 3 in the direction perpendicular to the disk plane is shown in Figure 4. The data shown in the middle and bottom panels clearly confirm a change in the counts behavior around $|Z| \sim$1-1.5 kpc, interpreted as evidence for an extended "thick" disk component by Gilmore & Reid (1983). When this additional more extended component becomes unable to explain counts at $|Z| \sim$5 kpc, another component – the stellar halo – is invoked to explain the data. Although these modern counts have exceedingly low statistical noise and fairly well understood systematics, the three-component fit to data shown in the bottom panel begs the question whether a single-component fit with some other function parametrized with fewer free parameters might suffice.

It turns out that the three components invoked to explain the counts display distinctive chemical and kinematic behavior, too. Figure 5 shows a panoramic view of the variation in the median $[Fe/H]$ over an unprecedentedly large Galaxy volume. The map is based on photometric metallicity for a sample of 2.5 million blue main sequence stars (most of F spectral type) selected using very simple color and flux limits. It is easily discernible that the median metallicity further than $\sim$ 5 kpc from the Galactic plane is very uniform and about 1 dex lower than for stars within $\sim$ 1 kpc from the plane.

The reason for a very fast decrease of the median metallicity with $|Z|$ for $|Z| < 5$ kpc, and very little variation further from the plane, is illustrated in the left panel in Figure 6. Two distinct distributions implying different Galaxy components, halo and disk, are clearly evident. High-metallicity disk stars dominate close

to the plane, while low-metallicity halo stars dominate beyond 3 kpc from the plane. The median metallicity for disk stars shows a gradient, while halo stars have spatially invariant metallicity distribution. As $|Z|$ increases from $|Z| \sim 2$ kpc to $|Z| \sim 4$ kpc, halo stars become more numerous than disk stars, and the median metalicity drops by ~ 1 dex. A more detailed and quantitative discussion of these metallicity distributions can be found in I08.

These two components with distinct metallicity distributions also have vastly different kinematic behavior, as shown in the right panel in Figure 6. The high-metallicity disk stars have large rotational velocity (about 200 km s^{-1}), while the low-metallicity halo stars display behavior consistent with no net rotation. Similarly to the behavior of their metallicity distributions, the rotational velocity for disk stars decreases with the distance from the Galactic plane, while it is constant for halo stars (see Figure 7).

Therefore, fairly clean samples of halo and disk stars can be defined using a simple metalicity boundary $[Fe/H] = -1$. We proceed with more detailed discussions of each component.

4.2   The Milky Way Disk

Recent massive datasets confirmed with exceedingly high statistical signal-to-noise the abrupt change of slope in the log(counts) vs. $|Z|$ plot around $|Z| \sim 1$ kpc for disk stars, discovered almost three decades ago by Gilmore & Reid (1983). A key question now is whether the two disk components required to explain the counts can also be used to explain chemical and kinematic measurements for the same stars. In other words, what is an optimal way to decompose disk into thin and thick disk components?

I08 showed that observed variations in metallicity and velocity distributions of disk stars over the $Z \sim 1 - 3$ kpc range can be modeled as smooth shifts of metallicity and velocity distributions that do not change their shape. They argued that this ability to describe observations using functions with *universal Z-independent shapes* has fundamental implications for disk origin: instead of two distinct components, the data can be interpreted with a single disk, albeit with metallicity and velocity distributions more complex than traditionally used Gaussians (an alternative is to use thin/thick disk decomposition, though it also requires non-Gaussian components). While the disk separation into thin and thick components may be a useful concept to describe the fairly abrupt change of number density around $|Z| \sim 1$ kpc, the disk spatial (counts) profile may simply indicate a complex structure (i.e. not a single exponential function), rather than two distinct entities with *different* formation and evolution history. The implication of their conclusions is that *different processes led to the observed metallicity and velocity distributions of disk stars,* rather than a single process, such as mergers or an increase of velocity dispersion due to scattering, that simultaneously shaped both distributions.

On other hand, I08 pointed out that stars from the solar neighborhood selected kinematically as thick-disk stars have larger α-element abundances, at the same $[Fe/H]$, than do thin-disk stars (e.g., Fuhrmann 2004; Bensby et al. 2004; Feltzing 2006; Reddy et al. 2006; Ramírez et al. 2007). In addition, the thick-disk stars, again selected kinematically, appear older than the thin-disk stars (e.g., Fuhrmann 2004; Bensby et al. 2004). They concluded that measurements of α-element abundances for samples of distant stars extending to several kpc from the midplane (as opposed to local samples) could resolve difficulties with traditional

thin-thick disk decomposition when applied to their data. Such a dataset was recently produced by Lee et al. (2011a) who showed that $[\alpha/Fe]$ ratio can be estimated using SDSS spectra: for stars with temperatures in the range 4500-7000 K and sufficient signal-to-noise ratios, $[\alpha/Fe]$ errors are below 0.1 dex.

4.2.1 THE HOLY GRAIL FOR THIN-THICK DISK DECOMPOSITION: $[\alpha/Fe]$

Lee et al. (2011b, hereafter L11) analyze a sample of ∼17,000 G dwarfs with $[\alpha/Fe]$ measurements, that were selected using simple colors and flux selection criteria. These data represent the first massive sample of stars at distances of several kpc, with reasonably accurate distance estimates, measurements of all three velocity components, measurements of both $[Fe/H]$ and $[\alpha/Fe]$, and selected using well-understood criteria over a large sky area. The L11 sample enabled several far-reaching observational breakthroughs:

1. The bimodal distribution of an unbiased sample of G dwarfs in the $[\alpha/Fe]$ vs. $[Fe/H]$ diagram (see Figure 8) strongly motivates the separation of the sample by an essentially simple $[\alpha/Fe]$ cut into two subsamples that closely resemble traditional thin and thick disks in their spatial distributions, $[Fe/H]$ distributions, and distributions of their rotational velocity (see Figure 1 in L11).

2. The low-$[\alpha/Fe]$, thin disk, subsample has an $[Fe/H]$ distribution that does not vary strongly with position within the probed volume ($|Z| < 3$ kpc and $7 < R/$kpc < 10), with a median value of $[Fe/H] \sim -0.2$. Similarly, the metallicity distribution for the high-$[\alpha/Fe]$, thick disk, subsample has a median value of $[Fe/H] \sim -0.6$, without a strong spatial variation (see Figure 4 in L11). The observed decrease of $[Fe/H]$ with distance from the Galactic plane (e.g., as reported by I08) appears to be the result of

increasing fraction of thick-disk stars.

3. The rotational velocity component, v_Φ, decreases linearly with the distance from the midplane, $|Z|$, with a gradient of $d|v_\Phi|/d|Z| \sim -10$ km s^{-1} kpc^{-1} for both thin and thick disk subsamples (see Figure 8 in L11). The difference between the mean values of v_Φ for the two subsamples of $\sim$35 km s^{-1} (asymmetric drift) is independent of $|Z|$, and explains the discrepancy between the $|Z|$ gradient of -10 km s^{-1} kpc^{-1} reported by L11, and gradients about 2-3 times steeper reported for the full disk by earlier studies (e.g., I08, Cassetii-Dinescu et al. 2011): as $|Z|$ increases from the midplane to 2-3 kpc, the fraction of thick disk stars increases from $\sim$10% to >90%, and the observed gradient when all stars are considered is affected by both the intrinsic gradient for each component, and the velocity lag of thick disk relative to thin disk stars.

4. The rotational velocity component does not show a gradient with respect to the radial coordinate, R, for thin disk stars (-0.1 ± 0.6 km s^{-1} kpc^{-1}; a "flat rotation curve"), and only a small and marginally detected gradient for thick disk stars (-5.6 ± 1.1 km s^{-1} kpc^{-1}).

5. The rotational velocity component and mean orbital radius are complex functions of the position in the $[\alpha/Fe]$ vs. $[Fe/H]$ diagram (see Figure 9). The rotational velocity component shows a linear dependence on metallicity for both thin and thick disk $[\alpha/Fe]$-selected subsamples (see Figure 10). The slopes of these v_Φ vs. $[Fe/H]$ correlations have *opposite* signs, $d|v_\Phi|/d[Fe/H] \sim -25$ km s^{-1}dex^{-1} for thin disk, and ~ 45 km s^{-1}dex^{-1} for thick disk, and do not strongly vary with distance from the midplane. These opposite gradients are partially responsible for the lack of correla-

tion between v_Φ and $[Fe/H]$ at $|Z| \sim 1$ kpc reported by I08 (for the full sample; the other reason for the lack of correlation are systematic errors in photometric metallicity estimator, see Appendix in L11).

6. Velocity dispersions for all three components increase with $[\alpha/Fe]$ as smooth functions, and continuously across the adopted thin/thick disk boundary (see Figure 3 in L11). Approximate values for velocity dispersions $(\sigma_R, \sigma_Z, \sigma_\Phi)$ are (40, 25, 25) km s^{-1} for thin disk and (60, 40, 40) km s^{-1} for thick disk.

7. Eccentricity distributions (model-dependent and determined using an analytic Stäckel-type gravitational potential from Chiba & Beers (2000)) are significantly different for the two $[\alpha/Fe]$-selected subsamples (see Figure 10 in L11), and show strong variation with the position and metallicity (see Figure 9 in L11). Notably, the shapes of the eccentricity distributions for the thin- and thick-disk populations are independent of distance from the plane, and include only a minute fraction of stars with eccentricity above 0.6.

In summary, Lee et al. (2011b) robustly demonstrated that disk stars can indeed be separated into thin and thick disk components. In addition to the bimodal distribution of $[\alpha/Fe]$ which motivates and enables this separation, further support for the decomposition into two components is provided by the fact that the spatial and kinematic distributions of each component display much simpler behavior than those for the full sample. It does not seem an overstatement to proclaim that $[\alpha/Fe]$ *measurements are the long-awaited holy grail for robust decomposition of disk stars into thin-disk and thick-disk components.*

On the other hand, a few words of caution are due here. The main results from L11 still need to be confirmed by independent datasets. It is somewhat

worrisome that the RAVE-based results from Burnett et al. (2011) for the disk $[Fe/H]$ distribution are different from L11 results. At $Z \sim 0$, the RAVE results are about 0.2 dex more metal rich (though we note that the SDSS result for the median $[Fe/H] = -0.2$ at $Z = 0$ is consistent with the results from Nordström et al. 2004), and the discrepancy increases to $\sim$0.3 dex at $Z \sim 2.5$ kpc. It is not clear yet whether discrepant results reported by RAVE and SDSS surveys are due to differences in metallicity scales, or due to unaccounted selection effects in RAVE analysis (see Section 6 in Burnett et al. 2011). Encouragingly, the spatial metallicity gradients at $Z \sim 1$ kpc, where the thick disk stars become more numerous than thin disk stars, are robustly detected and similar in both studies, $d[Fe/H]/d|Z| \sim -0.2$ dex/kpc. The median $[Fe/H]$ at $Z \sim 1$ kpc reported by Lee et al. (2011b) is -0.5 dex, about 0.2 dex lower than reported by Burnett et al. (2011)) using RAVE, and about 0.2 dex higher than reported by I08 using photometric metallicity from SDSS imaging survey. It remains to be seen how $[\alpha/Fe]$ measurements from SDSS and RAVE surveys compare to each other.

Burnett et al. (2011) study also reports age determination for RAVE stars (based on stellar models) with typical uncertainties of about a factor of 2 (see their Figure 7). They detect a remarkable age gradient between the Galactic midplane and $|Z| \sim 2$ kpc (see their Figures 16 and 17), which is at least qualitatively consistent with the variation of the $g-r$ color of turn-off stars seen by SDSS. They also detect a complex variation of metallicity distribution with stellar age (see their Figure 18). In particular, the oldest stars ($> 8-9$ Gyr) are predominantly low-metallicity ($[Fe/H] < -0.5$). These age data represent a valuable addition to L11 results. Nevertheless, determining age for individual stars is exceedingly hard (Pont & Eyer 2004, Soderblom 2010) and one needs to remember all the

caveats discussed by Burnett et al. at the end of their Section 7.

4.2.2   COMPARISONS OF OBSERVATIONS AND DISK FORMATION MODELS   Despite the three decades of thick disk studies, there is still no consensus on models for its formation and evolution (the thick disk is not unique to the Milky Way; for a review of thick disks in other galaxies, see van der Kruit & Freeman 2011). The proposed scenarios can be broadly divided into two groups: violent origin, such as heating of existing thin disk due to mergers, and secular evolution, such as heating due to scattering off molecular clouds and spiral arms (see L11 for a detailed discussion and references). In the first set of scenarios, the fraction of thick disk stars accreted from merged galaxies remains an important and still unconstrained parameter, and further complexity arises from the posssibility that some stars may have formed *in situ* when star formation is triggered in mergers of gas-rich galaxies (Brook et al. 2007, and references therein). In the second set of scenarios, the main modeling difficulty is the lack of detailed knowledge about relative importance of various scattering mechanisms. Over the last decade, the radial migration mechanism (Sellwood & Binney 2002; Roškar et al. 2008b; Schönrich & Binney 2009b; Minchev & Famaey 2010) has been developed as an attractive secular scenario. Due to various computational and other difficulties, numerical models that combine the main features of the violent and secular scenarios are scarce.

The recent observational material contains rich information for model testing, and is beginning to rule out some models. Modern data include simultaneous measurements of many observables for large numbers of stars, and enable qualitatively new approaches to tests of disk formation models. The more observables are measured, the more powerful are these tests because data can be "sliced"

along multiple axes in numerous ways, and small statistical errors can be main-
tained due to the large sample sizes. On the other hand, the complexity of such
tests can be formidable: even a minimalistic selection of observables, such as
coordinates R and Z, chemical parameters $[Fe/H]$ and $[\alpha/Fe]$, and essential
kinematic parameters, rotational velocity and orbital eccentricity, span a six-
dimensional space. The basic model vs. data comparisons for testing thick disk
formation and evolution scenarios include

1. Compare distribution of stars in the $[\alpha/Fe]$ vs. $[Fe/H]$ diagram, as a
   function of the position in the Galaxy (e.g., Can models reproduce the
   bimodal distribution seen in Figure 8? Does the fraction of sample in the
   high-$[\alpha/Fe]$ component increase with the distance from the midplane as
   observed?).

2. For subsamples defined using $[\alpha/Fe]$, compare the shapes of their metal-
   licity and kinematic distributions (e.g., Can models reproduce $[Fe/H]$ dis-
   tributions seen in Figure 4 from L11, or eccentricity distributions seen in
   their Figure 10?).

3. For subsamples defined using $[\alpha/Fe]$, compare the variations of their num-
   ber volume density and low-order statistics for metallicity and kinematic
   distributions (e.g., mean v_Φ, velocity dispersions, mean/mode/median ec-
   centricity) with the position in the Galaxy (e.g., Can models reproduce the
   spatial gradients of the mean v_Φ seen in Figure 8 from L11, or the spatial
   gradients of the mean eccentricity from their Figure 9?).

4. Compare high-order correlations between observables, such as the complex
   variation of the mean rotational velocity with the position in the $[\alpha/Fe]$
   vs. $[Fe/H]$ diagram (see Figures 9 and 10), or the variation of the orbital

eccentricity with metallicity (see Figure 9 in L11).

A few of such tests have already been performed. In a strict statistical sense, all the proposed models can be outright rejected because the observed distributions of various parameters have very low statistical noise, and the models are not sufficiently fine tuned (yet) to reproduce them (e.g., none of model eccentricity distributions comes even close to passing the Kolmogorov-Smirnov test). For this reason, most of model vs. data comparisons are still qualitative and only gross inconsistencies are used to reject certain scenarios.

Starting with Sales et al. (2009), a number of recent papers used the shape of eccentricity distribution as means to compare models to data from SDSS and RAVE surveys (Casetti-Dinescu et al. 2011, Di Matteo et al. 2011, Dierickx et al. 2010, Lee et al. 2011b, Loebman et al. 2011, Wilson et al. 2011). We note that orbital eccentricity is derived from observations in a model-dependent way (a gravitational potential must be assumed), and different assumptions may lead to systematic differences between observed and predicted distributions. In most of these studies, four published simulations of thick discs formed by a) accretion from disrupted satellites, (b) heating of a pre-existing thin disc by a minor merger, (c) radial migration and (d) gas-rich mergers (see Sales et al. for references), are confronted with data. The scenario a) predicts an eccentricity distribution that includes too many stars with high eccentricities (see Figure 3 in Sales et al. and Figure 10 in L11), and the scenario b) does not show the characteristic change of slope in the log(counts) vs. $|Z|$ plot (see Figure 1 in Sales et al.). These are the two main reasons for growing consensus that gas-rich mergers and radial migration scenarios are in best agreement (more precisely, least disagreement) with data.

Loebman et al. (2011) performed a number of data vs. model tests listed above in the limited context of radial migration models developed by Roškar et al. (2008a,b). They demonstrated that overall features seen in data, such as the gradients of metallicity and rotational velocity with distance from the midplane (see Figure 11), as well as the gradients of rotational velocity with metallicity (see their Figure 15), and the complex structure seen for the mean rotational velocity in the $[\alpha/Fe]$ vs. $[Fe/H]$ diagram (their Figure 14) are qualitatively reproduced by models (at detailed quantitative level there is room for improvement). We note an important implication of those models that $[\alpha/Fe]$ is an excellent proxy for age. Using a different numerical implementation of the radial migration scenario, Schönrich & Binney (2009a,b) demonstrated good agreement with local solar neighborhood data from the Geneva-Copenhagen survey (Nordström et al. 2004).

These model successes hint that the thick disk may be a ubiquitous Galactic feature generated by stellar migration (though note that a similar analysis was not done yet with gar-rich merger models). However, while these models at least qualitatively reproduce a lot of complex behavior seen in data, the radial migration cannot be the full story: there are counter-rotating disks observed in some galaxies (Yoachim & Dalcanton 2008), and remnants of merged galaxies are directly observed in the Milky Way (see the right column in Figure 3 and dicussion in §4.4 below).

4.2.3   A SUMMARY OF RECENT DISK STUDIES   To summarize, given the new SDSS, RAVE and other data, there is no doubt that the spatial and kinematic behavior of disk stars greatly varies as a function of their chemical composition parametrized by the position in the $[\alpha/Fe]$ and $[Fe/H]$ diagram. While quantitative details still differ somewhat between different analysis methods, and between

SDSS and RAVE datasets, robust conclusions are that the high-$[\alpha/Fe]$ subsample has all the characteristics traditionally assigned to thick disk: larger scale height, lower $[Fe/H]$, rotational velocity lag, and larger dispersions for all three velocity components, when compared to the low-$[\alpha/Fe]$ subsample. There is mounting evidence that ages of these stars are much higher than those in the low-$[\alpha/Fe]$ subsample, and similar to the age of Galaxy, though the interpretation of age data is much more prone to systematics than chemical and kinematic data.

Despite this tremendeous observational progress, there is still no consensus on theories for the origin of thick disk. The two main contenders remain gas-rich mergers and radial migration scenarios, while accretion scenario and disk heating appear to be in conflict with data. Nevertheless, no generic model/scenario should be fully rejected yet since the detailed comparison with data just began and the input model parameter space has not been fully explored. Assuming that SDSS measurements reported in L11 will survive further scrutiny (e.g., when compared to RAVE and other datasets), the modelers will be very busy for quite some time trying to explain the rich observational material collected over the last few years.

4.3 The Milky Way Halo

UNFINISHED

Studies of the Galactic halo provide unique insights in the formation history of the Milky Way, and the galaxy formation process in general...

Studies with main sequence turn-off stars to $\sim$10 kpc: spatial profiles from J08, $[Fe/H]$ distribution from I08, kinematics and velocity ellipsoid tilt from B10.

Indirect studies and dual halo from Carollo et al. papers.

Most distant halo: various luminous tracers, such as RR Lyrae variables, BHB

stars, and red giants are used to detect halo substructures.

BHB stars from DR8: Xue et al. (2011)

Summarize constraints on profile, $[Fe/H]$ and kinematics at distances beyond 30 kpc...

4.4   Streams and Other Substructures

UNFINISHED

Within the framework of hierarchical galaxy formation (Freeman & Bland-Hawthorn 2002b), the spheroidal component of the luminous matter should reveal substructures such as tidal tails and streams (Bullock, Kravtsov & Weinberg 2001; Harding et al. 2001; Helmi & White 1999; Johnston, Hernquist & Bolte 1996). The number of these substructures, due to mergers and accretion over the Galaxy's lifetime, may provide a crucial test for proposed solutions to the "missing satellite" problem (Bullock, Kravtsov & Weinberg 2000). Substructures are expected to be ubiquitous in the outer halo (galactocentric distance $> 15-20$ kpc), where the dynamical timescales are sufficiently long for them to remain spatially coherent (Johnston, Hernquist & Bolte 1996; Mayer et al. 2002), and indeed many have been discovered (e.g., Belokurov et al. 2007a,b, 2006, 2007c, Grillmair 2009, Grillmair & Dionatos 2006, Ivezić et al. 2000, Jurić et al. 2008, Newberg et al. 2007, 2002, Vivas & Zinn 2006, Yanny et al. 2000).

Streams (Grillmair!)... Mention Klement (2010) review

Sesar et al. Figure 17

The Cambridge group results, Wilman's work

Just how much substructure there is in the Milky Way halo? J08, also Bell et al. (2008)

Also substructure in the disk: large (unnamed) overdensities visible in the two middle panels in the right column in Figure 3, and Monoceros stream in Figure 18.

5   UNANSWERED QUESTIONS

UNFINISHED: here comes Mario's contribution.

6   THE ROAD AHEAD

UNFINISHED: ZI needs to finish (mostly copy & paste from other papers.

The results discussed here will be greatly extended by several upcoming large-scale, deep optical surveys, including the Dark Energy Survey (Flaugher 2008), Pan-STARRS (Kaiser et al. 2002), and ultimately the Large Synoptic Survey Telescope (Ivezić et al. 2008a). These surveys will extend the faint limit of the current surveys, such as SDSS, by up to 5 magnitudes. In addition, upcoming Gaia mission (Perryman et al. 2001, Wilkinson et al. 2005) will provide superb astrometric and photometric measurement accuracy for sources with $r < 20$ that will enable unprecedented science programs, and WISE mission will extend the probed wavelength range to 22 μm

6.1   Pan-STARRS, SkyMapper, and the Dark Energy Survey

Summarize PS, SM and DES...

6.2   WISE

Pasted...

NASA's Wide-field Infrared Survey Explorer (WISE; **?**) mapped the sky at

3.4, 4.6, 12, and 22 μm in 2010 with an angular of $6-12$ arcsec. WISE achieved 5σ point source sensitivities better than 0.08, 0.11, 1 and 6 mJy (corresponding to AB magnitudes of 19.1, 18.8, 16.4 and 14.5) in unconfused regions on the ecliptic in the four bands (for comparison, WISE represents an improvement over the IRAS survey's 12 μm band sensitivity by about a factor of 1000). The astrometric precision for high signal-to-noise sources is better than 150 mas. The survey sensitivity improves toward the ecliptic poles due to denser coverage and lower zodiacal background. Saturation affects photometry for sources brighter than approximately 8.0, 6.7, 3.8 and -0.4 mag (Vega) at 3.4, 4.6, 12 and 22 μm, respectively.

The WISE Preliminary Release[1] includes data from the first 105 days of WISE survey observations. Primary release data products include an Atlas of 10,464 calibrated, coadded Image Sets, a Source Catalog containing positional and photometric information for over 257 million objects detected on the WISE images, and an Explanatory Supplement that provides a user's guide to the WISE mission and format, content, characteristics and cautionary notes for the Release products.

From the viewpoint of stellar population studies, WISE...

6.3   Gaia

Perryman (2002)

Gaia is an ESA Cornerstone mission set for launch in 2012. Building on experience from HIPPARCOS, it will survey the sky to a magnitude limit of $r \sim 20$ (approximately, see the next section) and obtain astrometric and three-band pho-

---

[1]http://wise2.ipac.caltech.edu/docs/release/prelim/preview.html

tometric measurements for about 1 billion sources, as well as radial velocity and chemical composition measurements (using 847-874 nm wavelength range) for 150 million stars with $r < 18$. The final data product, the Gaia Catalogue, is expected to be published by 2020.

The Gaia's payload will include two telescopes sharing a common focal plane, with two $1.7° \times 0.6°$ viewing fields separated by a highly stable angle of $106.5°$. The focal plane includes a mosaic of 106 CCDs, with a total pixel count close to one billion. Due to spacecrafts' rotation and precession, the whole sky will be scanned in TDI (drift scanning) mode about 70 times on average during 5 years of operations. Gaia will produce broad-band G magnitudes with sensitivity in the wavelength range 330-1020 nm (FWHM points at $\sim$400 nm and $\sim$850 nm). The spectral energy distribution of each source will be sampled by a spectropho-tometric instrument providing low resolution spectra in the blue (BP, effective wavelength $\sim$520 nm) and the red (RP, effective wavelength $\sim$800 nm). In addition, the RVS instrument (radial velocity spectrograph) will disperse the light in the range 847–874 nm, for which it will include a dedicated filter. Therefore, there are four passbands associated with the Gaia instruments: G, G_{BP} , G_{RP} and G_{RVS}.

6.4   LSST

The Large Synoptic Survey Telescope (LSST) is the most ambitious currently planned ground-based optical survey, with a unique survey capability in the faint time domain. The LSST design is driven by four main science themes: probing dark energy and dark matter, taking an inventory of the Solar System, exploring the transient optical sky, and mapping the Milky Way. LSST will be a large,

wide-field ground-based system designed to obtain multiple images covering the sky that is visible from Cerro Pachón in Northern Chile. The current baseline design, with an 8.4m (6.7m effective) primary mirror, a 9.6 deg^2 field of view, and a 3.2 Gigapixel camera, will allow about 10,000 square degrees of sky to be covered using pairs of 15-second exposures twice per night every three nights on average, with typical 5σ depth for point sources of $r \sim 24.5$ (AB). The system is designed to yield high image quality as well as superb astrometric and photometric accuracy. The total survey area will include 30,000 deg^2 with $\delta < +34.5°$, and will be imaged multiple times in six bands, *ugrizy*, covering the wavelength range 320–1050 nm. The project is scheduled to begin the regular survey operations before the end of this decade. About 90% of the observing time will be devoted to a deep-wide-fast survey mode which will uniformly observe a 18,000 deg^2 region about 1000 times (summed over all six bands) during the anticipated 10 years of operations, and yield a coadded map to $r \sim 27.5$. These data will result in databases including 10 billion galaxies and a similar number of stars, and will serve the majority of the primary science programs. The remaining 10% of the observing time will be allocated to special projects such as a Very Deep and Fast time domain survey.

LSST will obtain proper motion measurements of comparable accuracy to those of Gaia at their faint limit, and smoothly extend the error vs. magnitude curve deeper by 5 mag (for details see Eyer et al., in preparation). With its u-band data, LSST will enable studies of metallicity and kinematics using the *same sample* of stars out to a distance of ~ 100 kpc (~ 200 million F/G main sequence stars brighter than $g = 23.5$, for a discussion see I08).

LSST will produce a massive and exquisitely accurate photometric and astro-

metric dataset for about 10 billion Milky Way stars. The coverage of the Galactic plane will yield data for numerous star-forming regions, and the y band data will penetrate through the interstellar dust layer. Photometric metallicity measurements will be available for about 200 million main-sequence F/G stars which will sample the halo to distances of 100 kpc (**?**). No other existing or planned survey will provide such a massive and powerful dataset to study the outer halo (including Gaia which is flux limited at $r = 20$, and Pan-STARRS which will not have the u band). The LSST in its standard surveying mode will be able to detect RR Lyrae and classical novae out to 400 kpc, and hence explore the extent and structure of the halo out to half the distance to M31. All together, the LSST will enable studies of the stellar distribution beyond the presumed edge of the Galactic halo, of their metallicity distribution throughout most of the halo, and of their kinematics beyond the thick disk/halo boundary (**?**).

In the context of Gaia, the LSST can be thought of as its deep complement. A comparison of LSST and Gaia performance is given in Figure 19. Gaia will provide an all-sky catalog with unsurpassed trigonometric parallax, proper motion and photometric measurements to $r \sim 20$ for about 10^9 stars. LSST will extend this map to $r \sim 27$ over half of the sky, detecting about 10^{10} stars. Because of Gaia's superb astrometric and photometric quality, and LSST's significantly deeper reach, the two surveys are highly complementary: Gaia will map the Milky Way's disk with unprecedented detail, and LSST will extend this map all the way to the halo edge (Eyer et al., in prep).
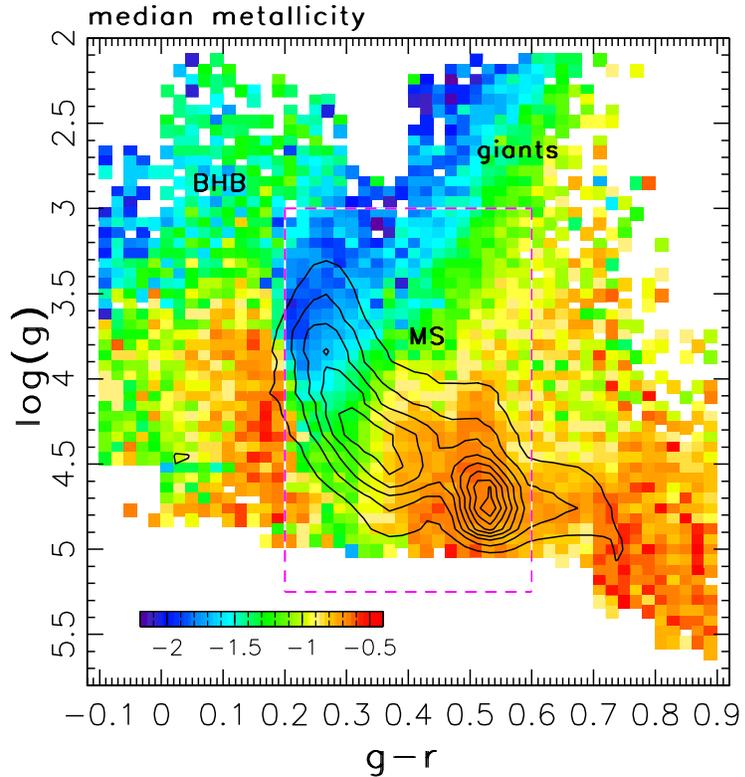
Figure 1: The stellar content of SDSS spectroscopic surveys (Figure 1 from Ivezić et al. (2008b)). Linearly spaced contours showing the distribution of ∼110,000 stars with $g < 19.5$ and $0.1 < g - r < 0.9$ (corresponding to effective temperatures in the range 45008200 K) in the log(g) vs. $g - r$ plane. The multimodal distribution is a result of the SDSS target selection algorithm. The color scheme shows the median metallicity in all 0.02 mag by 0.06 dex large pixels that contain at least 10 stars. The fraction of stars with log(g) < 3 (giants) is 4%, and they are mostly found in two color regions: $-0.1 < g - r < 0.2$ (BHB stars) and $0.4 < g - r < 0.65$ (red giants). They are dominated by low-metallicity stars ($[Fe/H] < -1$). The dashed lines outline the main-sequence (MS) region where photometric metallicity method can be applied.
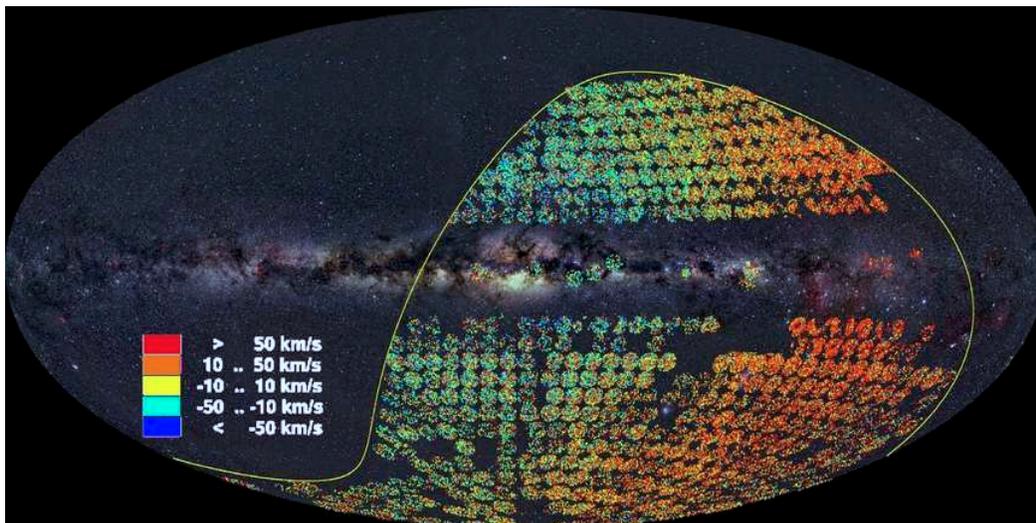
Figure 2: The sky coverage of RAVE's third data release shown as Aitoff projection in Galactic coordinates (Figure 17 from Siebert et al. 2011).
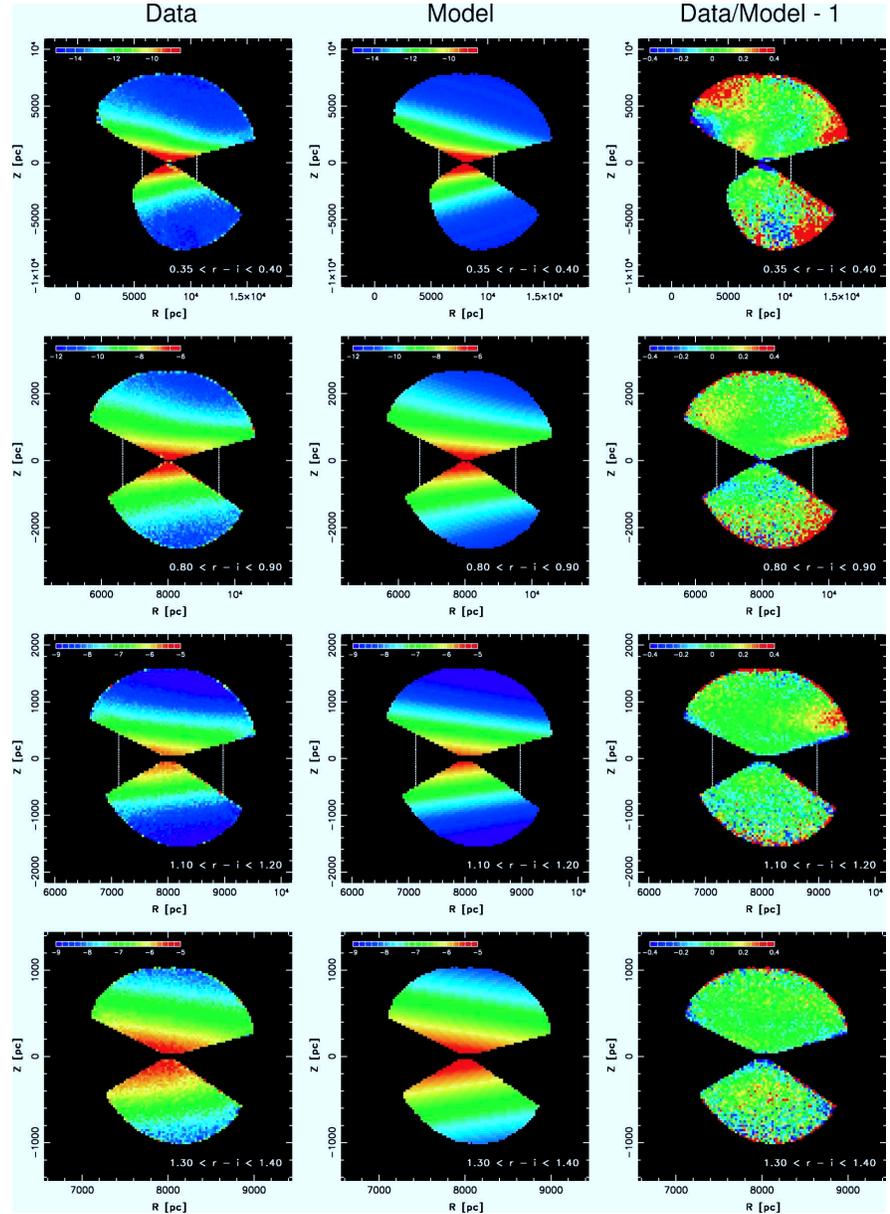
Figure 3: Figure 26 from Jurić et al. (2008). The panels in the left column show the measured stellar number density as a function of Galactic cylindrical coordinates, for stars selected from narrow ranges of the $r-i$ color ($0.35 < r-i < 0.40$ in the top row to $1.30 < r-i < 1.40$ in the bottom row). The panels in the middle column show the best-fit smooth models, and panels in the right column show the normalized (data-model) difference map. Note the large overdensities visible in the top three panels in the right column.
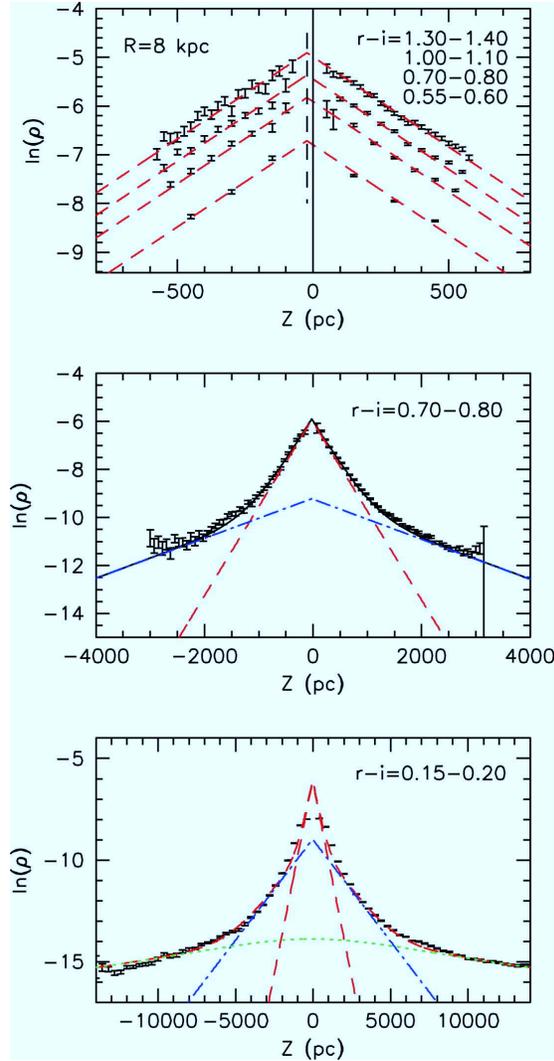
Figure 4: Cross sections through maps similar to those shown in Figure 3 showing vertical (Z) distribution at $R = 8$ kpc and for different $r - i$ color bins (Figure 15 from Jurić et al. 2008). The lines are exponential models fitted to the points. The dashed lines in the top panel correspond to a fit with a single, exponential disk. The dashed line in the middle panel correspond to a sum of two disks with scale heights of 270 pc and 1200 pc, and a relative normalization of 0.04 (the "thin" and the "thick" disks). The dashed line in the bottom panel (closely following the data points) corresponds to a sum of two disks and a power-law spherical halo. The dashed line and the dot-dashed line are the disk contributions, and the halo contribution is shown by the dotted line. For the final de-biased best-fit parameters see Table 10 in J08.
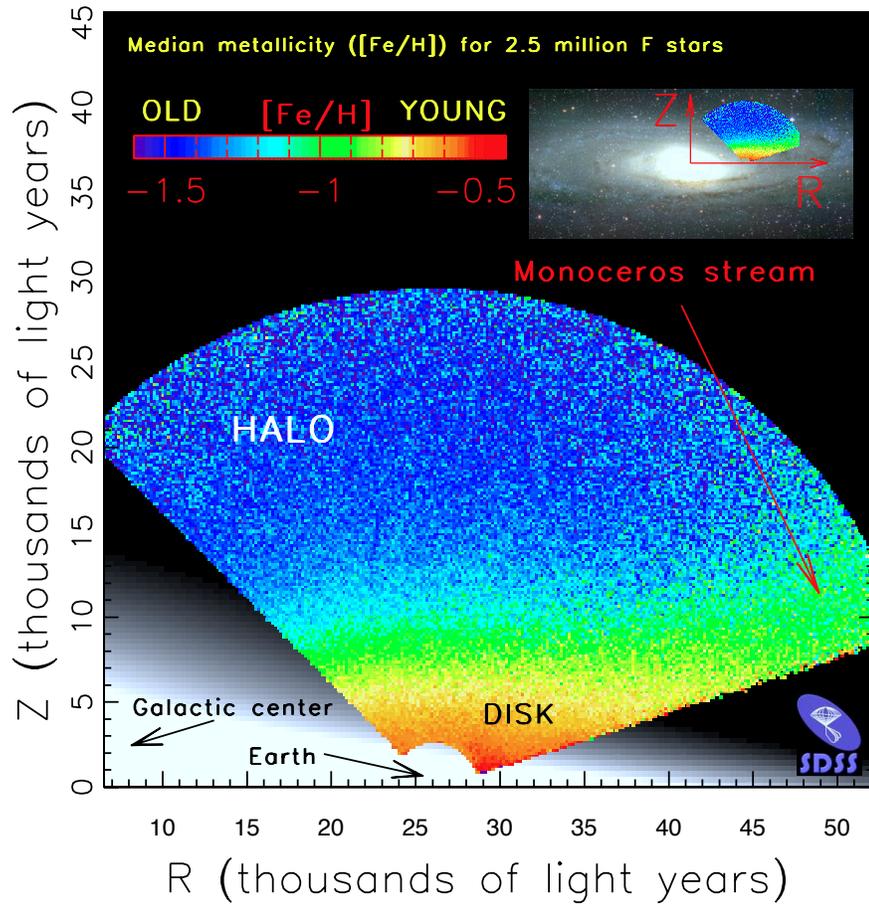
Figure 5: Variation of the median photometric metallicity for ∼2.5 million stars from SDSS with $14.5 < r < 20$ and $0.2 < g-r < 0.4$, and photometric distance in the 0.8-9 kpc range, in cylindrical Galactic coordinates R and $|Z|$. The ∼40,000 pixels (50 pc by 50 pc) contained in this map are colored according to the legend in the top left. Note that the gradient of the median metallicity is essentially parallel to the $|Z|$ axis, except in the Monoceros stream region, as marked. The gray scale background is the best-fit model for the stellar number density distribution from J08. The inset in the top right illustrates the extent of the data volume relative to the rest of Galaxy (the background image is Andromeda galaxy).
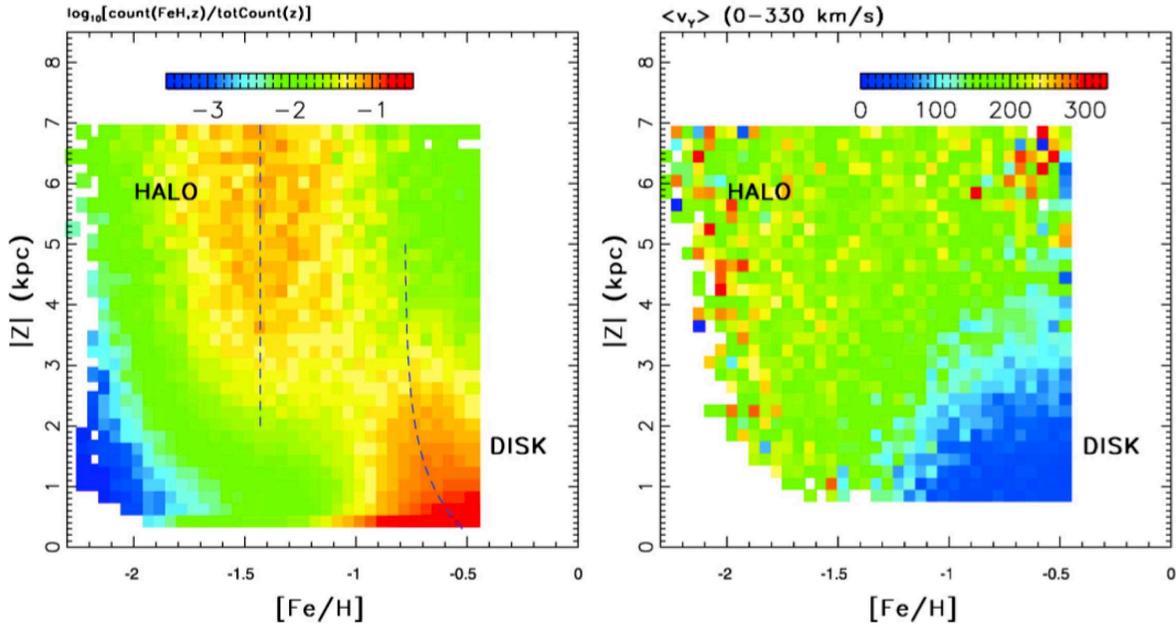
Figure 6: Figure 9 from Ivezić et al. (2008b). The left panel shows the conditional metallicity probability distribution (each row of pixels integrates to 1) for ∼60,000 stars from a cylinder perpendicular to the Galactic plane, centered on the Sun, and with a radius of 1 kpc. The values are color coded on a logarithmic scale according to the legend on top. An updated version of this map is shown in Figure A.3 from Bond et al. (2010). The right panel shows the median heliocentric rotational velocity component (the value of ∼220 km s^{-1} corresponds to no rotation) as a function of metallicity and distance from the Galactic plane for the ∼40,000 stars from the left panel that also satisfy $b > 80°$.
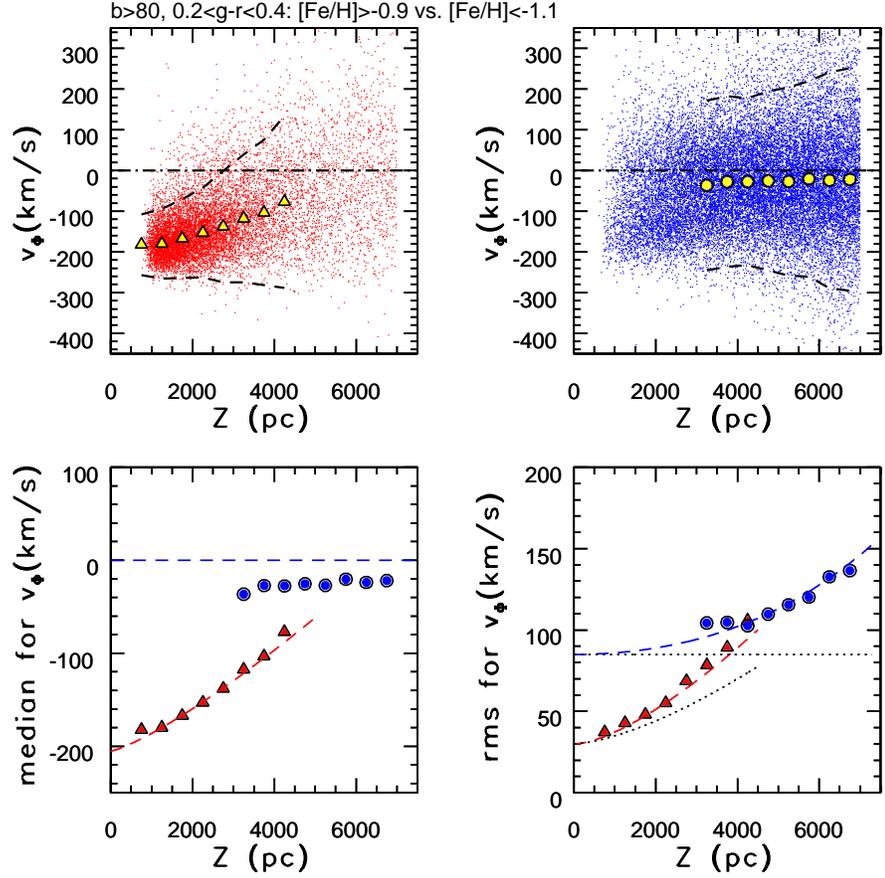
Figure 7: Figure 5 from Bond et al. (2010). A comparison of rotational velocity (see their eq. 8), v_Φ, on distance from the Galactic plane, Z, for 14,000 high-metallicity ($[Fe/H] > 0.9$; top-left panel) and 23,000 low-metallicity ($[Fe/H] < 1.1$; top right) stars with $b > 80°$. In the top two panels, individual stars are plotted as small dots, and the medians in bins of Z are plotted as large symbols. The 2σ envelope around the medians is shown by dashed lines. The bottom two panels compare the medians (left) and dispersions (right) for the two subsamples shown in the top panels, and the dashed lines in the bottom two panels show predictions of a kinematic model. The dotted lines in the bottom-right panel show model dispersions without a correction for measurement errors.
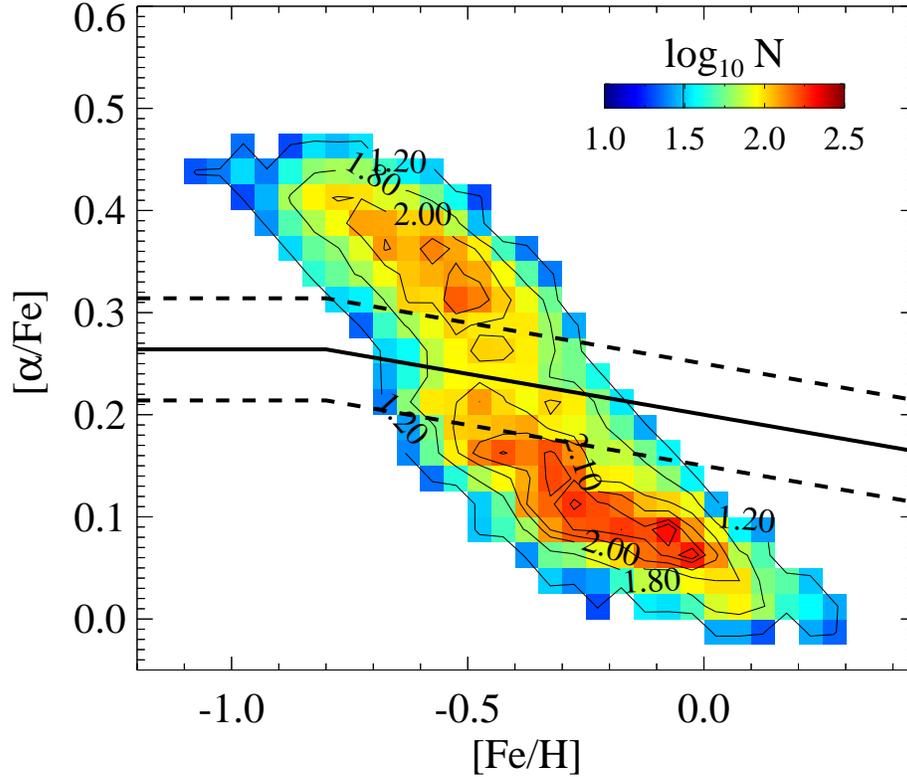
Figure 8: The $[\alpha/Fe]$ vs. $[Fe/H]$ distribution of G dwarfs within a few kpc from the Sun (Figure 2 from Lee et al. 2011b). The number density is shown on logarithmic scale according to the legend, and by isodensity contours. Each pixel (0.025 dex in $[\alpha/Fe]$ direction and 0.05 dex in $[Fe/H]$) contains at least 20 stars (with a median of 70 stars). The distribution of disk stars in this diagram can be described by two components (thin and thick disk, respectively) centered on $([Fe/H], [\alpha/Fe]) = (-0.2, 0.10)$ and $(-0.6, 0.35)$. The solid line is the fiducial for division into likely thin- and thick-disk populations; note that simple $[\alpha/Fe] = 0.24$ separation results in almost identical subsamples. The dashed lines show selection boundaries adopted by Lee et al. (2011b) which exclude the central region.
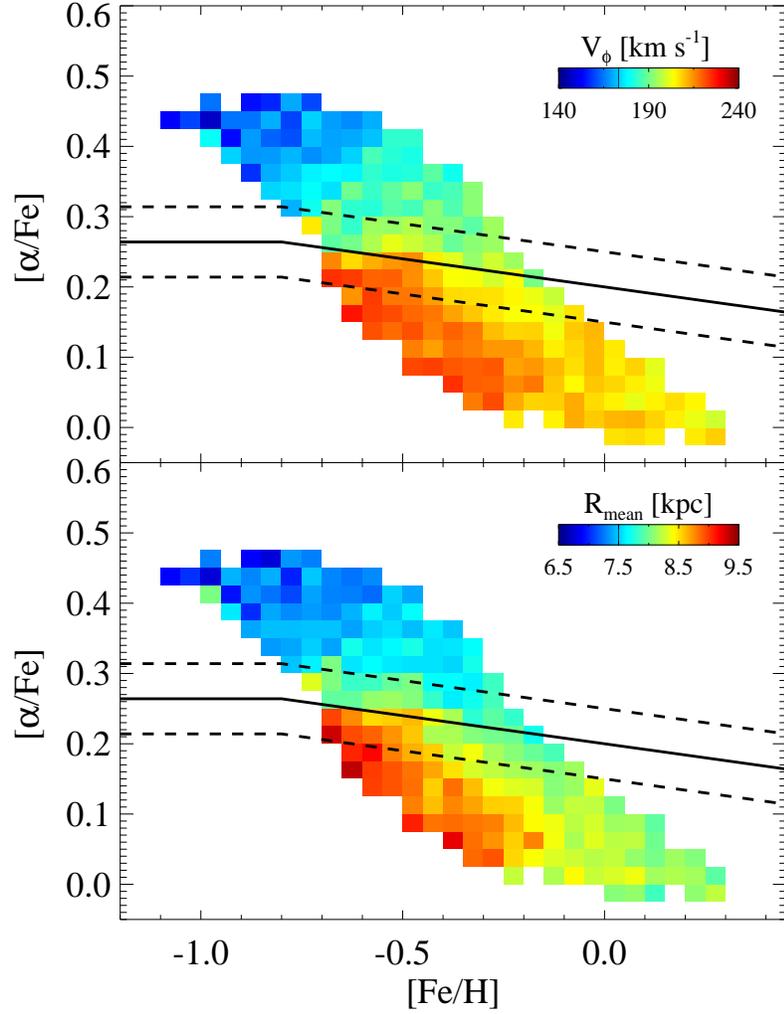
Figure 9: Distribution of mean rotational velocities (v_Φ, top panel) and the orbital radii (R_{mean}, bottom panel) for G dwarf sample from (Lee et al. 2011b, their Figure 5) the $[\alpha/Fe]$ vs. $[Fe/H]$ diagram (3σ-clipped mean values). Orbital parameters are computed using an analytic Stäckel-type gravitational potential from Chiba & Beers (2000). Note the rich structure in both panels.
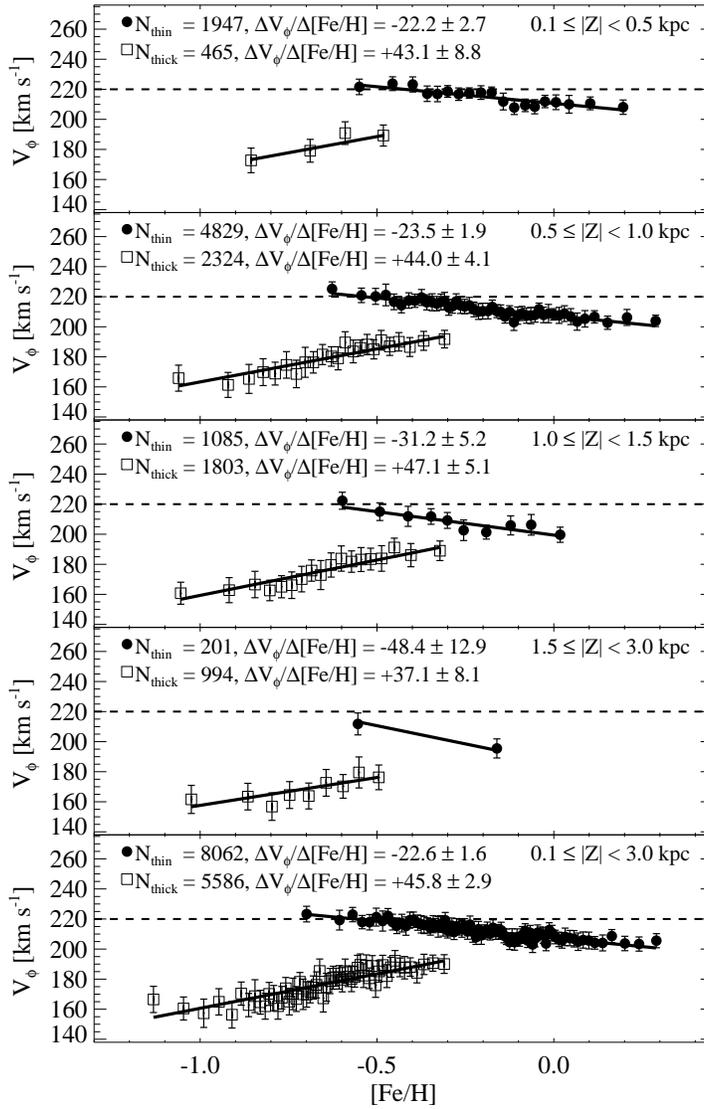
Figure 10: The variation of the mean rotational velocity with metallicity for different slices in distance from the Galactic plane (top four panels), for stars separated using $[\alpha/Fe]$ the thin-disk (black dots) and thick-disk (open squares) populations (Figure 7 from Lee et al. 2011b). Each dot represents a 3σ-clipped average of 100 stars. The bottom panel shows the results for the full samples of stars considered. Estimates of the slopes and their errors listed in the panels are computed for unbinned data.
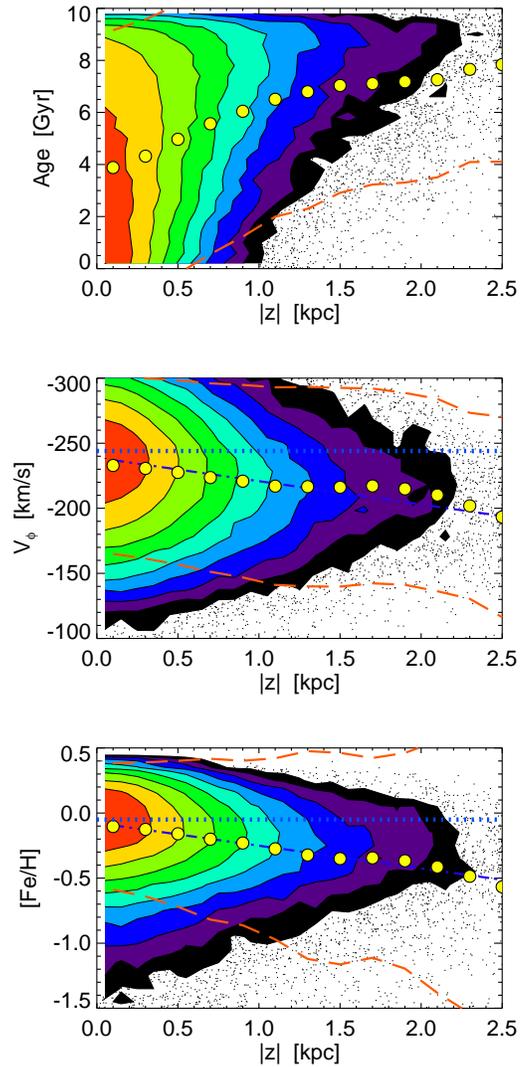
Figure 11: Predictions of the radial migration model from Roškar et al. (2008b) for the variation of stellar age, rotational velocity, and metallicity with distance from the Galactic plane for stars in the solar cylinder (Figure 8 from Loebman et al. (2011)). The simulated distributions are represented by color-coded contours (low to medium to high: black to green to red) in the regions of high density, and as individual points otherwise. The large symbols show the means for $|Z|$ bins, and the dashed lines show a 2σ envelope. The gradients seen in the bottom two panels are consistent with the SDSS-based results.
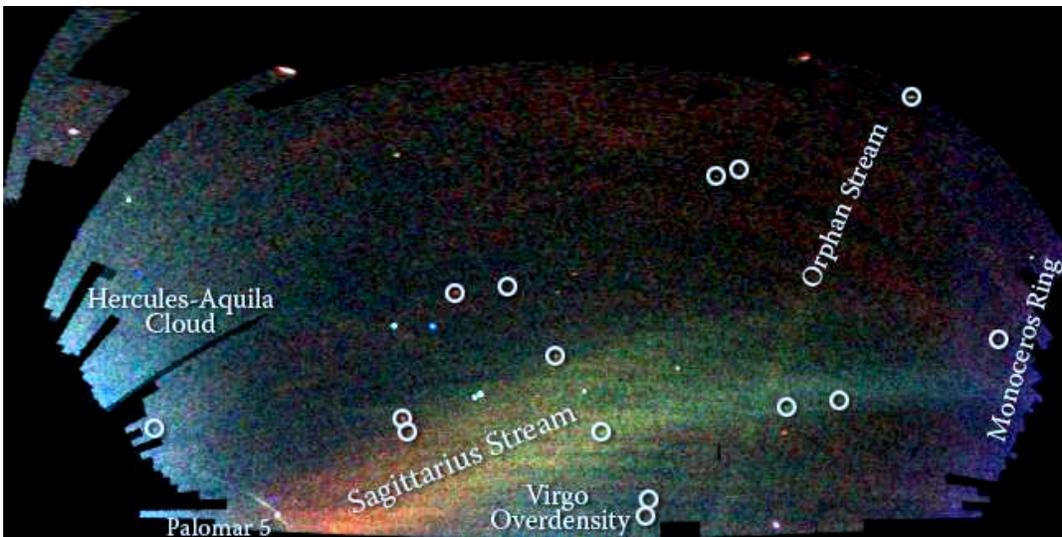
Figure 12: A map of stars in the outer regions of the Milky Way galaxy, derived from the SDSS images of the northern sky, shown in a Mercator-like projection. The color indicates the distance of the stars, while the intensity indicates the density of stars on the sky. There are several structures visible in this map, as marked, that demonstrate the halo is not a smooth structure. Circles enclose new Milky Way companions discovered by the SDSS; two of these are faint globular star clusters, while the others are faint dwarf galaxies.

metal-poor, 8 < D/kpc < 10, <μ_L>   metal-poor, 8 < D/kpc < 10, <μ_B>

−6 −4 −2 0   2   4   6

−3 −2 −1 0   1   2   3

metal-poor, 8 < D/kpc < 10, μ_L: <residuals>   metal-poor, 8 < D/kpc < 10, μ_B: <residuals>
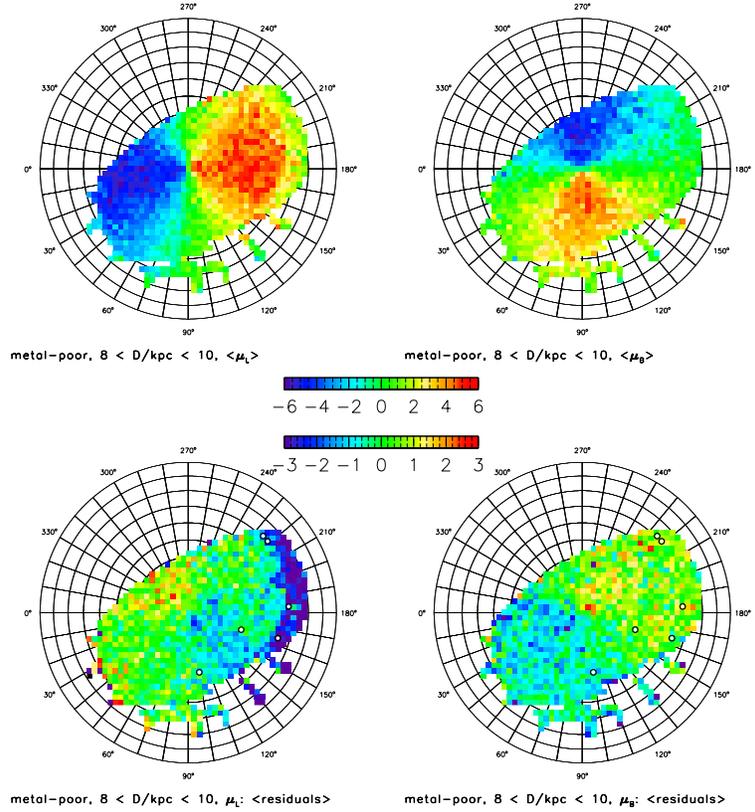
Figure 13: Figure 21 from Bond et al. (2010). Distribution of the median longitudinal proper motion in a Lambert projection of the North Galactic cap for low-metallicity (spectroscopic $[Fe/H] < -1.1$), blue $(0.2 < gr < 0.4)$ stars, with distances in the range 8-10 kpc. The top two panels show the median longitudinal (left) and latitudinal (right) proper motions, and the two bottom panels show the median difference between the observed and model-predicted values. The maps are color-coded according to the legends in the middle (mas yr^{-1}); note that the bottom scale has a harder stretch to emphasize structure in the residual maps). In the bottom panels, the white symbols show the positions of the six northern cold substructures identified by Schlaufman et al. (2009).
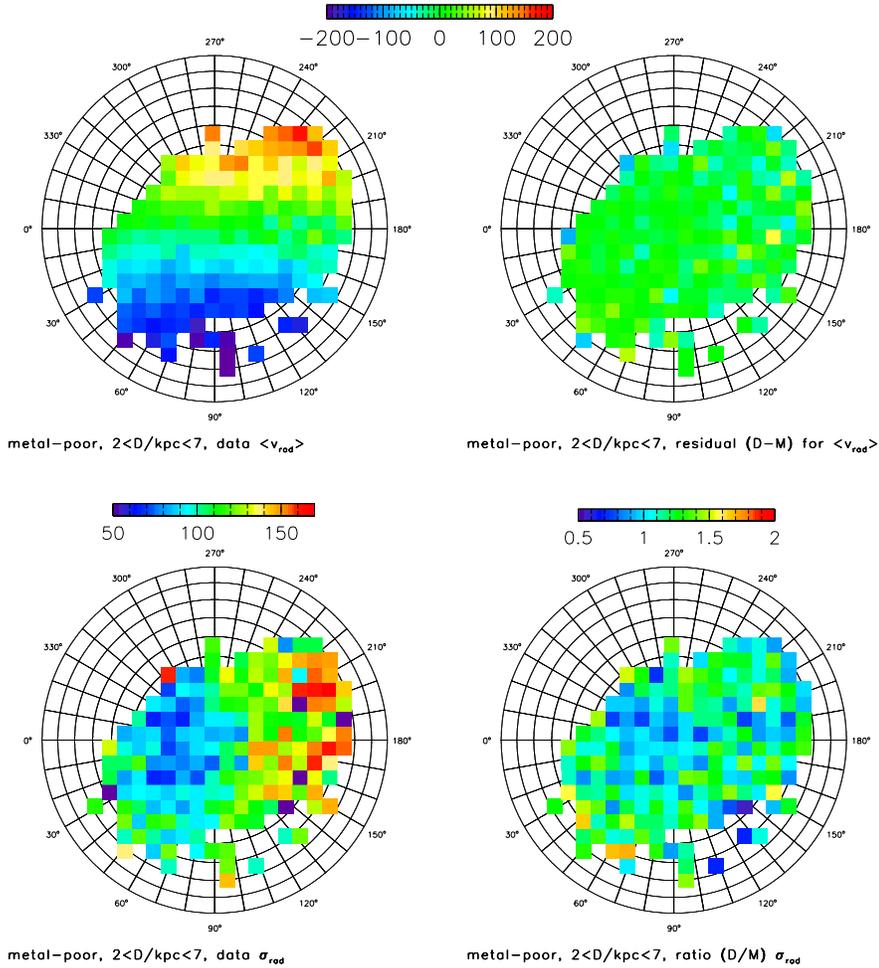
Figure 14: Figure 16 from Bond et al. (2010). Comparison of medians and dispersions for the measured and modeled radial velocities of 20,000 blue ($0.2 < gr < 0.4$) halo stars (spectroscopic $[Fe/H] < 1.1$) at distances, $D = 27$ kpc, and $b > 20°$. The top-left panel shows the median measured radial velocity in each pixel, color-coded according to the legend shown at the top (units are km s^{-1}). The top-right panel shows the difference between this map and an analogous map based on model-generated values of radial velocity, using the same scale as in the top-left panel. The bottom-left panel shows the dispersion of measured radial velocities, color-coded according to the legend above it. The bottom-right panel shows the ratio of this map and an analogous map based on model-generated values of radial velocity, color-coded according to the legend above it.
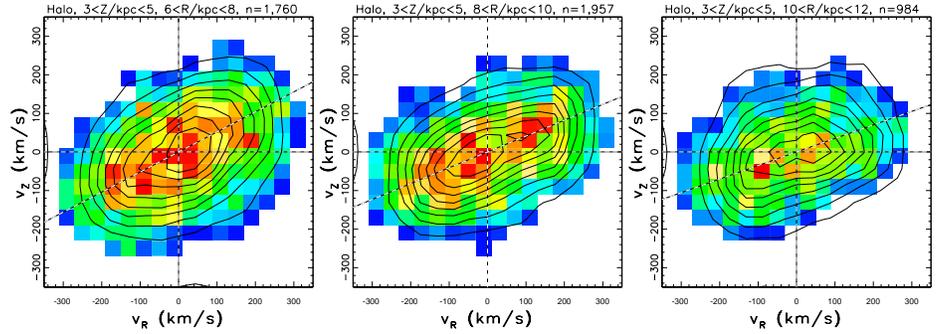
Figure 15: Figure 13 from Bond et al. (2010). The two-dimensional v_Z vs. v_R projections of the velocity distribution for three subsamples of candidate halo stars selected using spectroscopic metallicity ($3 < [Fe/H] < 1.1$), with $3 < Z/\text{kpc} < 5$, and $6 < R/\text{kpc} < 8$ (left), $8 < R/\text{kpc} < 10$ (middle), and $10 < R/\text{kpc} < 12$ (right). The distributions are shown using linearly spaced contours, and with a color-coded map showing smoothed counts in pixels (low to high from blue to red). The measurement errors are typically 60 kms^{-1}, and the dashed lines show the median direction toward the Galactic center. Note the strong evidence for a velocity-ellipsoid tilt, and the variation of the tilt with R so that the ellipsoid always points towards the Galactic center.
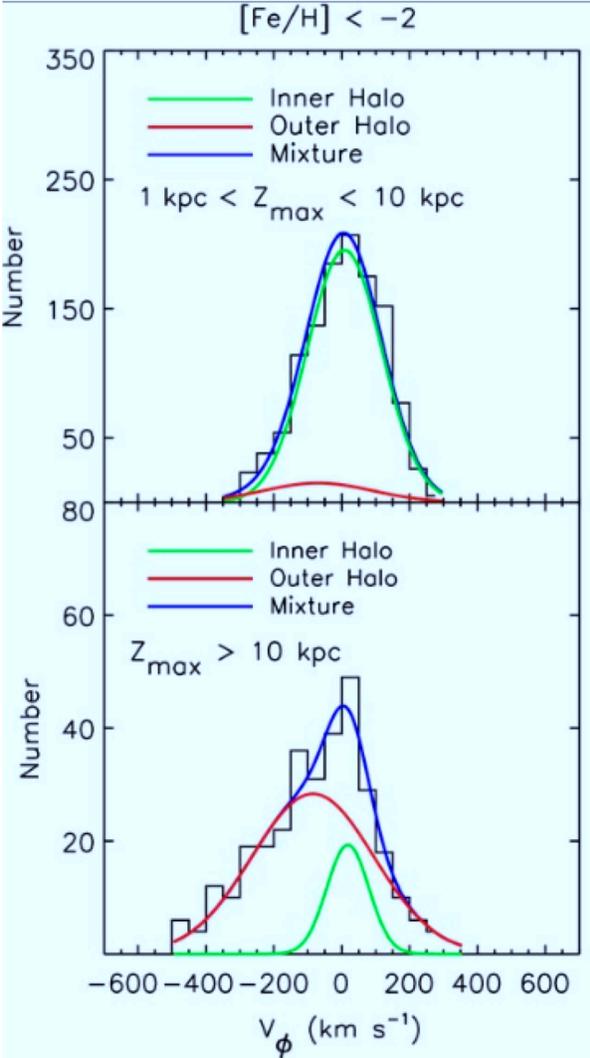
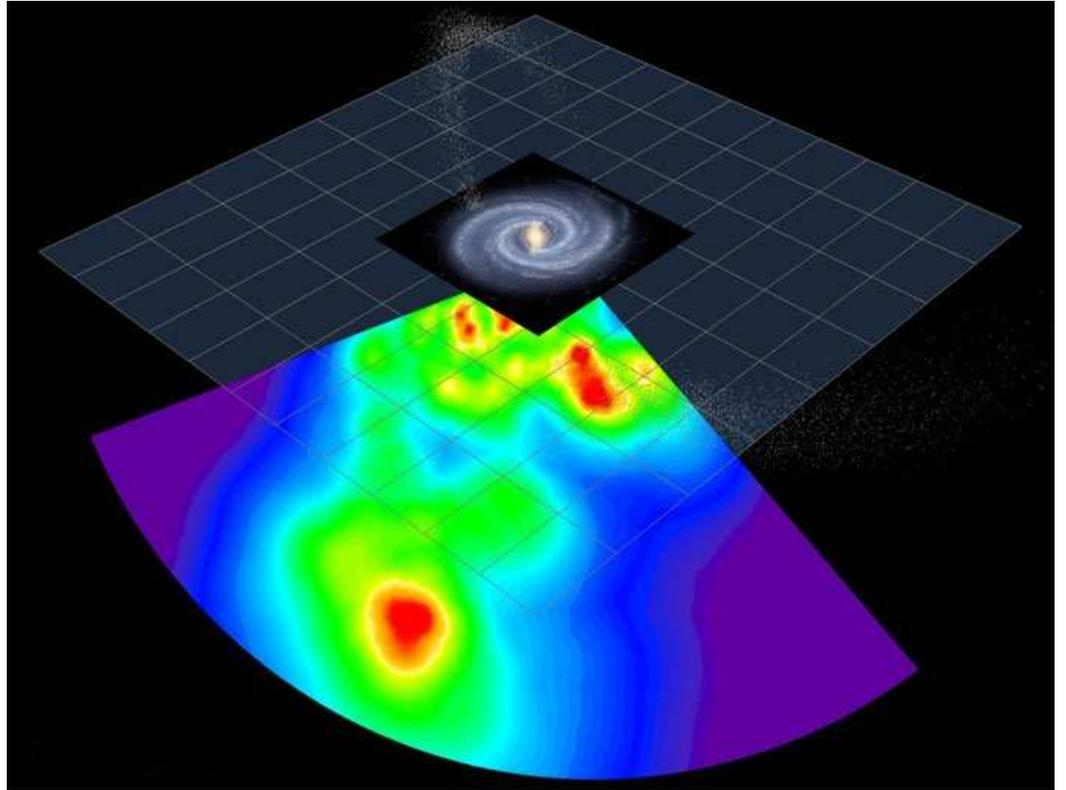Figure 16: Figure 10 from Carollo et al.

Figure 17: Figure 12 from Sesar et al. (2010). The stripe 82 plane and artists concept of the disk plane. The white dots show the Sagittarius dSph and its tidal streams, as modeled by the Law, Johnston & Majewski (2005) "spherical" model.
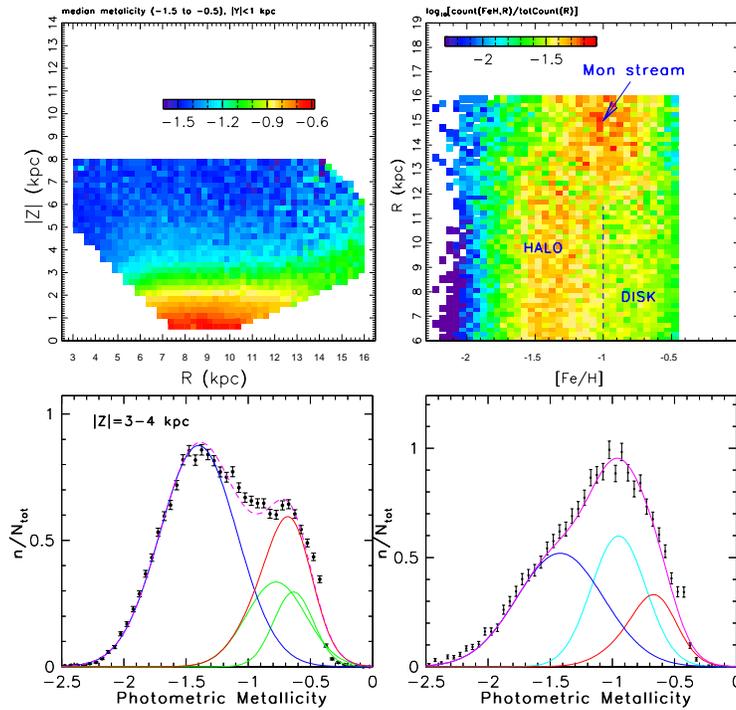
Figure 18: Figure 18 from Ivezić et al. (2008b). Top left panel: Dependence of the median photometric metallicity for ~1.04 million stars with $14.5 < r < 20$, $0.2 < g-r < 0.4$, and $|Y| < 1$ kpc, in cylindrical Galactic coordinates R and $|Z|$. This Y range is selected to include the Monoceros stream, which represents an overdensity by a factor of ~1.52 in a region around $R \sim 15$ kpc and $|Z| \sim 3 - 4$ kpc. As discernible from the map, this region has a larger median metallicity than expected for this $|Z|$ range based on extrapolation from smaller R. Top right panel: Conditional metallicity probability distribution for a subsample of ~111,000 stars with $3 < |Z|/\mathrm{kpc} < 4$. The strong overdensity at $R > 12$ kpc is the Monoceros stream. The bottom panels show the metallicity distribution (symbols with error bars) for a subsample of ~40,000 stars with $6 < R/\mathrm{kpc} < 9$ (left) and for ~12,000 stars with $13 < R/\mathrm{kpc} < 16$ (right). The lines represent empirical fits discussed Ivezić et al. (2008b). The cyan line in the bottom right panel is a 0.22 dex wide Gaussian centered on $[Fe/H] = -0.95$. It accounts for 33% of stars in the sample that presumably belong to the Monoceros stream.

6.5   New Analysis Challenges for Massive Datasets

Talk about Massive Data Mining, Knowledge Discovery, and Data Visualization.

References

1. Abazajian KN, et al. 2009. *Astrophysical Journal Supplement* 182:543–558

2. Allende Prieto C, Sivarani T, Beers TC, Lee YS, Koesterke L, et al. 2008. *Astronomical Journal* 136:2070–2082

3. Bahcall JN, Soneira RM. 1980. *Astrophysical Journal Supplement* 44:73–110

4. Beers TC, Lee Y, Sivarani T, Allende Prieto C, Wilhelm R, et al. 2006. *Mem.S.A.It.* 77:1171–+

5. Bell EF, Zucker DB, Belokurov V, Sharma S, Johnston KV, et al. 2008. *Astrophysical Journal* 680:295–311

6. Belokurov V, Evans NW, Bell EF, Irwin MJ, Hewett PC, et al. 2007a. *Astrophysical Journal Letters* 657:L89–L92

7. Belokurov V, Evans NW, Irwin MJ, Lynden-Bell D, Yanny B, et al. 2007b. *Astrophysical Journal* 658:337–344

8. Belokurov V, Zucker DB, Evans NW, Gilmore G, Vidrih S, et al. 2006. *Astrophysical Journal Letters* 642:L137–L140

9. Belokurov V, Zucker DB, Evans NW, Kleyna JT, Koposov S, et al. 2007c. *Astrophysical Journal* 654:897–906

10. Bond NA, Ivezić Ž, Sesar B, Jurić M, Munn JA, et al. 2010. *Astrophysical Journal* 716:1–29

11. Brook C, Richard S, Kawata D, Martel H, Gibson BK. 2007. *Astrophysical Journal* 658:60–64

12. Brook CB, Gibson BK, Martel H, Kawata D. 2005. *Astrophysical Journal* 630:298–308

13. Bullock JS, Johnston KV. 2005. *Astrophysical Journal* 635:931–949

14. Bullock JS, Kravtsov AV, Weinberg DH. 2000. *Astrophysical Journal* 539:517–521

15. Bullock JS, Kravtsov AV, Weinberg DH. 2001. *Astrophysical Journal* 548:33–46

16. Burnett B, Binney J, Sharma S, Williams M, Zwitter T, et al. 2011. *Astronomy & Astrophysics* 532:A113+

17. Casetti-Dinescu DI, Girard TM, Korchagin VI, van Altena WF. 2011. *Astrophysical Journal* 728:7–+

18. Chiba M, Beers TC. 2000. *Astronomical Journal* 119:2843–2865

19. Covey KR, Ivezić Ž, Schlegel D, Finkbeiner D, Padmanabhan N, et al. 2007. *Astronomical Journal* 134:2398–2417

20. Di Matteo P, Lehnert MD, Qu Y, van Driel W. 2011. *Astronomy & Astrophysics* 525:L3+

21. Dierickx M, Klement R, Rix HW, Liu C. 2010. *Astrophysical Journal Letters* 725:L186–L190

22. Eggen OJ, Lynden-Bell D, Sandage AR. 1962. *Astrophysical Journal* 136:748–+

23. Flaugher B. 2008. In *A Decade of Dark Energy: Spring Symposium, Proceedings of the conferences held May 5-8, 2008 in Baltimore, Maryland. (USA). Edited by Norbert Pirzkal and Henry Ferguson. http://www.stsci.edu/institute/conference/spring2008*

24. Freeman K, Bland-Hawthorn J. 2002a. *Annu. Rev. Astro. Astrophys.* 40:487–537

25. Freeman K, Bland-Hawthorn J. 2002b. *Annu. Rev. Astro. Astrophys.* 40:487–537

26. Gilmore G, Reid N. 1983. *Monthly Notices of the R.A.S.* 202:1025–1047

27. Gilmore G, Wyse RFG, Kuijken K. 1989. *Annu. Rev. Astro. Astrophys.* 27:555–627

28. Governato F, Mayer L, Wadsley J, Gardner JP, Willman B, et al. 2004. *Astrophysical Journal* 607:688–696

29. Governato F, Willman B, Mayer L, Brooks A, Stinson G, et al. 2007. *Monthly Notices of the R.A.S.* 374:1479–1494

30. Grillmair CJ. 2009. *Astrophysical Journal* 693:1118–1127

31. Grillmair CJ, Dionatos O. 2006. *Astrophysical Journal Letters* 643:L17–L20

32. Harding P, Morrison HL, Olszewski EW, Arabadjis J, Mateo M, et al. 2001. *Astronomical Journal* 122:1397–1419

33. Helmi A. 2008. *AAPR?* 15:145–188

34. Helmi A, White SDM. 1999. *Monthly Notices of the R.A.S.* 307:495–517

35. Ivezić Ž, Goldston J, Finlator K, Knapp GR, Yanny B, et al. 2000. *Astronomical Journal* 120:963–977

36. Ivezić Z, et al. 2008a. *ArXiv:0805.2366*

37. Ivezić Ž, et al. 2008b. *Astrophysical Journal* 684:287–325

38. Johnston KV, Bullock JS, Sharma S, Font A, Robertson BE, Leitner SN. 2008. *Astrophysical Journal* 689:936–957

39. Johnston KV, Hernquist L, Bolte M. 1996. *Astrophysical Journal* 465:278–+

40. Jurić M, et al. 2008. *Astrophysical Journal* 673:864–914

41. Kaiser N, et al. 2002. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, eds. JA Tyson, S Wolff, vol. 4836 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*

42. Klement RJ. 2010. *AAPR?* 18:567–594

43. Law DR, Johnston KV, Majewski SR. 2005. *Astrophysical Journal* 619:807–823

44. Lee YS, Beers TC, Allende Prieto C, Lai DK, Rockosi CM, et al. 2011a. *Astronomical Journal* 141:90–+

45. Lee YS, Beers TC, An D, Ivezic Z, Just A, et al. 2011b. *ArXiv e-prints*

46. Lee YS, Beers TC, Sivarani T, Allende Prieto C, Koesterke L, et al. 2008a. *Astronomical Journal* 136:2022–2049

47. Lee YS, Beers TC, Sivarani T, Johnson JA, An D, et al. 2008b. *Astronomical Journal* 136:2050–2069

48. Loebman SR, Roškar R, Debattista VP, Ivezić Ž, Quinn TR, Wadsley J. 2011. *Astrophysical Journal* 737:8–+

49. Majewski SR. 1992. *Astrophysical Journal Supplement* 78:87–152

50. Majewski SR. 1993. *Annu. Rev. Astro. Astrophys.* 31:575–638

51. Majewski SR. 2010. In *IAU Symposium*, ed. G. Bruzual & S. Charlot, vol. 262 of *IAU Symposium*

52. Majewski SR, Skrutskie MF, Weinberg MD, Ostheimer JC. 2003. *Astrophysical Journal* 599:1082–1115

53. Mayer L, Moore B, Quinn T, Governato F, Stadel J. 2002. *Monthly Notices of the R.A.S.* 336:119–130

54. Minchev I, Famaey B. 2010. *Astrophysical Journal* 722:112–121

55. Monet DG, Levine SE, Canzian B, Ables HD, Bird AR, et al. 2003. *Astronomical Journal* 125:984–993

56. Munn JA, et al. 2004. *Astronomical Journal* 127:3034–3042

57. Newberg HJ, Yanny B, Cole N, Beers TC, Re Fiorentin P, et al. 2007. *Astrophysical Journal* 668:221–235

58. Newberg HJ, Yanny B, Rockosi C, Grebel EK, Rix HW, et al. 2002. *Astrophysical Journal* 569:245–274

59. Nordström B, Mayor M, Andersen J, Holmberg J, Pont F, et al. 2004. *Astronomy & Astrophysics* 418:989–1019

60. Perryman MAC. 2002. *Astronomy & Astrophysics Supp. Ser.* 280:1–10

61. Perryman MAC, Lindegren L, Kovalevsky J, Hoeg E, Bastian U, et al. 1997. *Astronomy & Astrophysics* 323:L49–L52

62. Perryman MAC, et al. 2001. *Astronomy & Astrophysics* 369:339–363

63. Pont F, Eyer L. 2004. *Monthly Notices of the R.A.S.* 351:487–504

64. Pourbaix D, Knapp GR, Szkody P, Ivezić Ž, Kleinman SJ, et al. 2005. *Astronomy & Astrophysics* 444:643–649

65. Robin AC, Reylé C, Derrière S, Picaud S. 2003. *Astronomy & Astrophysics* 409:523–540

66. Roškar R, Debattista VP, Quinn TR, Stinson GS, Wadsley J. 2008a. *Astrophysical Journal Letters* 684:L79–L82

67. Roškar R, Debattista VP, Stinson GS, Quinn TR, Kaufmann T, Wadsley J. 2008b. *Astrophysical Journal Letters* 675:L65–L68

68. Sales LV, Helmi A, Abadi MG, Brook CB, Gómez FA, et al. 2009. *Monthly Notices of the R.A.S.* 400:L61–L65

69. Schlaufman KC, et al. 2009. *ArXiv e-prints*

70. Schönrich R, Binney J. 2009a. *Monthly Notices of the R.A.S.* 396:203–222

71. Schönrich R, Binney J. 2009b. *Monthly Notices of the R.A.S.* 399:1145–1156

72. Searle L, Zinn R. 1978. *Astrophysical Journal* 225:357–379

73. Sellwood JA, Binney JJ. 2002. *Monthly Notices of the R.A.S.* 336:785–796

74. Sesar B, Ivezić Ž, Grammer SH, Morgan DP, Becker AC, et al. 2010. *Astrophysical Journal* 708:717–741

75. Sesar B, Svilković D, Ivezić Ž, Lupton RH, Munn JA, et al. 2006. *Astronomical Journal* 131:2801–2825

76. Siebert A, Williams MEK, Siviero A, Reid W, Boeche C, et al. 2011. *Astronomical Journal* 141:187–+

77. Skrutskie MF, Cutri RM, Stiening R, Weinberg MD, Schneider S, et al. 2006. *Astronomical Journal* 131:1163–1183

78. Soderblom DR. 2010. *Annu. Rev. Astro. Astrophys.* 48:581–629

79. Sommer-Larsen J, Götz M, Portinari L. 2003. *Astrophysical Journal* 596:47–66

80. Steinmetz M, Navarro JF. 2002. *New Astronomy* 7:155–160

81. Steinmetz M, Zwitter T, Siebert A, Watson FG, Freeman KC, et al. 2006. *Astronomical Journal* 132:1645–1668

82. van der Kruit PC, Freeman KC. 2011. *Annu. Rev. Astro. Astrophys.* 49:301–371

83. Vivas AK, Zinn R. 2006. *Astronomical Journal* 132:714–728

84. Wilkinson MI, et al. 2005. *Monthly Notices of the R.A.S.* 359:1306–1335

85. Wilson ML, Helmi A, Morrison HL, Breddels MA, Bienaymé O, et al. 2011. *Monthly Notices of the R.A.S.* 413:2235–2241

86. Xue XX, Rix HW, Yanny B, Beers TC, Bell EF, et al. 2011. *Astrophysical Journal* 738:79–+

87. Yanny B, Newberg HJ, Kent S, Laurent-Muehleisen SA, Pier JR, et al. 2000. *Astrophysical Journal* 540:825–841

88. Yanny B, Rockosi C, Newberg HJ, Knapp GR, Adelman-McCarthy JK, et al. 2009. *Astronomical Journal* 137:4377–4399

89. Yoachim P, Dalcanton JJ. 2008. *Astrophysical Journal* 682:1004–1019

90. York DG, et al. 2000. *Astronomical Journal* 120:1579–1587

91. Zwitter T, Matijevič G, Breddels MA, Smith MC, Helmi A, et al. 2010. *Astronomy & Astrophysics* 522:A54+
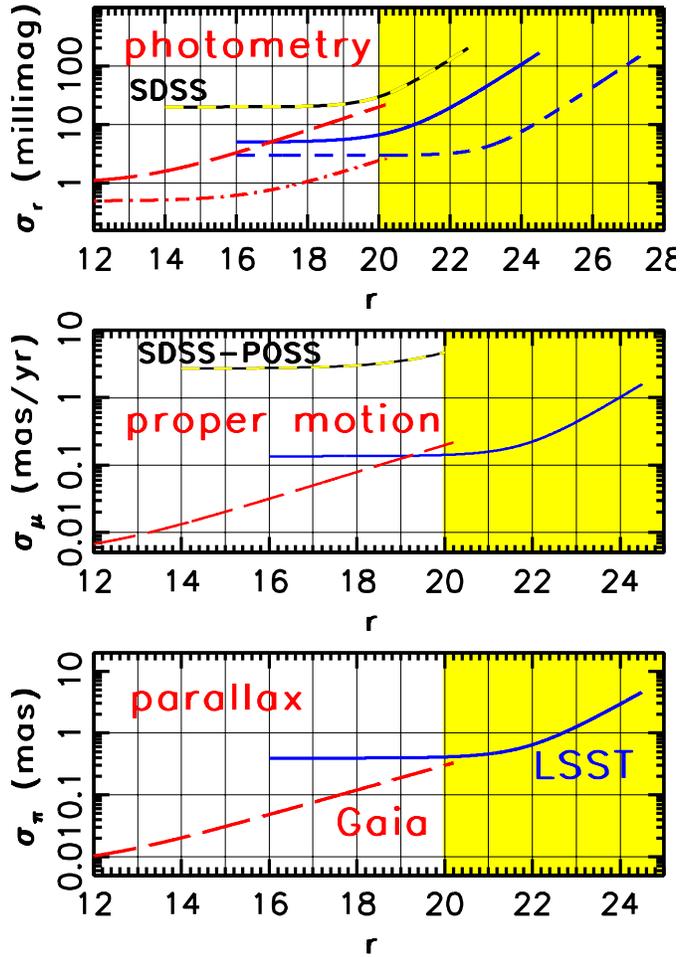
Figure 19: A comparison of photometric, proper motion and parallax errors for SDSS, Gaia and LSST, as a function of apparent magnitude r, for a G2V star (Eyer et al, in prep). In the top panel, the curve marked "SDSS" corresponds to a single SDSS observation. The red curves correspond to Gaia; the long-dashed curve shows a single *transit* accuracy, and the dot-dashed curve the end of mission accuracy (assuming 70 transits). The blue curves correspond to LSST; the solid curve shows a single *visit* accuracy, and the short-dashed curve shows accuracy for co-added data (assuming 230 visits in the r band). The curve marked "SDSS-POSS" in the middle panel shows accuracy delivered by the proper motion catalog of Munn et al. (2004).