Preprint typeset using LaTeX style emulateapj v. 5/25/10

KINEMATIC MODELLING OF THE MILKY WAY USING THE RAVE AND GCS STELLAR SURVEYS

S. SHARMA [1], J. BLAND-HAWTHORN [1], J. BINNEY [2], K. C. FREEMAN [3], M. STEINMETZ [4], C. BOECHE [5], O. BIENAYMÉ [6], B. K. GIBSON [7], G. F. GILMORE [8], E. K. GREBEL [5], A. HELMI [9], G. KORDOPATIS [8], U. MUNARI [10], J. F. NAVARRO [11], Q. A. PARKER [12], W. A. REID [12], G. M. SEABROKE [13], A. SIEBERT [6], F. WATSON [14], M. E. K. WILLIAMS [4], R. F. G. WYSE [15] T. ZWITTER [16]

ABSTRACT

We investigate the kinematic parameters of the Milky Way disc using the Radial Velocity (RAVE) and Geneva-Copenhagen (GCS) stellar surveys. We do this by fitting a kinematic model to the data taking the selection function of the data into account. Using Markov Chain Monte Carlo (MCMC) technique, we investigate the full posterior distributions of the parameters given the data. We investigate the 'age-velocity dispersion' relation (AVR) for the three kinematic components ($\sigma_R, \sigma_\phi, \sigma_z$), the radial dependence of the velocity dispersions, the Solar peculiar motion ($U_\odot, V_\odot, W_\odot$), the circular velocity (v_0) at the Sun and its fall with height above the mid-plane. For RAVE, our analysis uses *only* the angular position and radial velocity of stars since these quantities are determined to high accuracy. We confirm that the Besançon-style Gaussian model accurately fits the GCS data, but fails to match the details of the more spatially extended RAVE survey. In particular, our use of the Shu distribution function handles non-circular orbits more accurately and provides a better fit to the kinematic data. The Gaussian distribution function not only fits the data poorly but systematically underestimates the fall of velocity dispersion with radius. The measured Solar peculiar motion, $U_\odot = 11.0 \pm 0.2, V_\odot = 7.5 \pm 0.2$, and $W_\odot = 7.5 \pm 0.1$, is in good agreement with recent estimates except for value of $V_\odot$ which we find to be lower. We find the circular velocity to be $v_0 = 232 \pm 2$ km s^{-1} (with $R_\odot = 8$ kpc) or $\Omega_\odot = 29.9 \pm 0.3\,\mathrm{km\,s^{-1}\,kpc^{-1}}$ which is in good agreement with the Sgr A* proper motion. If we explicitly adopt the Sgr A* proper motion as a prior, we find the gradient of circular velocity to be $0.65 \pm 0.2\,\mathrm{km\,s^{-1}\,kpc^{-1}}$. Additional systematic uncertainties in $V_\odot$ and v_0 of the order of those in the adopted priors on spatial distribution of stars are expected. We find that, for an extended sample of stars, v_0 is underestimated if the vertical dependence of the effective circular velocity is neglected. **Proper treatment of the third integral of motion still remains an issue.** We also find that the radial scale length of the velocity dispersion profile of the thick disc is smaller than that of the thin disc.

Subject headings: galaxies:kinematics and dynamics – fundamental parameters – formation – methods: data analysis – numerical – statistical

1. INTRODUCTION

Understanding the origin and evolution of disc galaxies is one of the major goals of modern astronomy. The disc is a prominent feature of late type galaxies like the Milky Way. As compared to distant galaxies, for which one can only measure the gross properties, the Milky Way offers the opportunity to study the disc in great detail. For the Milky Way, we can determine 6-dimensional phase space information, combined with photometric and stellar parameters, for a huge sample of stars. This has led to large observational programs to catalog the stars in the Milky Way in order to compare them with theoretical models.

The Milky Way stellar system is broadly composed of four distinct parts although in reality there is likely to be considerable overlap between them: the thin disc, the thick disc, the stellar halo and the bulge. In this paper, we mainly concentrate on understanding the disc components which are the dominant stellar populations. In the Milky Way, the thick disc was originally identified as the second exponential required to fit vertical star counts (Gilmore & Reid 1983; Reid & Majewski 1993; Jurić et al. 2008). Thick discs are also ubiquitous features of late type galaxies (Yoachim & Dalcanton 2006). But whether the thick disc is a separate component with a distinct formation mechanism is highly debatable and a difficult question to answer.

Since the Gilmore & Reid (1983) result, various attempts have been made to characterize the thick disc. Some studies suggest that thick disc stars have distinct properties: they are

[1] Sydney Institute for Astronomy, School of Physics, University of Sydney, NSW 2006, Australia
[2] Rudolf Pierls Center for Theoretical Physics, University of Oxford, 1 Keble Road, Oxford OX1 3NP, UK
[3] RSAA Australian National University, Mount Stromlo Observatory, Cotter Road, Weston Creek, Canberra, ACT 72611, Australia
[4] Leibniz Institut für Astrophysik Potsdam (AIP), An der Sterwarte 16, D-14482 Potsdam, Germany
[5] Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, D-69120 Heidelberg, Germany
[6] Observatoire astronomique de Strasbourg, Université de Strasbourg, CNRS, UMR 7550, Strasbourg, France
[7] Jeremiah Horrocks Institute for Astrophysics & Super-computing, University of Central Lancashire, Preston, UK
[8] Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
[9] Kapteyn Astronomical Institute, University of Groningen, Postbus 800, 9700 AV Groningen, Netherlands
[10] INAF - Astronomical Observatory of Padova, 36012 Asiago (VI), Italy
[11] University of Victoria, P.O. Box 3055, Station CSC, Victoria, BC V8W 3P6, Canada
[12] Department of Physics and Astronomy, Macquarie University, Sydney, NSW 2109, Australia
[13] Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, RH5 6NT, UK
[14] Australian Astronomical Observatory, PO Box 296, Epping, NSW 1710, Australia
[15] Johns Hopkins University, 3400 N Charles Street, Baltimore, MD 21218, USA
[16] Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, Ljubljana, Slovenia

old and metal poor (Chiba & Beers 2000) and α enhanced (Fuhrmann 1998; Bensby et al. 2005, 2003). Chemical evolution models by Chiappini et al. (1997) advocate a hiatus in star formation to explain this. Jurić et al. (2008) fit the SDSS star counts using a two component model and find that the thick disc has a larger scale-length than the thin disc. In contrast, Bovy et al. (2012d) using SDSS and SEGUE data find the opposite when they associate the thick disc with the α-enhanced component. Finally, the idea of a separate thick disc has recently been challenged. Schönrich & Binney (2009b,a) showed that chemical evolutionary models with radial migration and mixing can replicate the properties of the thick disc (see also Loebman et al. 2011, that explore radial mixing using N-body simulations). Ivezić et al. (2008) do not find the expected separation between metallicity and kinematics for F, G stars in the SDSS survey, and Bovy et al. (2012b,c) argue that the thick disc is a smooth continuation of the thin disc.

Opinions regarding the formation of a thick disc are equally divided. Various mechanisms have been proposed: accretion of stars from disrupted galaxies (Abadi et al. 2003), heating of discs by minor mergers (Quinn et al. 1993; Kazantzidis et al. 2008, 2009; Villalobos & Helmi 2008; Di Matteo et al. 2011), radial migration of stars by spiral arms (Schönrich & Binney 2009b,a; Loebman et al. 2011), a gas-rich merger at high redshift (Brook et al. 2004), and gravitationally unstable disc evolution (Bournaud et al. 2009), inter alia. Recently, Forbes et al. (2012) have suggested that the thick disc can form without secular heating, mainly because stars forming at higher redshift had a higher velocity dispersion to begin with. Another possibility as proposed by Roškar et al. (2010) is misaligned angular momentum of in-falling gas. How the angular momentum of halo gas becomes misaligned is described in Sharma et al. (2012). However, Aumer & White (2013) and Sales et al. (2012) suggest that misaligned gas can destroy the discs.

The only way to test the different thick disc theories is to compare the kinematic and chemical abundance distributions of the thick disc stars with those of different models. Since, the thin and thick disc stars strongly overlap both in space and kinematics, it is difficult to separate them using just position and velocity. To really isolate and study the thick disc, one needs a tag that is unique to its formation mechanism. Age is a possible tag but it is difficult to get reliable age estimates of stars. Chemical composition is another promising tag that can be used, but this requires high resolution spectroscopy of a large number of stars. In near future, surveys like GALAH using the HERMES spectrograph should be able to fill this void (Freeman & Bland-Hawthorn 2008). In our first analysis, we restrict ourselves to a differential kinematic study of the disc components. We plan to treat the more difficult problem of chemo-dynamics in future.

The simplest way to describe the kinematics of the Milky Way stars in the Solar neighborhood is by using Gaussian distribution functions. If a single component disc is used, then in cylindrical co-ordinates only three components of velocity dispersion σ_R, σ_ϕ and σ_z and the mean azimuthal velocity $\overline{v_\phi}$ need to be known [17]. If a thick disc is included, one requires 5 additional parameters, one of them being the fraction of stars in thick disc. If stars are sampled from an extended volume and not just the Solar neighborhood, then one needs to specify the radial dependence of the dispersions. The velocity dispersion of a disc stellar population is known to increase with age. Spitzer & Schwarzschild (1953) showed that scattering of stars by gas clouds can cause an increase in velocity dispersion with age.[18] To account for this, one has to model the thin disc with an age velocity dispersion relation, traditionally assumed to be a power law (although see Edvardsson et al. 1993; Quillen & Garnett 2001; Seabroke & Gilmore 2007). The exponents β_R, β_ϕ and β_z of these power laws may not be the same for all three components.

The ratio of σ_z/σ_R and σ_ϕ/σ_R and the values of β_R, β_ϕ and β_z are useful for understanding the physical processes responsible for heating the disc. Hänninen & Flynn (2002) showed that one gets different predictions depending upon how the perturbers are distributed in space, e.g., in plane or in a sphere. Another way to heat up the disc is by Lindblad resonances of transient spiral arms that can scatter the stars. This process specifically increases the in-plane dispersions (Carlberg & Sellwood 1985; Sellwood 2013). Multiple spiral density waves Minchev & Quillen (2006) or a combination of bar and spirals can also heat up the disc (Minchev & Famaey 2010).

In reality, multiple processes work together to heat up the disc. Ideally, one would like to study the disc evolution using fully self-consistent cosmological simulations, where the relative importance of different processes can be studied. Current cosmological simulations of Milky Way-sized galaxies predict velocity dispersions that are too high as compared to observations (House et al. 2011). Also, a floor in dispersion profiles is observed which is related to the density threshold used for star formation in the simulations. Recently, Minchev et al. (2013) investigate the age-velocity dispersion relation (AVR) for stars in simulations of disc galaxies and find it to be in rough agreement with observations. Disconcertingly, Sellwood (2013) has shown that 2-body relaxation can artificially increase the vertical dispersion and this may have affected many published simulations.

While simulations are the most accurate way to understand disc galaxies, it is not feasible to run multiple simulations to fit the observational data. Hence, fitting analytical models to data is important. Characterizing the properties of the Milky Way disc by fitting a suitable analytical model helps to summarize large amounts of data but its usefulness extends beyond this. The more elaborate the model, and the more physical processes it captures, the better it can help us to understand the formation of the disc galaxies. McMillan & Binney (2012) discuss in detail the virtues of fitting equilibrium analytical models to data.

The first generation of stellar population models characterized the density distribution of stars using photometric surveys. The earliest such attempt was by Bahcall & Soneira (1980a,b, 1984) where they assumed an exponential disc with magnitude-dependent scale heights. An evolutionary model using population synthesis techniques was presented by Robin & Creze (1986). Given a star formation rate (SFR) and an initial mass function (IMF), one calculates the resulting stellar populations using theoretical evolutionary tracks. The important step forward was that the properties of the disc, like scale height, density laws and velocity dispersions, were assumed to be a function of age rather than being color-magnitude dependent terms. Bienayme et al. (1987) later introduced dynamical self-consistency to link disc scale and

[17] In reality a knowledge of the full velocity dispersion tensor is required. The cross terms can be ignored only for stars close to mid plane

[18] Strictly speaking, the increased dispersion comes from in-plane scattering processes; the clouds serve to provide the vertical dispersion by redirecting the stars (Sellwood 2013).

vertical velocity dispersions via the gravitational potential. Haywood et al. (1997a,b) further improved the constraints on SFR and IMF of the disc. The present state of the art is described in Robin et al. (2003) and is known as the Besançon model. Here, the disc is constructed from a set of isothermal populations that are assumed to be in equilibrium. Analytic functions for density distributions, the age/metallicity relation and the IMF are provided for each population. A similar scheme is also used by the codes TRILEGAL Girardi et al. (2005) and *Galaxia* (Sharma et al. 2011).

There is a crucial distinction between kinematic and dynamical models. In a kinematic model, one specifies the stellar motions according to a simple analytic formula based on local position; in a dynamical model, the spatial density distribution of stars and their kinematics are self-consistently linked by the potential in which the stars move, with the assumption that the system is in steady state. The action and angle formalism is especially well suited for constructing dynamical models. Here, the problem reduces from six dimensions of phase space to three dimensions of action space. This has led to torus modelling (McMillan & Binney 2008; Binney & McMillan 2011) where the distribution functions that are simple analytical functions of actions provide a powerful and simple way to construct dynamical models (Binney 2010, 2012b,a).

Today we have data from large photometric surveys like DENIS (Epchtein et al. 1999), 2MASS (Skrutskie et al. 2006) and SDSS (Abazajian et al. 2009). The SDSS survey in particular has been used to provide an empirical model of the Milky Way stars (Jurić et al. 2008; Ivezić et al. 2008; Bond et al. 2010). Since kinematic data for a large number of stars was not available at the time the Besançon model was constructed, the kinematic parameters of the model have not been tested as extensively as the star counts. The Hipparcos satellite (Perryman et al. 1997) and the UCAC2 catalog (Zacharias et al. 2004) provided proper motions and parallaxes for $\sim 10^5$ stars in the Solar neighborhood. Dehnen & Binney (1998b) used the Hipparcos data to study stellar kinematics as a function of color. They also determined the Solar motion with respect to the LSR and the axial ratios of the velocity ellipsoid. Binney et al. (2000) also using Hipparcos stars found the velocity dispersion to vary with function of age as $\tau^{0.33}$. More recently, Aumer & Binney (2009) using data from a new reduction of the Hipparcos mission estimated the Solar motion and the AVR for all three velocity components. The AVR is assumed to be a power law with exponents β_R, β_ϕ and β_z for the three velocity components in the galactocentric cylindrical coordinate system. They found $(\beta_R, \beta_\phi, \beta_z) = (0.30, 0.43, 0.44)$. They also investigated the star formation rate (SFR) and found it to be declining from past to present. However, it should be noted that a degeneracy exists between the SFR and the slope of the IMF (Haywood et al. 1997b), and constraining both of them together is challenging.

The GCS survey (Nordström et al. 2004) combined the Hipparcos and Tycho-2 (Høg et al. 2000) proper motions with radial velocity measurements to create a kinematically unbiased sample of 16682 F and G stars in the Solar neighborhood. The data contains full 6D phase space information along with estimates of ages. The temperature, metallicity and ages were further improved by Holmberg et al. (2007) and distances and kinematics were improved by Holmberg et al. (2009) using revised Hipparcos parallaxes. They investigated the AVR and found $(\beta_R, \beta_\phi, \beta_z) = (0.39, 0.40, 0.53)$ which are at odds with Aumer & Binney (2009). Casagrande et al. (2011) used the infrared flux method to improve the temperature, metallicity and age estimates for the GCS survey. The uncertainty in estimated ages is an ongoing concern for studies that attempt to derive the AVR directly from the GCS data.

With the advent of large spectroscopic surveys like RAVE (Steinmetz et al. 2006) and SDSS/SEGUE (Yanny et al. 2009), we can now get the radial velocity and stellar parameters for a large number of stars and even for stars beyond the Solar neighborhood. Bovy et al. (2012c,b,d) used SDSS/SEGUE to find that mono-abundance populations can be fit by simple double exponentials, and the vertical velocity dispersion follows an exponential decline with radius with no other z dependence. Finally, they also find that the thick disc seems to be a continuation of the thin disc rather than a separate entity.

The RAVE survey has also been used to study the stellar kinematics of the Milky Way disc. Pasetto et al. (2012a,b) study the velocity dispersion and mean motion of the thin and thick disc stars in the (R, z) plane. They use the technique of singular value decomposition to compute the moments of the velocity distribution. Their analysis clearly shows that velocity dispersions fall as a function of distance from the Galactic Center. Williams et al. (2013) explored the kinematics using red clump stars from RAVE and found complex structures in velocity space. A detailed comparison with the prediction from the code *Galaxia* was done, taking the selection function of RAVE into account. The trend of dispersions in the (R, z) plane showed a good match with the model. However, the mean velocities showed significant differences. Boeche et al. (2013) studied the relation between kinematics and the chemical abundances of stars. By computing stellar orbits they deduced the maximum vertical distance $z_{\max}$ and eccentricity e of stars. Next they studied the chemical properties of stars by binning them up in $(z_{\max}, e)$ plane. They found that stars with $z_{\max} < 1$ and $0.4 < e < 0.6$ have two populations with distinct chemical properties, which hints at radial migration.

Stellar kinematics allow us to measure the peculiar motion $(U_\odot, V_\odot, W_\odot)$ of the Sun with respect to the local standard of rest (LSR), and also the LSR value itself (in other words, the circular velocity at the location of Sun, $v_{\rm c}(R_0)$). There have been as many determinations of these as there have been new data, one of the earliest being $(U_\odot, V_\odot, W_\odot) = (9, 12, 7)$ km s^{-1} by Delhaye (1965). The most accurate measurement of these till now are from the Hipparcos proper motions and the Geneva Copenhagen survey. Dehnen & Binney (1998b) and Aumer & Binney (2009) using Hipparcos proper motions got $(U_\odot, V_\odot, W_\odot) = (9.96 \pm 0.33, 5.25 \pm 0.54, 7.07 \pm 0.37)$ km s^{-1}. A revision of $V_\odot$ was suggested by (Binney 2010) and (McMillan & Binney 2010). Later Schönrich et al. (2010) explained why the previous estimates that were using colors as a proxy for age gave wrong results. Using a chemo-dynamical model on GCS data they found $(U_\odot, V_\odot, W_\odot) = (11.1 \pm 0.72, 12.24 \pm 0.47, 7.07 \pm 0.36)$ km s^{-1}. Schönrich (2012) described a model-independent method and suggests that $U_\odot$ could be as high as 14 km s^{-1}. As further evidence of an unsettled situation, Bovy et al. (2012a) find $v_{\rm c} = 218 \pm 6$, $V_\odot = 26 \pm 3$ and $U_\odot = 10.5$ km s^{-1} and also suggest a revision of the LSR reference frame.

In this paper, we refine the kinematic parameters of the Milky Way, using first a simple model based on Gaussian functions, and then a model based on the Shu distribution function (DF). We explore the age-velocity dispersion relation, the radial gradient in dispersions, the Solar motion and

the circular velocity. A full exploration of the parameter space using Markov chain Monte Carlo (MCMC) techniques has not been done before, even for a sample as small as the GCS. If the full 6D phase space information is not available, one has to marginalize over unknown variables, a process that is computationally expensive. For Gaussian models, an analytic form exists for the marginal distribution of line-of-sight (LOS) velocities when integrated over tangential velocities, but not for Shu DF models. Bovy et al. (2012a) recently carried out a detailed investigation using 3500 APOGEE stars, but the AVR was not investigated and only Gaussian models were considered. In this paper, we fit a kinematic model to the 280,000 RAVE stars, where we marginalize over age, metallicity, mass, distance and proper motions. The marginalization is done taking into account the photometric selection function of RAVE. To handle the large data size, we discuss two new MCMC model-fitting techniques. Our aim is to encapsulate the main kinematic properties of the Milky Way disc by means of simple analytical models, which in future would be also useful for making detailed comparison with simulations.

The paper is organized as follows. In §2, we introduce the analytic framework employed for modelling. In §3, we describe the data that we use and its selection functions. In §4, we describe MCMC model fitting technique employed here. In §5, we present our results and discuss their implications in §6. Finally, in §7 we summarize our findings and look forward to the next stages of the project.

2. ANALYTIC FRAMEWORK FOR MODELLING THE GALAXY

We first describe the analytic framework used to model the Galaxy (Sharma et al. 2011). The stellar content of the Galaxy is modeled as a set of distinct components: the thin disc, the thick disc, the stellar halo and the bulge. The distribution function, i.e., the number density of stars as a function of position ($\mathbf{r}$), velocity ($\mathbf{v}$), age (τ), metallicity (Z), and mass (m) of stars for each component is assumed to be specified a priori. This can be expressed in general as

$$f_j = f_j(\mathbf{r}, \mathbf{v}, \tau, Z, m) \tag{1}$$

where $j\ (= 1, 2, 3, 4)$ runs over components. The correct form of Equation (1) that describes all the properties of the Galaxy and is self-consistent is still an open question. However, over the past few decades considerable progress has been made in identifying a working model dependent on a few simple assumptions (Robin & Creze 1986; Bienayme et al. 1987; Haywood et al. 1997a,b; Girardi et al. 2005; Robin et al. 2003). Our analytical framework brings together these models as we describe this below.

For a given Galactic component, let the stars form at a rate $\Psi(\tau)$ and the mass distribution of stars $\xi(m|\tau)$ (IMF) be a parameterized function of age τ only. Let the present day spatial distribution of stars $p(\mathbf{r}|\tau)$ be conditional on age only. Finally, assuming that the velocity distribution to be $p(\mathbf{v}|\mathbf{r}, \tau)$ and the metallicity distribution to be $p(Z|\tau)$

$$f(\mathbf{r}, \mathbf{v}, \tau, m, Z) = \frac{\Psi(\tau)}{\langle m \rangle} \xi(m|\tau) p(\mathbf{r}|\tau) p(\mathbf{v}|\mathbf{r}, \tau) p(Z|\tau). \tag{2}$$

The functions conditional on age can take different forms for different Galactic components. The IMF here is normalized such that $\int_{m_{\rm min}}^{m_{\rm max}} \xi(m|\tau) dm = 1$ and $\langle m \rangle = \int_{m_{\rm min}}^{m_{\rm max}} m \xi(m|\tau) dm$ is the mean stellar mass.

The metallicity distribution is modeled as a log-normal distribution,

$$p(Z, |\tau) = \frac{1}{\sigma_{\log Z}(\tau)\sqrt{2\pi}} \exp\left[-\frac{(\log Z - \log \bar{Z}(\tau))}{(2\sigma^2_{\log Z}(\tau))} \right], \tag{3}$$

the mean and dispersion of which are given by age-dependent functions $\bar{Z}(\tau)$ and $\sigma_{\log Z}(\tau)$. The $\bar{Z}(\tau)$ is widely referred to as the age-metallicity relation (AMR). Functional forms for each of the expressions in Equation (2) are given in Sharma et al. (2011) (see also Robin et al. 2003).

We now discuss the functional forms of the velocity distribution $p(\mathbf{v}|\mathbf{r}, \tau)$ which we aim to derive in this paper.

2.1. *Gaussian velocity ellipsoid model*

In this model, the velocity distribution is assumed to be a triaxial Gaussian,

$$p(\mathbf{v}|\mathbf{r}, \tau) = \frac{1}{\sigma_R \sigma_\phi \sigma_z (2\pi)^{3/2}} \exp\left[-\frac{v_R^2}{2\sigma_R^2} \right] \exp\left[-\frac{v_z^2}{2\sigma_z^2} \right] \times \exp\left[-\frac{(v_\phi - \overline{v_\phi})^2}{2\sigma_\phi^2} \right], \tag{4}$$

where R, ϕ, z are the cylindrical coordinates. $v_{\rm c}(R)$ the circular velocity as a function of cylindrical radius R. The $\overline{v_\phi}$ is the asymmetric drift and is given by

$$\overline{v_\phi}^2(\tau, R) = v_{\rm c}^2(R) + \sigma_R^2 \times \left(\frac{d \ln \rho}{d \ln R} + \frac{d \ln \sigma_R^2}{d \ln R} + 1 - \frac{\sigma_\phi^2}{\sigma_R^2} + 1 - \frac{\sigma_z^2}{\sigma_R^2} \right) \tag{5}$$

This follows from Equation 4.227 in Binney & Tremaine (2008) assuming $\overline{v_R\, v_z} = (v_R^2 - v_z^2)(z/R)$. This is valid for the case where the principal axes of velocity ellipsoid are aligned with the (r, θ, ϕ) spherical coordinate system. If the velocity ellipsoid is aligned with the cylindrical (R, z, ϕ) coordinate system then $\overline{v_R\, v_z} = 0$. Recent results using the RAVE data suggest that the velocity ellipsoid is aligned with the spherical coordinates (Siebert et al. 2008). One can parameterize our ignorance by writing the asymmetric drift as follows:

$$\overline{v_\phi}^2(\tau, R) = v_{\rm c}^2(R) + \sigma_R^2 \left(\frac{d \ln \rho}{d \ln R} + \frac{d \ln \sigma_R^2}{d \ln R} + 1 - k_{\rm ad}^2 \right) \tag{6}$$

This is the form that is also used by Bovy et al. (2012a).

The dispersions of the R, z and ϕ components of velocity increase as a function age due to secular heating in the disc, and there is a radial dependence such that the dispersion increases towards the Galactic Center. We model these effects after Aumer & Binney (2009) and Binney (2010) using the functional form

$$\sigma^{\rm thin}_{R,\phi,z}(R, \tau) = \sigma^{\rm thin}_{R,\phi,z,\odot} \exp\left[-\frac{q_{\rm thin}(R - R_0)}{R_d} \right] \times \left(\frac{\tau + \tau_{\rm min}}{\tau_{\rm max} + \tau_{\rm min}} \right)^{\beta_{R,\phi,z}} \tag{7}$$

$$\sigma^{\rm thick}_{R,\phi,z}(R) = \sigma^{\rm thick}_{R,\phi,z,\odot} \exp\left[-\frac{q_{\rm thick}(R - R_0)}{R_d} \right] \tag{8}$$

where Σ is the surface density of the disc. The choice of the radial dependence is motivated by the desire to produce discs

in which the scale height is independent of radius. For example, under the epicyclic approximation, if σ_z/σ_R is assumed to be constant, then the scale height is independent of radius for $q = 0.5$ (van der Kruit & Searle 1982; van der Kruit 1988; van der Kruit & Freeman 2011). In reality there is also a z dependence of velocity dispersions which we have chosen to ignore in our present analysis. This means that for a given mono age population the asymmetric drift is independent of z. However, the velocity dispersion and asymmetric drift of the combined population of stars are a function of z. This is because the scale height of stars for a given isothermal population is a monotonically increasing function of its vertical velocity dispersion.

2.2. *Shu distribution function model*

The Gaussian velocity ellipsoid model has its limitations. In reality, the distribution of angular momentum L is not Gaussian but is skewed. A better approximation to the velocity spread is the Shu (1969) distribution function. Another advantage of the Shu DF is that given a potential, density distribution and velocity dispersion profile it gives a self consistent distribution of velocities. In contrast, in the Gaussian model the potential density and velocities are independent of each other.

Strictly speaking, the validity of the Shu model is restricted to motion in a plane. However, assuming vertical motions are independent of planar motion, one can write the full distribution function as follows:

$$f(E_R, L, v_z) = \frac{F(L)}{\sigma^2(L)} \exp\left[-\frac{E_R}{\sigma^2(L)}\right] \frac{\exp\left[-v_z^2/(2\sigma_z^2)\right]}{\sigma_z\sqrt{2\pi}} \quad (9)$$

where

$$\begin{aligned} E_R &= \frac{1}{2}v_R^2 + \Phi_{\rm eff}(R, L) - \Phi_{\rm eff}(R_L, L) \\ &= \frac{1}{2}v_R^2 + \Delta\Phi_{\rm eff}(R, L) \end{aligned} \quad (10)$$

with

$$\Phi_{\rm eff}(R, L) = \frac{L^2}{2R^2} + \Phi(R) = \frac{L^2}{2R^2} + v_c^2 \ln R \quad (11)$$

being the effective potential. Let $R_g(L) = L/v_{\rm c}$ be the radius of a circular orbit with specific angular momentum L. In Schönrich & Binney (2012) (see also Sharma & Bland-Hawthorn 2013) it was shown that joint distribution of R and R_g can be written as

$$P(R, R_g) = \frac{(2\pi)^2\Sigma(R_g)}{g(\frac{1}{2a^2})} \exp\left[\frac{2\ln(R_g/R) + 1 - R_g^2/R^2}{2a^2}\right] \quad (12)$$

where

$$a = \sigma(R_g)/v_{\rm c} \quad (13)$$

$$g(c) = \frac{e^c\Gamma(c - 1/2)}{2c^{c-1/2}} \quad (14)$$

We assume a to be specified as

$$\begin{aligned} a &= a_0(\tau)\exp\left[-\frac{qR_g}{R_d}\right] \\ &= \frac{\sigma_{R,\odot}}{v_{\rm c}} \left(\frac{\tau + \tau_{\rm min}}{\tau_{\rm max} + \tau_{\rm min}}\right)^{\beta_R} \exp\left[-\frac{q(R_g - R_0)}{R_d}\right] \end{aligned} \quad (15)$$

TABLE 1
DESCRIPTION OF MODEL PARAMETERS

Model Parameter	Description
$U_\odot$	Solar motion with respect to LSR
$V_\odot$	Solar motion with respect to LSR
$W_\odot$	Solar motion with respect to LSR
$\sigma_\phi^{\rm thin}$	The velocity dispersion at 10 Gyr Normalization of thin disc AVR (Eq 8)
$\sigma_z^{\rm thin}$	The velocity dispersion at 10 Gyr Normalization of thin disc AVR (Eq 8)
$\sigma_R^{\rm thin}$	The velocity dispersion at 10 Gyr Normalization of thin disc AVR (Eq 8)
$\sigma_\phi^{\rm thick}$	The velocity dispersion of thick disc (Eq 8)
$\sigma_z^{\rm thick}$	The velocity dispersion of thick disc (Eq 8)
$\sigma_R^{\rm thick}$	The velocity dispersion of thick disc (Eq 8)
β_R	The exponent of thin disc AVR (Eq 8)
β_z	The exponent of thin disc AVR (Eq 8)
β_ϕ	The exponent of thin disc AVR (Eq 8)
$q_{\rm thin}$	The exponent of the velocity dispersion profile for thin disc (Eq 8)
$q_{\rm thick}$	The exponent of the velocity dispersion profile for thick disc (Eq 8)
R_0	Distance of Sun from the Galactic Center
v_0	The circular velocity at Sun
α_z	A parameter for vertical fall of circular velocity
α_R	The radial gradient of circular velocity at Sun

and σ_z to be specified as

$$\sigma_{z0}(R_g, \tau) = \sigma_{z,\odot} \left(\frac{\tau + \tau_{\rm min}}{\tau_{\rm max} + \tau_{\rm min}}\right)^{\beta_z} \exp\left[-\frac{q(R_g - R_0)}{R_d}\right] \quad (16)$$

Now this leaves us to choose $\Sigma(R_g)$. This should be done so as to produce discs that satisfy the observational constraint given by $\Sigma(R)$, i.e., an exponential disc (or discs) with scale length R_d. A simple way to do this is to let

$$\Sigma(R_g) = \frac{1}{2\pi R_d^2} \exp\left[-\frac{R_g}{R_d}\right]. \quad (17)$$

However, this matches the target surface density only approximately. A better way to do this is to use the empirical formula proposed in Sharma & Bland-Hawthorn (2013) such that

$$\Sigma(R_g) = \frac{e^{-R_g/R_d}}{2\pi R_d^2} - \frac{0.00976a_0^{2.29}}{R_d^2} s\left[\frac{R_g}{(3.74R_d(1 + q/0.523)}\right] \quad (18)$$

where s is a function of the following form

$$s(x) = ke^{-x/b}((x/a)^2 - 1), \quad (19)$$

with $(k, a, b) = (31.53, 0.6719, 0.2743)$. This is the scheme that we employ in this paper.

2.3. *Model for circular velocity*

So far we have described kinematic models in which the circular velocity is constant. However, the circular velocity can have both a radial and a vertical dependence. We model it as follows.

$$v_{\rm c}(R, z) = \sqrt{R\frac{\partial\Phi}{\partial R}} = (v_0 + \alpha_R R)\frac{1}{1 + \alpha_z|z|^{1.34}} \quad (20)$$

The parameters α_R and α_z control the radial and vertical dependence respectively. The motivation for the vertical term comes from the fact that the above formula with $\alpha_z = 0.0374$

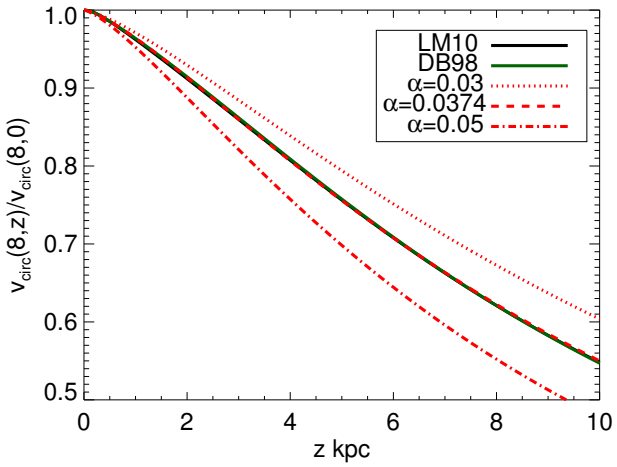

FIG. 1.— Circular velocity as a function of height z above the mid plane for models of the Milky Way consisting of bulge, halo and disc. The non solid red lines are for the fitting formula with different values of α_z. The larger the α_z the steeper is the fall of circular velocity.

provides a good fit to the $v_c(R_\odot, z)$ profile of Milky Way potential by Dehnen & Binney (1998a) as well as Law & Majewski (2010) (see Figure 1). Both of them have bulge, halo and discs. The former has two double exponential discs while the later has a Miyamoto-Nagai disc. Setting $\alpha_z > 0.03744$ one can also approximately take into account the increase of asymmetric drift of a mono age population as a function of height z. This can be easily incorporated by changing v_c to $v_c(R, z)$ in Equation (5) for a Gaussian model and in Equation (13) for Shu model. Strictly speaking, for the Shu model the above prescription is correct only for $\alpha_R = 0$. When a radial dependence is present, $P(R, R_g)$ is no longer analytical (Sharma & Bland-Hawthorn 2013). However, in this paper we are mainly interested in cases where α_R is either zero or small and the above prescription should be a good first order estimate.

2.4. *Models and parameters explored*

We now give a description of the parameters and models that we explore. We investigate up to 18 parameters (see Table 1 for a summary). These are the Solar motion $(U_\odot, V_\odot, W_\odot)$, the logarithmic slope of age-dispersion relations $(\beta_R, \beta_z, \beta_\phi)$, the logarithmic slope of radial dependence of velocity dispersions $(q_{\rm thin}, q_{\rm thick})$, the velocity dispersions at $R = R_0$ of the thin disc $(\sigma_\phi^{\rm thin}, \sigma_z^{\rm thin}, \sigma_R^{\rm thin})$ and the thick disc $(\sigma_\phi^{\rm thick}, \sigma_z^{\rm thick}, \sigma_R^{\rm thick})$; for simplicity the subscript $\odot$ is dropped here. The Gaussian models are denoted by GU whereas models based on the Shu DF are denoted by SHU. For models based on the Shu DF, the azimuthal motion is coupled to the radial motion, hence β_ϕ, $\sigma_\phi^{\rm thin}$ and $\sigma_\phi^{\rm thick}$ are not required. When the circular velocity is kept fixed we assume its value to be 226.84 km s^{-1}. In some cases, we also keep the parameters β_z and $q_{\rm thin}$ fixed. While reporting the results we highlight the fixed parameters using the magenta color.

In our analysis the distance of the Sun from the galactic center, R_0, is assumed to be 8.0 kpc. To gauge the sensitivity of our results to R_0, we also provide results for cases with $R_0 = 7.5$ and 8.5 kpc. The true value of R_0 is still debatable ranging from 6.5 to 9 kpc. Recent results from studies of orbit of stars near the Galactic Center give $R_0 = 8.33 \pm 0.35$ (Gillessen et al. 2009). The classically accepted value of 8 ± 0.5 kpc is a weighted average given in a review by Reid (1993). The main reason we keep R_0 fixed is as follows. Given that we do not make use of explicit distances, proper motion or external constraints like proper motion of Sgr A* star, it is clear we will not be able to constrain R_0 well, specially if v_c is kept free. For example McMillan & Binney (2010) using parallax, proper motion and line of sight velocity of masers in high star forming regions, show that constraining both v_0 and R_0 independently is difficult.

3. OBSERVATIONAL DATA AND SELECTION FUNCTIONS

In this paper we analyze data from two surveys, the Radial Velocity Experiment, RAVE (Steinmetz et al. 2006; Zwitter et al. 2008; Siebert et al. 2011; Kordopatis et al. 2013) and the Geneva Copenhagen Survey, GCS (Nordström et al. 2004; Holmberg et al. 2009). For fitting theoretical models to data from stellar surveys, it is important to take into account the selection biases that were introduced when observing the stars. This is especially important for spectroscopic surveys which are not unbiased and observe only a subset of the all possible stars defined within a color-magnitude range. So we also analyze the selection function for the RAVE and GCS survey.

3.1. *RAVE survey*

The RAVE survey collected spectra of 482430 stars between April 2004 and December 2012 and stellar parameters, radial velocity, abundance and distances have been determined for 425561 stars. In this paper we used the internal release of RAVE from May 2012, which consisted of 458412 observations. The final explored sample after applying various selection criteria consists of 280128 unique stars. This data is available in the DR4 public release (Kordopatis et al. 2013) where an extended discussion of the sample is also presented.

For RAVE we only make use of the ℓ, b and $v_{\rm los}$ of stars. The $I_{\rm DENIS}$ and 2MASS $J-K_s$ colors are used for marginalization over age, metallicity and mass of stars taking into account the photometric selection function of RAVE. We do not use proper motions, or stellar parameters which could in principle provide tighter constraints, but then one has to worry about the systematics introduced by their use. For example, in a recent kinematic analysis of RAVE stars by (Williams et al. 2013), they found systematic differences between different proper motion catalogs like PPMXL (Röser et al. 2008), SPM4 (Girard et al. 2011) and UCAC3 (Zacharias et al. 2010). As for stellar parameters, although they are reliable, but no pipeline can claim to be free of unknown systematics specially when working with low signal to noise data. Hence, as a first step it is always instructive to work with data that is least ambiguous and then in the next step check the results by adding more information. As we will show later, for the type of models that we consider even using only ℓ, b and $v_{\rm los}$ can provide sufficiently good constraints on the model parameters.

We now discuss the selection function of RAVE. The RAVE survey was designed to be a magnitude-limited survey in the I band. This means that theoretically it has one of the simplest selection functions although, in practice, for a multitude of reasons, some biases were introduced. First, the DENIS and 2MASS surveys were not fully available when the survey started. Hence, the first input catalog (IC1) had stars from Tycho and SuperCOSMOS. For Tycho stars, I magnitudes were estimated from V_T and B_T magnitudes. On the other hand,

the SuperCOSMOS stars had I magnitudes but an offset was later detected with respect to $I_{\rm DENIS}$. Later on, as DENIS and 2MASS became available, the second input catalog IC2 was created. With the availability of DENIS, it became possible to have a direct I mag measurement which was free from offsets like those observed in SuperCOSMOS. But DENIS itself had its own share of problems – saturation at the bright end, duplicate entries, missing stripes in the sky, inter alia. To solve the problem of duplicate entries, the DENIS catalog was cross-matched with 2MASS to within a tolerance of $1''$. This helped clean up the color-color diagram of $(I_{\rm DENIS} - K_{\rm 2MASS})$ vs $(J_{\rm 2MASS} - K_{\rm 2MASS})$ in particular (Seabroke 2008).

Given this history, the question arises how can we compute the selection function. Since accurate I mag photometry is not available for IC1, the first cut we make is to select stars from IC2 only. Then we removed the duplicates– among multiple observations one of them was selected randomly. To weed out stars with large errors in radial velocity, we made some additional cuts:

$$\text{Signal to Noise STN} > 20$$
$$\text{Tonry} - \text{Davis Correlation Coefficient} > 5.$$

For brighter magnitudes, $I_{\rm DENIS} < 10$, $I_{\rm DENIS}$ suffers from saturation effects. One could either get rid of these stars to be more accurate or ignore the saturation. In the present analysis we ignore the saturation effects. Note, the observed stars in the input catalog are not necessarily randomly sampled from the IC2. Stars were divided into four bins in $I_{\rm mag}$ and stars in each bin were randomly selected to observe at a given time. However it seems later on this division was not strictly maintained (probably due to the observation of calibration stars and some extra stars going to brighter magnitudes). This means the selection function has to be computed as a function of $I_{\rm DENIS}$ in much finer bins. Assuming the DENIS I magnitudes are correct and the cross-matching is correct, the only thing that needs to be taken into account is the angular completeness of the DENIS survey (missing stripes). To this end, we grid the observed and IC2 stars in $(\ell, b, I_{\rm DENIS})$ space and compute a probability map. To grid the angular co-ordinates we use the HEALPIX pixelization scheme (Górski et al. 2005). The resolution of HEALPIX is specified by the number $n_{\rm side}$ and the total number of pixels is given by $12n_{\rm side}^2$. For our purpose, we use $n_{\rm side} = 16$ which gives a pixel size of 13.42 $\deg^2$, which is smaller than the RAVE field of view of 28.3 $\deg^2$. For magnitudes, we use a bin size of 0.1 mag, which again is much smaller than the magnitude ranges employed for each observation. Given the fine resolution of the probability map, the angular and magnitude dependent selection biases are adequately handled. Note, in the range $(225^\circ < \ell < 315^\circ) \& (5^\circ < |b| < 25^\circ)$, a color selection of $(J - K_s) > 0.5$ was used to selectively target giants, and we take this into account in our analysis.

Arce & Goodman (1999) suggest that the Schlegel et al. (1998) maps overestimate reddening by a factor of 1.3-1.5 in regions with smooth extinction $A_V > 0.5$, i.e., $E_{B-V} > 0.15$ (see also Cambrésy et al. 2005). In Figure 2 the color and temperature distribution of RAVE stars that we analyze are compared with predictions from *Galaxia* given the selection above. First we compare the red and black lines. It can be seen that the temperature and color distributions match up well at high latitudes. However, at low latitudes the model color distribution is shifted to the right. The fact that the temperature distribution at low latitude do not show such a shift suggests

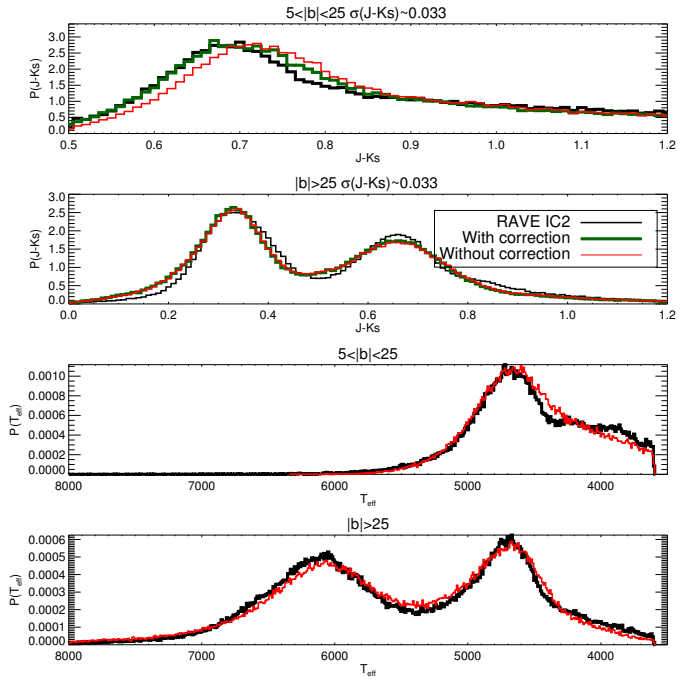

FIG. 2.— The color and temperature distribution (from DR3 pipeline) of RAVE stars compared with *Galaxia* simulations with properly matched selection and statistical sampling. The effect of our new correction formula for the Schlegel extinction map is also shown. The results for $|b| < 25^\circ$ and $|b| > 25^\circ$ are shown separately.

that the problem is related to the modelling of extinction. To account for this we modify the Schlegel E_{B-V} as follows

$$f_{\rm corr} = 0.6 + 0.2\left(1 - \tanh\left[\frac{E_{B-V} - 0.15}{0.1}\right]\right) \quad (21)$$

The formula above reduces extinction by 40% for high extinction regions; the transition occurs around $E_{B-V} \sim 0.15$ and is smoothly controlled by the tanh function. It can be seen that the proposed correction to Schlegel maps, although not perfect, reduces the discrepancy between the model and data for low latitude stars (top panel). The effect of the correction is negligible for high latitude stars as expected (second panel).

In Figure 2, we applied only an $I_{\rm DENIS}$ magnitude selection, the fact that the temperature and color distributions match up so well is encouraging. This means that the spatial distribution of stars as specified in the *Galaxia* model do satisfy one of the necessary observational constraint.

3.2. *GCS survey*

The GCS stars were sampled as in Sharma et al. (2011). We use the data from the Geneva-Copenhagen Survey, GCS, (Nordström et al. 2004; Holmberg et al. 2009), which is a selection of 16682 F and G type main-sequence stars, out of which velocities and temperatures are available for 13382 stars. We found that while *Galaxia* predicts less than one halo star in the GCS sample for a distance less than 120 pc, when plotted in $([{\rm Fe/H}], v_y)$ plane, the GCS has 29 stars with $[{\rm Fe/H}] < -1.2$ that have highly negative values of v_y (as expected for halo stars). Just like Schönrich et al. (2010), we identify these as halo stars and exclude them from our analysis.

The GCS catalog is complete for F and G type stars for a volume given by $r < 40$ pc and $V \sim 8$ in magnitude; within these limits there are only 1342 stars. But GCS being a color

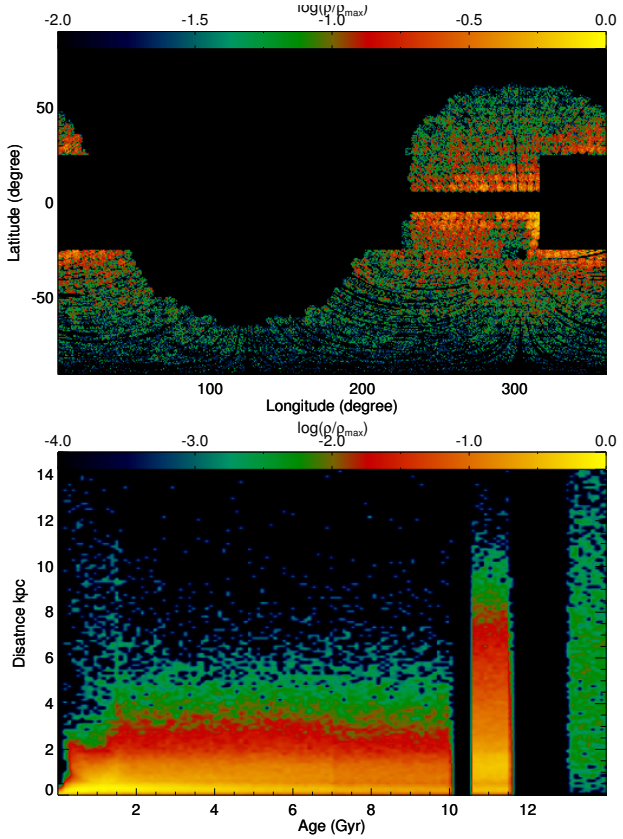

FIG. 3.— Probability distribution of RAVE stars analyzed in this paper in (ℓ, b) space (*Top*) and (Age, Distance) space (*Bottom*). The age-distance distributions are predictions from the *Galaxia* model for stars satisfying the RAVE selection criteria.

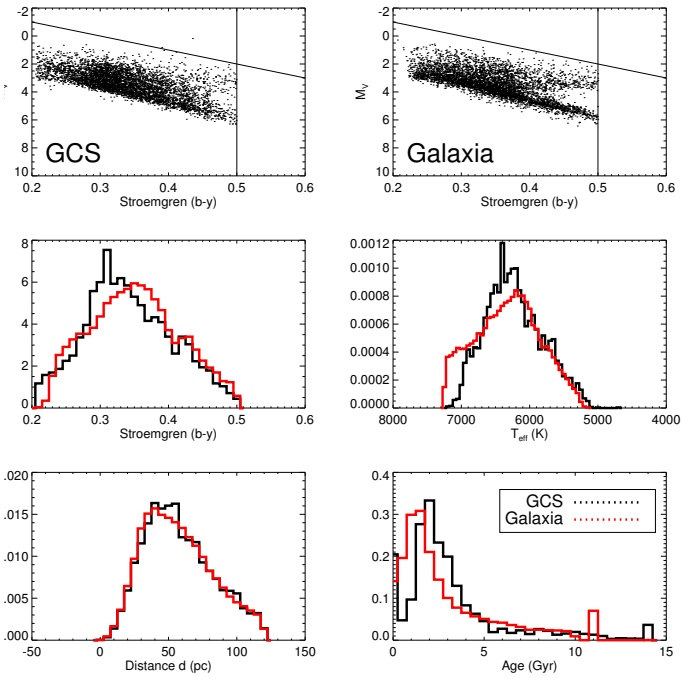

FIG. 4.— Distribution of GCS stars as a function of color, temperature, distance and age. Shown alongside are results of a mock sample created using *Galaxia* but without observational uncertainties. The top panel shows the distribution in the $(b-y, M_V)$ plane; the colors span the range $0.205 < (b-y) < 0.5$. The magnitude limits are a function of color and are taken from Nordström et al. (2004). The line represents the equation $M_V = 10(b-y)-3$ and is used to mimic the selective avoidance of giants in GCS. A selection of $d < 0.12$ kpc and $T_{\rm eff} > 7244$ K is also applied. The temperature and ages (maximum likelihood Padova) are from Casagrande et al. (2011).

magnitude limited survey, there is no need to restrict the analysis to a volume complete sample. In Nordström et al. (2004) magnitude completeness as a function of color is provided and we use this (their §2.2). There is some ambiguity about the coolest dwarfs which were added for declination $\delta < -26°$; from information gleaned from Nordström et al. (2004), we could not find a suitable way to take this into account.

We also applied some additional restrictions on the sample. For example, we restrict our analysis to stars with distance less than 120 pc, so as to avoid stars with large distance errors. The GCS survey selectively avoids giants. To mimic this we use the following selection function $M_V < 10(b-y)-3$. The predicted temperature distributions show a mismatch with models, in particular, there are too many hot stars. Using Casagrande et al. (2011) temperatures, which are more accurate, we found an upper limit on $T_{\rm eff}$ of 7244 K, which was applied to the models.

After the above mentioned cuts, the final sample consisted of 5201 stars. In Figure 4, we show the distribution of GCS and model stars after applying the above mentioned selection functions. The temperature distribution at the hot end still shows some difference but the distance distribution is correctly reproduced. The age distribution is qualitatively correct given that we do not convolve with uncertainties or take systematics into account. The peak in the model at 11 Gyr is due to the thick disc having a constant age. The peaks in the data at 0 and 14 Gyr are most likely due to caps employed while estimating ages. The color distribution in GCS shows a peak at around $b - y = 0.3$, which could be due to an unknown selection effect. The bump at $b - y \sim 0.43$ which is also seen in models is due to turnoff stars. Overall, we think our modelling reproduces to a good degree the selection function of the GCS stars.

4. MODEL FITTING TECHNIQUES

If y_i are the observed properties of a star, we can describe the observed data by $y = \{y_i \in \mathbb{R}^d, 0 < i < N\}$. Also, let θ be the set of parameters that define the model. Our job is to compute

$$p(\theta|y) \propto p(y|\theta)p(\theta) \tag{22}$$

where $p(y|\theta) = \prod_i p(y_i|\theta)$. We employ an MCMC scheme to estimate $p(\theta|y)$ and assume a uniform prior on θ. We now discuss how to compute $p(y_i|\theta)$.

Generally, a model of a galaxy gives the probability density $p(\mathbf{r}, \mathbf{v}, \tau, Z, m|\theta)$. For RAVE, the observed quantities are $v_{\rm los}, l$ and b, while for GCS it is l, b, r, v_l, v_b and $v_{\rm los}$. Since quantities like τ, Z and m are not known, one has to compute the marginal probability density by integration. For RAVE, the required marginal density is

$$p(\ell, b, v_{\rm los}|\theta) = \int p(\ell, b, r, \tau, Z, m, v_l, v_b, v_{\rm los}|\theta) \times S(\ell, b, \tau, Z, m)\, dr\, d\tau\, dZ\, dm\, dv_l\, dv_b, \tag{23}$$

and for GCS it is

$$p(\ell, b, r, v_l, v_b, v_{\rm los}|\theta) = \int p(\ell, b, r, \tau, Z, m, v_l, v_b, v_{\rm los}|\theta) \times S(\ell, b, \tau, Z, m)\, d\tau\, dZ\, dm. \tag{24}$$

Here $S(\ell, b, \tau, Z, m)$ is the selection function specifying how the stars were preselected in the data. The actual selection is on photometric magnitude which in turn is a function of τ, Z and m.

For the kinds of models explored here, the computations are considerably simplified due to the fact that

$$p(\ell, b, r, \tau, Z, m, v_l, v_b, v_{\rm los}|\theta) = p(v_l, v_b, v_{\rm los}|\ell, b, r, \tau, \theta) \times p(\ell, b, r, \tau, Z, m|\theta_S), \quad (25)$$

for which θ_S is the set of model parameters that govern the spatial distribution of stars and θ is the set of model parameters that govern the kinematic distribution of stars. The term $p(\ell, b, r, \tau, Z, m|\theta_S)$ is invariant in our analysis, and this is the main assumption that we make. In other words we assume a star formation history (SFR), initial mass function (IMF), scale length of disc, an age-scale height relation and an age-metallicity relation for the disc. This can be constrained by the photometry of the stars. The $p(v_l, v_b, v_{\rm los}|\ell, b, r, \tau, \theta)$ represents the kinematics which is what we explore. It should be noted that the model $p(\ell, b, r, \tau, Z, m|\theta_S)$ that we use has been shown to satisfy the number count of stars (Robin et al. 2003; Sharma et al. 2011). Ideally, in a fully self consistent model, the scale height is related to the stellar velocity dispersion and this is something we would like to address in future.

We can now integrate the first term in Equation (25) over m and Z such that

$$p(\ell, b, r, v_l, v_b, v_{\rm los}, \tau|\theta) = p(v_l, v_b, v_{\rm los}|\ell, b, r, \tau, \theta) \times p(\ell, b, r, \tau|\theta_S, S) \quad (26)$$

where

$$p(\ell, b, r, \tau|\theta_S, S) = \int\int p(\ell, b, r, \tau, Z, m|\theta_S) \times S(\ell, b, \tau, Z, m)\, dZ\, dm. \quad (27)$$

The term $p(\ell, b, r, \tau|\theta_S, S)$ is computed numerically using the code *Galaxia* (Sharma et al. 2011). Galaxia, uses isochrones from the Padova database to compute photometric magnitude of the model stars (Marigo et al. 2008; Bertelli et al. 1994). We first generate a fiducial set of stars satisfying the color magnitude range of the survey. Then we apply the selection function and reject stars that do not satisfy the constraints of the survey. The accepted stars are then binned in (ℓ, b, r, τ) space. Since, the GCS is local to Sun, we use the following approximation $p(\ell, b, r, \tau|\theta_S, S) \propto p(\tau|\theta_S, S)$. The probability distribution in (ℓ, b, r, τ) space for RAVE is shown in Figure 3.

For RAVE, we have to integrate over four variables (r, τ, v_l, v_b), but for GCS we integrate over only τ. The 4D marginalization for RAVE poses a serious computational challenge for data as large as the RAVE survey. For Gaussian distribution functions, the integral over v_l and v_b can be carried out to give an analytic expression for $p(v_{\rm los}|\ell, b, r, \tau, Z, \theta)$, but in general it cannot be done. Hence, we try two new methods. The first method is fast but has inflated uncertainties. The second method is slower to converge but gives correct estimates of uncertainties. Given these strengths and limitations, we use a combined strategy that makes best use of both the methods.

We use the first 'sampling and projection' method to get an initial estimate of θ and also its covariance matrix. These are then used in the second 'data augmentation' method. The initial estimate reduces the 'burn in' time, while the covariance matrix eliminates the need to tune the widths of the proposal distributions. In general we use an adaptive MCMC scheme, which avoids manual tuning of the widths of the proposal distributions (Andrieu & Thoms 2008). At regular intervals, we compute the covariance matrix and scale it so as to achieve the desired acceptance ratio for the given number of parameters Gelman et al. (1996). We now discuss the two methods in more detail.

4.1. *MCMC using sampling and projection*

Instead of doing the computationally intensive marginalization, we generate a sample of stars by Monte Carlo sampling satisfying the given distribution function and the selection function. Taking a histogram of these stars in $(\ell, b, v_{\rm los})$ space then gives $p(\ell, b, v_{\rm los}|\theta)$. We then run a Markov Chain Monte Carlo simulation to estimate the likelihood distribution of the model parameters. Note that, given the stochastic nature of our model distribution function, the standard Metropolis-Hastings algorithm had to be altered to avoid the simulation from getting stuck at a stochastic maximum of the likelihood.

4.2. *MCMC using data augmentation*

Instead of marginalizing one can treat the nuisance parameters as unknown parameters and estimate them alongside other parameters. This constitutes what is known as, a sampling based approach for computing the marginal densities, the basic form of this scheme was introduced by Tanner & Wong (1987) and later on extended in (Gelfand & Smith 1990). Let $x = \left\{x_i \in \mathbb{R}^d, 0 < i < N\right\}$ be an extra set of variables that are needed by the model to compute the probability density. Then we can write

$$p(\theta, x|y) = p(x, y|\theta)p(\theta). \quad (28)$$

where $p(x, y|\theta) = \prod_i p(x_i, y_y|\theta)$, and $p(x_i, y_i|\theta)$ is a function which is known and relatively easy to compute. For example, for the RAVE data $y_i = \{l_i, b_i, v_{i,\rm los}\}$ and $x = \{r_i, \tau_i, v_{i,l}, v_{i,b}\}$. Due to the unusually large number of parameters, it is difficult to get satisfactory acceptance rates with the standard Metropolis-Hastings scheme without making the widths of the proposal distributions extremely small. Thus the chains would take an unusually long time to mix. To solve this, one uses the Metropolis scheme with Gibbs sampling (MWG) (Tierney 1994). The MWG scheme is also useful for solving hierarchical Bayesian models, and its application for 3D extinction mapping is discussed in Sale (2012). In our case, the Gibbs step consists of first sampling x from the conditional density $p(x|y, \theta)$ and then θ from the conditional density $p(\theta|y, x)$. The sampling in each Gibbs step is done using the Metropolis-Hastings algorithm.

4.3. *Tests using synthetic data*

We now run tests where the data is sampled from the distribution function and then fitted using the MCMC machinery. These tests serve two main purposes. First, they determine if our MCMC scheme works correctly or not. Secondly, they tell us which parameters can be recovered and with what accuracy. We study two classes of models based on (1) the Gaussian DF and (2) the Shu DF. Additionally, we study two types of mock data, one corresponding to the RAVE survey and the other to the GCS survey. For GCS we also study models where $v_{\rm c}$ is fixed. Altogether this leads to 6 different types of tests.

The results of these tests are summarized in Tables 2 and 3. The difference of a parameter p from input values divided by uncertainty σ_p measures the confidence of recovering the parameter. To aid the comparison, we color the values if they differ significantly from the input values: $|\delta p|/\sigma_p < 2$ (black), $2 < |\delta p|/\sigma_p < 3$ (blue). It can be seen that all parameters are recovered within the 3σ range as given by the error bars. Ideally to check the systematics, the fitting should be repeated multiple times and the mean values should be compared with input values. However, the MCMC simulations being computationally very expensive we report results with only one independent data sample for each of the test cases.

It can be seen that GCS type data cannot properly constrain v_c. This is because the GCS sample is very local to the Sun. Keeping v_c free also has the undesirable effect of increasing the uncertainty of $q_{\rm thin}$ and $q_{\rm thick}$. For Gaussian models, it is easy to see from Equation (5) that the effect of changing v_c can be compensated by a change in $q_{\rm thin}$ and $q_{\rm thick}$. Given these limitations, when analyzing GCS we keep v_c fixed to 226.87 km s^{-1}, a value that was used by Sharma et al. (2011) in the *Galaxia* code.

The Solar motion is constrained well by both surveys, but better by RAVE. The RAVE is also clearly better in constraining thick disc parameters than GCS, mainly because the GCS has very few thick disc stars (*Galaxia* estimates it to be 6% of the overall GCS sample). *Across all parameters, for Shu models β_z is the only parameter which is constrained better by GCS than by RAVE*. This is because RAVE only has radial velocities. This means that only those stars that lie towards the pole can carry meaningful information about the vertical motion, and such stars constitute a much smaller subset out of the whole RAVE sample. This suggests that one can use the β_z value from GCS when fitting the RAVE data, as we show below.

5. CONSTRAINTS ON KINEMATIC PARAMETERS

First, we discuss the fiducial parametric model for the Galaxy developed a decade ago by Robin et al. (2003). The so-called Besançcon model is based on Gaussian velocity ellipsoid functions. In the *Galaxia* code, the tabulated functions of Robin et al. (2003) were replaced by analytic expressions, the parameters of which are given in Table 4. One main difference between the *Galaxia* and Besançon models is the value of R_0 and the Solar motion with respect to the LSR. Also, *Galaxia* uses slightly different values of q. The q values corresponding to $R_0 = 8.0$ kpc for the Besançon model are shown in brackets. In the Besançon model, the velocity dispersions are assumed to saturate abruptly at around $\tau_{\rm sat} 6.5$ Gyr. Moreover, the velocity dispersion of the thick disc does not have any radial dependence, hence the value of $q_{\rm thick}$ only contributes to the calculation of the asymmetric drift. Neither of these Ansätze are assumed in our analysis.

Finally, in the Besançon model, the metallicity [Fe/H] of the thick disc is assumed to be -0.78 with a spread of 0.3 dex. The spread is not taken into account when assigning magnitude and color from isochrones. This was done so as to prevent the thick disc from having a horizontal branch. We do *not* make this ad hoc assumption. Since our data do not have a strong color-sensitive selection, this has a negligible impact on our kinematic study.

We now discuss the results obtained from fitting models to the RAVE and the GCS data. The best-fit parameters and their uncertainties obtained using MCMC simulation for different models and data are shown in Table 5 and Table 6. We begin by discussing results from the Gaussian distribution function before proceeding to the Shu distribution function.

TABLE 2
TESTS ON MOCK DATA: CONSTRAINTS ON MODEL PARAMETERS WITH GAUSSIAN DISTRIBUTION FUNCTION. THE MODEL RUNS ARE NAMED AS FOLLOWS; SURVEY NAME AS RAVE OR GCS, TYPE OF MODEL AS GU FOR GAUSSIAN AND SHU FOR SHU. VELOCITIES ARE IN km s^{-1} AND DISTANCES IN kpc

Model	GCS GU	GCS GU	RAVE GU	Input
$U_\odot$	$11.12^{+0.43}_{-0.41}$	$11.17^{+0.39}_{-0.39}$	$11.22^{+0.15}_{-0.16}$	11.1
$V_\odot$	$5.8^{+1.8}_{-1.9}$	$8.6^{+1.3}_{-1.3}$	$8.16^{+0.29}_{-0.24}$	7.5
$W_\odot$	$7.14^{+0.19}_{-0.19}$	$7.35^{+0.19}_{-0.18}$	$7.377^{+0.092}_{-0.087}$	7.25
$\sigma_\phi^{\rm thin}$	$28.8^{+1.1}_{-1}$	$28.4^{+1.1}_{-1.1}$	$27.7^{+0.42}_{-0.5}$	28.3
$\sigma_z^{\rm thin}$	$25.03^{+0.86}_{-0.84}$	$25.89^{+0.87}_{-0.86}$	$25.09^{+0.6}_{-0.72}$	25
$\sigma_R^{\rm thin}$	$38.5^{+1.7}_{-1.6}$	$42.7^{+1.6}_{-1.6}$	$40.45^{+0.56}_{-0.84}$	40
$\sigma_\phi^{\rm thick}$	$47.8^{+3.1}_{-2.9}$	$43.7^{+3.2}_{-3.1}$	$42.02^{+0.45}_{-0.4}$	42.4
$\sigma_z^{\rm thick}$	$34.1^{+2.3}_{-2.1}$	$32.8^{+2.3}_{-2.3}$	$35.19^{+0.58}_{-0.52}$	35
$\sigma_R^{\rm thick}$	$63.3^{+3.8}_{-3.8}$	$55.3^{+4.2}_{-4}$	$60.62^{+0.55}_{-0.68}$	60
β_R	$0.183^{+0.025}_{-0.025}$	$0.249^{+0.021}_{-0.023}$	$0.2079^{+0.0094}_{-0.015}$	0.2
β_z	$0.38^{+0.022}_{-0.022}$	$0.401^{+0.02}_{-0.022}$	$0.368^{+0.025}_{-0.03}$	0.37
β_ϕ	$0.216^{+0.023}_{-0.022}$	$0.197^{+0.022}_{-0.023}$	$0.177^{+0.013}_{-0.016}$	0.2
$q_{\rm thin}$	$0.36^{+0.17}_{-0.16}$	$0.14^{+0.19}_{-0.15}$	$0.18^{+0.012}_{-0.014}$	0.18
$q_{\rm thick}$	$0.267^{+0.099}_{-0.086}$	$0.33^{+0.16}_{-0.19}$	$0.3352^{+0.0072}_{-0.0072}$	0.33
v_0	233	265^{+63}_{-60}	$236^{+1.7}_{-1.4}$	233
$R_\odot$	8	8	8	8
α_z	0.047	0.047	$0.0432^{+0.0015}_{-0.0019}$	0.047
α_R	0	0	0	0
χ^2	1.09	1.00	0.935	

TABLE 3
TESTS ON MOCK DATA: CONSTRAINTS ON MODEL PARAMETERS WITH SHU DISTRIBUTION FUNCTION

Model	GCS SHU	GCS SHU	RAVE SHU	Input
$U_\odot$	$11.28^{+0.42}_{-0.41}$	$11.16^{+0.42}_{-0.41}$	$11.27^{+0.12}_{-0.14}$	11.1
$V_\odot$	$7.14^{+0.34}_{-0.36}$	$7.35^{+0.79}_{-0.67}$	$7.94^{+0.17}_{-0.15}$	7.5
$W_\odot$	$6.95^{+0.19}_{-0.2}$	$6.99^{+0.2}_{-0.2}$	$7.26^{+0.079}_{-0.088}$	7.25
$\sigma_z^{\rm thin}$	$25.18^{+0.84}_{-0.84}$	$24.9^{+0.95}_{-0.92}$	$24.62^{+0.81}_{-0.65}$	25
$\sigma_R^{\rm thin}$	$41^{+1.1}_{-1.1}$	$40.7^{+1.1}_{-1.2}$	$41.19^{+0.47}_{-0.6}$	40
$\sigma_z^{\rm thick}$	$36.8^{+2.6}_{-2.4}$	$32.3^{+2.4}_{-2.5}$	$34.3^{+0.52}_{-0.51}$	35
$\sigma_R^{\rm thick}$	$45.2^{+3.6}_{-3.5}$	$44.3^{+3.9}_{-4}$	$46.1^{+0.61}_{-0.58}$	45
β_R	$0.203^{+0.016}_{-0.016}$	$0.201^{+0.017}_{-0.017}$	$0.211^{+0.01}_{-0.013}$	0.2
β_z	$0.379^{+0.021}_{-0.021}$	$0.371^{+0.023}_{-0.024}$	$0.331^{+0.036}_{-0.025}$	0.37
$q_{\rm thin}$	$0.174^{+0.018}_{-0.019}$	$0.186^{+0.038}_{-0.027}$	$0.1706^{+0.0068}_{-0.0066}$	0.18
$q_{\rm thick}$	$0.333^{+0.04}_{-0.039}$	$0.326^{+0.044}_{-0.041}$	$0.3268^{+0.0063}_{-0.0068}$	0.33
v_0	233	224^{+33}_{-20}	$235.1^{+1.3}_{-1.3}$	233
$R_\odot$	8	8	8	8
α_z	0.047	0.047	$0.0427^{+0.0019}_{-0.0018}$	0.047
α_R	0	0	0	0
χ^2	0.960	0.996	0.928	

5.1. *Gaussian models*

First we concentrate on GCS data (column 1 of Table 5). For GCS we find that all the values are well constrained. However, percentage wise $q_{\rm thin}, q_{\rm thick}$ and $V_\odot$ have larger uncertainties as compared to other parameters. In Figure 5, where fits from column 1 are plotted it can be seen that the model is an acceptable fit to the data. The reduced χ^2 values are quite high especially in comparison to the mock models. This is mainly due to significant amount of structure in (U, V) velocity space (see Figure 5). The β_z, $\sigma_z^{\rm thin}$, $\sigma_z^{\rm thick}$ and $\sigma_R^{\rm thick}$ parameters are close to the corresponding Besançon values

TABLE 4
FIDUCIAL MODEL PARAMETERS: VELOCITY IN UNITS OF $\mathrm{km\,s^{-1}}$

Model	*Galaxia*	Besançon
$U_\odot$	11.1	10.3
$V_\odot$	12.24	6.3
$W_\odot$	7.25	5.9
$\sigma_\phi^{\rm thin}$	32.3	32.3
$\sigma_z^{\rm thin}$	21	21
$\sigma_R^{\rm thin}$	50	50
$\sigma_\phi^{\rm thick}$	51	51
$\sigma_z^{\rm thick}$	42	42
$\sigma_R^{\rm thick}$	67	67
β_R	0.33	0.33
β_z	0.33	0.33
β_ϕ	0.33	0.33
$\tau_{\rm sat}$	6.5 Gyr	6.5 Gyr
$q_{\rm thin}$	0.33	0.24(0.285)
$q_{\rm thick}$	0.33	0.44(0.5)
R_0	8.0 kpc	8.5 kpc
$v_{\rm c}(R_0)$	226.84	220.0

but other show differences. The most notable differences are that our value for $q_{\rm thin}$ is higher, $q^{\rm thick}$ is lower and so is $\sigma_\phi^{\rm thick}$. Other minor differences are as follows. Our β_R and β_ϕ are lower and so are the velocity dispersions $\sigma_R^{\rm thin}$, $\sigma_\phi^{\rm thin}$. The thin disc velocity dispersions are strongly correlated to β values, so fixing β to higher values will drive the corresponding thin disc velocity dispersions closer to the Besançon values. The second column in Table 5 shows the results for the case where a separate thick disc is not assumed (the thick disc stars are labelled as thin disc in the model). In this case, β, σ and $q_{\rm thin}$ are found to increase, , which is expected since the thin disc has to accommodate for the warmer thick disc component.

We now discuss results for the RAVE data, beginning with the model where $\alpha_z = 0$ (column 4 of Table 5). Surprisingly, $q_{\rm thin}$ is found to be negative, whereas the $q_{\rm thick}$ is positive. The value of v_0 is found to be significantly less than that reported in literature. The β_R and β_ϕ values are also too small. We note that the β_z value in RAVE has more uncertainty than that in GCS, which we had also noted in the tests on mock data. From now on we keep $\beta_z = 0.37$, a value we get in GCS. We checked and found that fixing β_z has negligible impact on other parameters.

We now let α_z free and this results in higher value of v_0. The $\Omega_\odot$ is now close to Sgr A* proper motion. Allowing for a vertical dependence of effective circular velocity increases $q_{\rm thin}$, β_R and β_ϕ. However, these values are still too low compared to GCS values. It can be seen from red lines in Figure 6 that the model does not fit well the projected V component of the velocity. Clearly there are some problems with this model.

We now compare RAVE and GCS results using columns 6 and 3 where we fix $q_{\rm thin}$, $q_{\rm thick}$, α_z to values that we will later get from Shu model. Having the same value of q in both RAVE and GCS makes it easier to compare the other parameters. Note, fixing some of the variables generally leads to an increased χ^2 and this is expected since we are moving away from best fit values. We find that most of the values agree to within 4σ of each other. The two exceptions are β_ϕ and $V_\odot$ which are higher for GCS.

To summarize, we find that the model parameters that best fit the RAVE data show important differences from those for GCS. The models differ mostly in their values of $q_{\rm thin}$ and $q_{\rm thick}$, with the RAVE values being systematically too low. If $q_{\rm thin}$ and $q_{\rm thick}$ are fixed to be same then $V_\odot$ in RAVE is found to be lower by about $2\,\mathrm{km\,s^{-1}}$. The β_ϕ and β_R are also slightly lower in RAVE and are better constrained than β_z.

5.2. *Shu models*

First, we discuss RAVE results for the case where most of the parameters were kept free (column 6 of Table 6). We find that $q_{\rm thin}$ is positive and well above zero unlike for the Gaussian model. It can be seen from Figure 6 that the wings of the V component of velocity are fit better by the Shu model as compared to the Gaussian model. Another important feature is that σ_R for the thick disc is almost the same as that of the thin disc. The σ_z values are also not too far apart. Apparently, as compared to Gaussian model, the velocity dispersions for the thick disc are very similar to that of old thin disc in the Shu model. However, $q_{\rm thick}$ is larger than $q_{\rm thin}$. If α_z is set to zero then v_0 is found to be underestimated (column 4). Setting the prior of Sgr A* proper motion also allows to constrain the radial gradient of circular velocity which is found to be less than $1\,\mathrm{km\,s^{-1}\,kpc^{-1}}$ (column 7). Comparing, columns 5 and 6 it can be seen that fixing β_z to 0.37 mainly changes $\sigma_z^{\rm thin}$ while the other parameters are relatively unaffected.

The thick disc parameters for the GCS sample (column 1 of Table 6) differ significantly from those of RAVE sample. This is mainly due to the GCS having very few thick disc stars. We next fix $q_{\rm thick} = 0.33$ and $q_{\rm thin} = 0.1825$ for GCS. Doing so improves the agreement between the two sets for thick disc while the change in χ^2 is very little (column 2). Most RAVE parameters agree to within 4σ of GCS except for $V_\odot$, which is lower by about $2\,\mathrm{km\,s^{-1}}$ for RAVE. Finally we also test models where the thick disc is ignored (column 3). As in the case of Gaussian models, this leads to an increase in β, σ and $q_{\rm thin}$.

In Figure 5 the best fit Gaussian and Shu models for GCS are compared. Unlike RAVE both models provide good fits. In fact to discriminate the models one requires a large number of warm stars that can sample the wings of the V distributions with adequate resolution. The GCS sample clearly lacks these characteristics. Next, in Figure 7 we plot the GCS Shu model alongside the RAVE Shu model (columns 2 and 6 of Table 6) and compare them with the GCS velocities. It can be seen that both are acceptable fits. However, the RAVE Shu model slightly overestimates the right wing of the GCS V distribution. Note, in Figure 6 a slight mismatch at $V' \sim 0$ can be seen, the cause for this is not yet clear.

6. DISCUSSION

6.1. *Correlations and degeneracies*

Not all parameters are independent. The dominant correlations are shown in Figure 10, Figure 11, Figure 12 and Figure 13 where pairwise posterior distributions of parameters are plotted. The implication of any correlation is that a change in one of the values also changes the other value *without affecting the quality of the fit.* In other words, a precise value of one correlated quantity needs to be known in order to determine the other. We find that the β values are strongly correlated with the corresponding $\sigma^{\rm thin}$ values. This is mainly because we do not have enough information in the data to estimate the ages of the stars. The model specifies the prior on the ages of stars and the data gives the velocities. The degeneracy reflects the fact that during fitting β can be adjusted while keeping the mean velocity dispersion constant.

TABLE 5
CONSTRAINTS ON MODEL PARAMETERS WITH THE GAUSSIAN DISTRIBUTION FUNCTION. PARAMETERS THAT DO NOT HAVE ERROR BARS WERE FIXED. MISSING VALUES IMPLY PARAMETERS THAT ARE NOT APPLICABLE FOR THAT MODEL. THE MODEL RUNS ARE NAMED AS FOLLOWS; SURVEY NAME AS RAVE OR GCS, TYPE OF MODEL AS GU FOR GAUSSIAN AND SHU FOR SHU. VELOCITIES ARE IN $\mathrm{km\,s^{-1}}$ AND DISTANCES IN kpc

Model	GCS GU	GCS GU	GCS GU	RAVE GU	RAVE GU	RAVE GU
$U_\odot$	$10.16^{+0.41}_{-0.42}$	$10.28^{+0.43}_{-0.43}$	$10.34^{+0.42}_{-0.42}$	$11.66^{+0.16}_{-0.15}$	$11.45^{+0.14}_{-0.14}$	$11.25^{+0.15}_{-0.15}$
$V_\odot$	$6.6^{+1.3}_{-1.4}$	$6.33^{+0.93}_{-0.97}$	$9.68^{+0.26}_{-0.26}$	$15.01^{+0.37}_{-0.42}$	$8^{+0.3}_{-0.28}$	$7.38^{+0.1}_{-0.12}$
$W_\odot$	$7.14^{+0.19}_{-0.18}$	$7.11^{+0.19}_{-0.19}$	$7.14^{+0.18}_{-0.18}$	$7.692^{+0.099}_{-0.082}$	$7.688^{+0.085}_{-0.091}$	$7.625^{+0.088}_{-0.082}$
$\sigma_\phi^{\rm thin}$	$27.12^{+0.89}_{-0.86}$	$31.61^{+0.8}_{-0.79}$	$27.83^{+0.88}_{-0.88}$	$24.97^{+0.43}_{-0.36}$	$25.56^{+0.33}_{-0.37}$	$25.34^{+0.35}_{-0.33}$
$\sigma_z^{\rm thin}$	$23.74^{+0.79}_{-0.74}$	$27.28^{+0.64}_{-0.63}$	$23.89^{+0.79}_{-0.74}$	$24.22^{+0.64}_{-0.47}$	$25.69^{+0.22}_{-0.2}$	$25.92^{+0.21}_{-0.2}$
$\sigma_R^{\rm thin}$	$41.2^{+1.4}_{-1.3}$	$47^{+1.1}_{-1.1}$	$41.5^{+1.4}_{-1.3}$	$36.6^{+1}_{-1.1}$	$39.26^{+0.67}_{-0.69}$	$39.69^{+0.62}_{-0.65}$
$\sigma_\phi^{\rm thick}$	$40.9^{+3.3}_{-3.1}$		$40^{+2.9}_{-2.8}$	$40.47^{+0.51}_{-0.48}$	$37.16^{+0.5}_{-0.53}$	$38.37^{+0.48}_{-0.54}$
$\sigma_z^{\rm thick}$	$38.5^{+2.8}_{-2.5}$		$38.7^{+2.7}_{-2.6}$	$40.55^{+0.46}_{-0.49}$	$40.4^{+0.5}_{-0.5}$	$39.41^{+0.48}_{-0.48}$
$\sigma_R^{\rm thick}$	$65.9^{+4.1}_{-3.7}$		$67.7^{+2.7}_{-2.7}$	$58.74^{+0.91}_{-0.79}$	$58.43^{+0.86}_{-0.76}$	$57.87^{+0.58}_{-0.56}$
β_R	$0.201^{+0.019}_{-0.019}$	$0.268^{+0.015}_{-0.014}$	$0.204^{+0.019}_{-0.019}$	$0.06^{+0.023}_{-0.029}$	$0.135^{+0.015}_{-0.015}$	$0.164^{+0.012}_{-0.013}$
β_z	$0.36^{+0.02}_{-0.021}$	$0.432^{+0.015}_{-0.016}$	$0.365^{+0.02}_{-0.02}$	$0.312^{+0.026}_{-0.02}$	0.37	0.37
β_ϕ	$0.271^{+0.019}_{-0.019}$	$0.349^{+0.016}_{-0.015}$	$0.284^{+0.019}_{-0.019}$	$0.132^{+0.014}_{-0.013}$	$0.17^{+0.012}_{-0.012}$	$0.164^{+0.012}_{-0.012}$
$q_{\rm thin}$	$0.43^{+0.11}_{-0.11}$	$0.447^{+0.069}_{-0.068}$	0.1825	$-0.139^{+0.02}_{-0.019}$	$0.047^{+0.014}_{-0.013}$	0.1825
$q_{\rm thick}$	$0.37^{+0.099}_{-0.087}$		0.33	$0.281^{+0.011}_{-0.011}$	$0.2266^{+0.0088}_{-0.0089}$	0.33
v_0	226.84	226.84	233	$207.2^{+1.9}_{-1.9}$	$229.2^{+1.8}_{-2}$	$234.1^{+1.4}_{-1.4}$
$R_\odot$	8	8	8	8	8	8
α_z	0	0	0.047	0	$0.0738^{+0.0021}_{-0.0023}$	0.047
α_R	0	0	0	0	0	0
χ^2 RAVE	2.55	3.19	2.49	1.89	1.64	1.79
χ^2 GCS	3.09	3.48	3.15	6.60	5.81	5.10

TABLE 6
CONSTRAINTS ON MODEL PARAMETERS WITH THE SHU DISTRIBUTION FUNCTION. SEE TABLE 5 FOR FURTHER DESCRIPTION.

Model	GCS SHU	GCS SHU	GCS SHU	RAVE SHU	RAVE SHU	RAVE SHU	RAVE SHU
$U_\odot$	$10.02^{+0.39}_{-0.4}$	$10.16^{+0.39}_{-0.4}$	$10.23^{+0.39}_{-0.4}$	$11.2^{+0.13}_{-0.13}$	$10.92^{+0.13}_{-0.14}$	$10.96^{+0.14}_{-0.13}$	$11.05^{+0.15}_{-0.16}$
$V_\odot$	$9.95^{+0.3}_{-0.3}$	$9.81^{+0.28}_{-0.28}$	$9.83^{+0.3}_{-0.29}$	$9.71^{+0.12}_{-0.11}$	$7.53^{+0.16}_{-0.16}$	$7.53^{+0.16}_{-0.16}$	$7.62^{+0.13}_{-0.16}$
$W_\odot$	$7.14^{+0.19}_{-0.19}$	$7.13^{+0.18}_{-0.19}$	$7.12^{+0.18}_{-0.19}$	$7.536^{+0.085}_{-0.086}$	$7.542^{+0.089}_{-0.093}$	$7.539^{+0.095}_{-0.09}$	$7.553^{+0.086}_{-0.09}$
$\sigma_z^{\rm thin}$	$23.39^{+0.77}_{-0.73}$	$23.63^{+0.85}_{-0.8}$	$25.91^{+0.64}_{-0.6}$	$26.85^{+0.85}_{-0.92}$	$24.7^{+0.66}_{-0.66}$	$25.73^{+0.21}_{-0.21}$	$25.72^{+0.23}_{-0.25}$
$\sigma_R^{\rm thin}$	$38.14^{+0.98}_{-0.94}$	$39.99^{+0.91}_{-0.91}$	$42.71^{+0.83}_{-0.8}$	$42.37^{+0.61}_{-0.66}$	$39.78^{+0.81}_{-0.73}$	$39.67^{+0.63}_{-0.72}$	$39.56^{+0.66}_{-0.7}$
$\sigma_z^{\rm thick}$	$39^{+3.1}_{-3.3}$	$32.6^{+2.3}_{-2.2}$		$29.15^{+0.87}_{-0.79}$	$34.66^{+0.61}_{-0.58}$	$34.3^{+0.51}_{-0.57}$	$34.48^{+0.54}_{-0.53}$
$\sigma_R^{\rm thick}$	$70.1^{+3.7}_{-5.5}$	$45.9^{+1.8}_{-1.8}$		$38.84^{+1.2}_{-0.96}$	$42.31^{+1}_{-0.9}$	$42.43^{+0.95}_{-1}$	$43.23^{+0.96}_{-1.1}$
β_R	$0.213^{+0.014}_{-0.014}$	$0.237^{+0.013}_{-0.013}$	$0.273^{+0.011}_{-0.011}$	$0.236^{+0.011}_{-0.011}$	$0.198^{+0.014}_{-0.014}$	$0.195^{+0.011}_{-0.013}$	$0.192^{+0.012}_{-0.013}$
β_z	$0.361^{+0.02}_{-0.02}$	$0.366^{+0.021}_{-0.021}$	$0.415^{+0.016}_{-0.016}$	$0.398^{+0.03}_{-0.029}$	$0.328^{+0.027}_{-0.024}$	0.37	0.37
$q_{\rm thin}$	$0.166^{+0.021}_{-0.021}$	0.1825	$0.193^{+0.015}_{-0.015}$	$0.1682^{+0.0071}_{-0.0071}$	$0.1805^{+0.0086}_{-0.008}$	$0.1825^{+0.0092}_{-0.0075}$	$0.1809^{+0.0078}_{-0.0078}$
$q_{\rm thick}$	$0.022^{+0.054}_{-0.016}$	0.33		$0.389^{+0.011}_{-0.016}$	$0.334^{+0.012}_{-0.014}$	$0.332^{+0.012}_{-0.013}$	$0.325^{+0.014}_{-0.012}$
v_0	226.84	232	226.84	$212.6^{+1.4}_{-1.3}$	$232.8^{+1.7}_{-1.6}$	$231.9^{+1.4}_{-1.5}$	$235.02^{+0.86}_{-0.83}$
$R_\odot$	8	8	8	8	8	8	8
α_z	0	0.0471	0	0	$0.048^{+0.0019}_{-0.0018}$	$0.0471^{+0.0016}_{-0.0019}$	$0.0471^{+0.0019}_{-0.0019}$
α_R	0	0	0	0	0	0	$0.67^{+0.25}_{-0.26}$
χ^2 RAVE	2.07	1.80	2.40	1.52	1.43	1.42	1.42
χ^2 GCS	3.85	3.86	4.08	5.15	5.57	5.42	5.46

The $\sigma_R^{\rm thick}$ is anti-correlated with $q_{\rm thick}$ and $\sigma_R^{\rm thin}$ with $q_{\rm thin}$. These two anti-correlations are stronger for the Shu model as compared to the Gaussian model. To get a good estimate of q, which controls the radial gradient of dispersions, ideally one would require a sample of stars distributed over a large volume. In the absence of an extended sample the constraint on q comes from the fact that it also determines the asymmetric drift. The amount of asymmetric drift increases with an increase in both σ_R and q (see Equation (5)). If the asymmetric drift is fixed, this naturally leads to the anti-correlation between q and σ_R. In the Shu model the effective velocity dispersion $\sqrt{\langle v_R^2 \rangle}$ is not only proportional to σ_R but also increase with increase in q. So one can keep the effective velocity dispersion constant by increasing q and decreasing σ_R at the same time. This makes the anti-correlation in Shu model stronger.

Also, $V_\odot$ is anti-correlated with $q_{\rm thin}$ and this relation is stronger for the Gaussian model. This makes it difficult to determine $V_\odot$ and q reliably using the Gaussian models. The Shu model does not have this problem because in it the azimuthal motion is coupled to the radial motion, so it has 3 fewer parameters, i.e., has less degree of freedom. This helps to resolve the $q_{\rm thin} - V_\odot$ degeneracy.

When fitting Shu models to RAVE we find an anti-correlation exists between thin and thick disc parameters, e.g., $(\sigma_R^{\rm thin}, \sigma_R^{\rm thick})$, $(\sigma_z^{\rm thin}, \sigma_z^{\rm thick})$ and $(q_{\rm thin}, q_{\rm thick})$. This mainly because we do not have any useful information about the ages in the data.

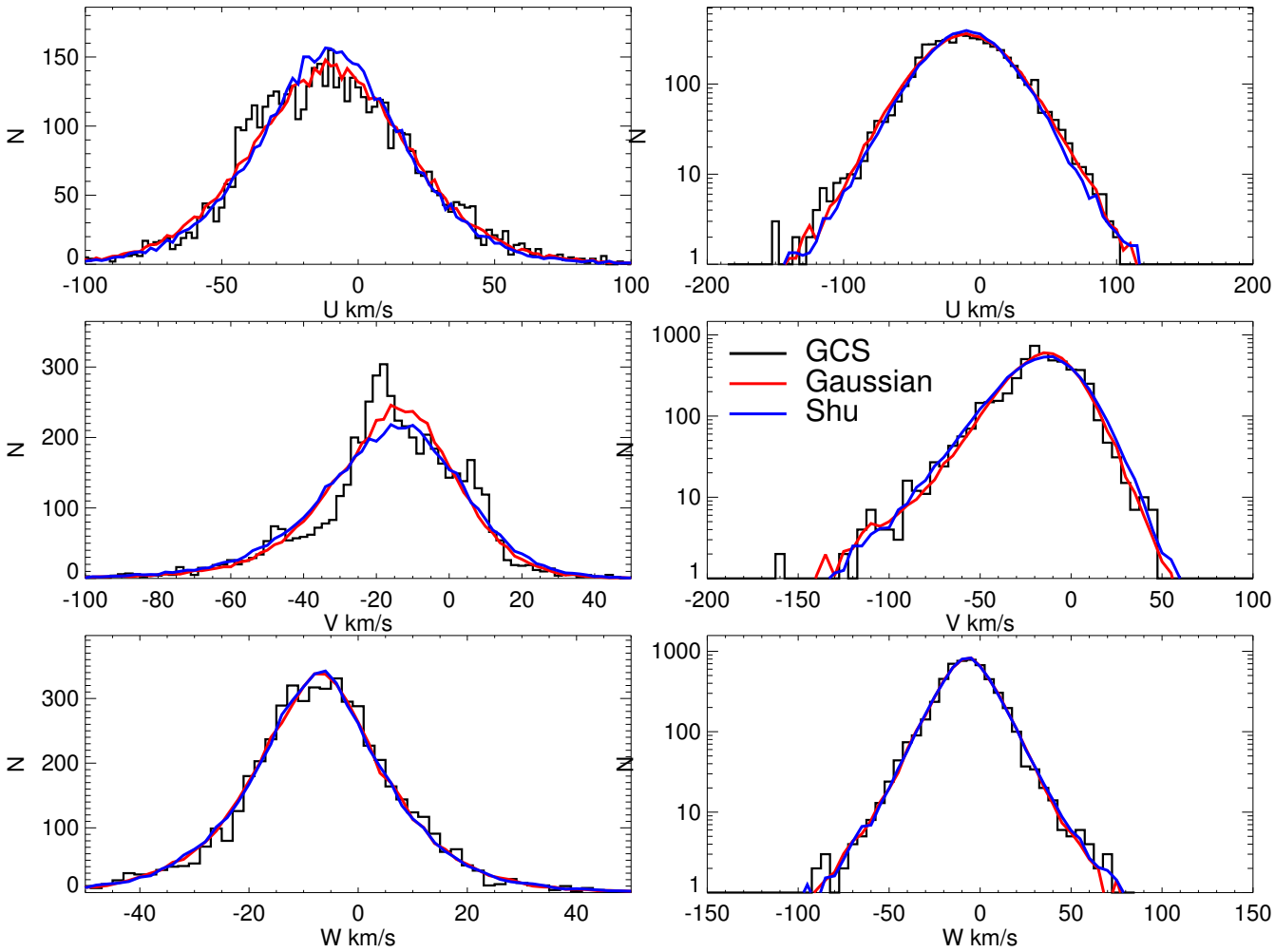

FIG. 5.— Comparison of model velocity distributions with that of GCS data. The right panels differ from the left only in range and scale of axes. The model used is the best fit Gaussian (column 1 of Table 5) and the Shu model (column 1 of Table 6) for the GCS data. Both the models are acceptable fits to the data. Significant structures can be seen in the velocity space.

We now discuss parameters v_0 and α_z which were kept free only in RAVE data. The α_z is correlated with v_0 and anti-correlated with $V_\odot$. The v_0 parameter is anti-correlated with $U_\odot$ and correlated with $q_{\rm thin}$. For the GCS data, the $(v_0, q_{\rm thin})$ correlation is so strong that it is difficult to get meaningful constraints on v_0 so the later was kept fixed.

6.2. *Solar peculiar motion*

Among the three components of Solar motion, the $U_\odot$ and $W_\odot$ are only weakly correlated with other variables and give similar values for both Gaussian and Shu models. The only major dependence of $U_\odot$ is for RAVE where it is anti-correlated with v_0 by about -0.5. So models with $\alpha_z = 0$ that underestimate v_0, will overestimate $U_\odot$. For RAVE we get $W_\odot = 7.54 \pm 0.1$ km s^{-1} and $U_\odot = 10.96 \pm 0.14$ km s^{-1} (column 6 of Table 6). GCS values for $W_\odot$ and $U_\odot$ are lower by about 0.4 and 0.8 km s^{-1} respectively but their $3-\sigma$ range matches with RAVE (column 2 of Table 6). The small mismatch could be either due to large-scale gradients in the mean motion of stars (Williams et al. 2013) in RAVE or due to kinematic substructures in GCS.

Our GCS results (column 2 of Table 6) are in excellent agreement with Dehnen & Binney (1998b), but differ with Schönrich et al. (2010) for $U_\odot$ by 1.0 km s^{-1}. Nevertheless, $U_\odot$ is well within their quoted 2σ range. The RAVE $U_\odot$ agrees with Schönrich et al. (2010). Interestingly, with the aid of a model-independent approach, Schönrich (2012) finds $U_\odot = 14.0 \pm 0.3$ km s^{-1} but with a systematic uncertainty of 1.5 km s^{-1}. The systematic errors in distances and proper motion can bias this result. Additionally, the analyzed sample not being local, their results can also be biased if there are large-scale streaming motions.

We now discuss our results for $V_\odot$. For Gaussian models the estimated $V_\odot$ value depends strongly upon the choice of q values and it is difficult to get a reliable value for either of them. For the Shu model, $V_\odot$ depends upon if α_z is kept fixed or free. For $\alpha_z = 0$, the GCS and RAVE $V_\odot$ agree with each other, but when α_z is kept free RAVE $V_\odot$ is 2 km s^{-1} lower than that of GCS (comparing columns 2 and 6 of Table 6). The $\alpha_z = 0$ model not only has a higher χ^2 but as we will discuss later also yields a low value of v_0, so we consider this model less useful. The most likely cause for the difference between RAVE and GCS $V_\odot$ is the significant amount of kinematic substructures in the distribution of the V component of the GCS velocities (Figure 7). It can be seen in Figure 7 that the best fit RAVE model, in spite of apparently having low $V_\odot$, is still a good description of the GCS data. Moreover, in GCS a dominant kinematic structure can be seen at $V \sim -20$ km s^{-1} (the Hyades and the Pleiades), lending further support to the idea that GCS probably overestimates $V_\odot$.

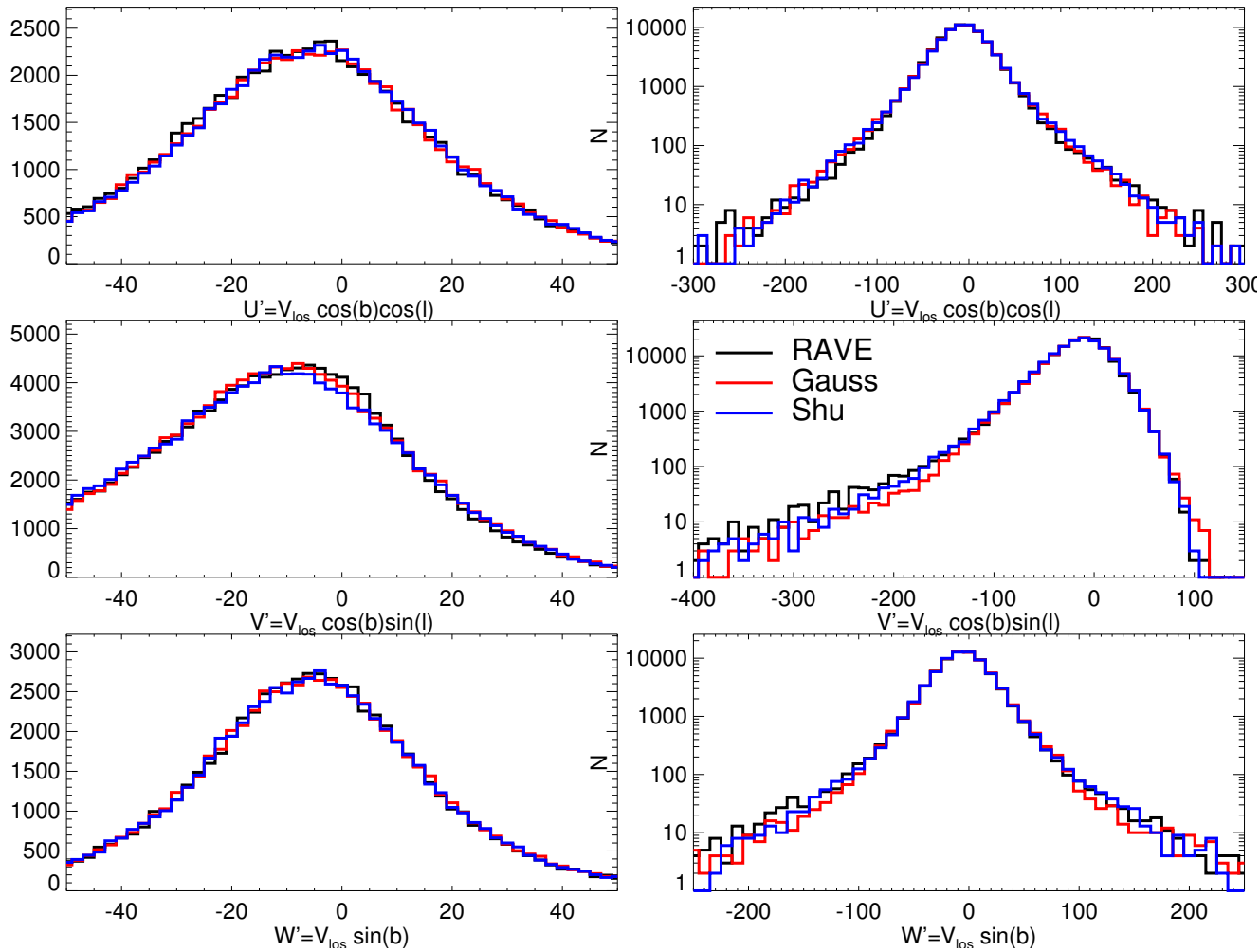

FIG. 6.— Comparison of model velocity distributions with that of RAVE data. Projection of radial velocity along U, V and W directions are shown. The right panels differ from the left only in range and scale of axes. The top panel is for stars with $(|b| < 45)\&((|l| > 45)||(|l - 180| > 45))$, the middle panel is for stars with $(|b| < 45)\&((|l| < 45)||(|l - 180| < 45))$ and the bottom panel is for stars with $|b| > 45$. The model used is the best fit Gaussian (column 5 in Table 5) and the Shu model (column 6 of Table 6) for the RAVE data. The Shu model clearly models the wings of V' better than the Gaussian model, especially in region -200 km s$^{-1} < V' < -150$ km s^{-1} and $V' > 80$ km s^{-1} which is dominated by thick disc. A slight mismatch at $V' \sim 0$ is also seen.

The need to revise the $V_\odot$ upwards from the value of 5.2 km s^{-1} given by Dehnen & Binney (1998b) has been extensively discussed (McMillan & Binney 2010; Binney 2010; Schönrich et al. 2010). Binney (2010) suggests a value of 11.0 km s^{-1} after randomizing some of the stars to reduce the impact of streams while Schönrich et al. (2010) get $V_\odot = 12.24 \pm 0.47$ km s^{-1}. Our RAVE value of $V_\odot = 7.5 \pm 0.2$ is significantly lower that this (column 6 of Table 6). Our GCS value of $V_\odot = 9.8 \pm 0.3$ km s^{-1} is also lower than both of them (column 2 of Table 6).

Recently, Golubov et al. (2013) determined $V_\odot = 3.06 \pm 0.68$ by binning the local RAVE stars in color and metallicity bins and applying an improved version of the Stromberg relation. Although their estimate is on the lower side like ours, but it is significantly lower than our estimates. The application of the Stromberg relation demands the identification of subpopulations that are relaxed, are in dynamical equilibrium and have same value for the slope in the relation. The fact that binning by color fails to satisfy this has already been discussed in Schönrich et al. (2010). Although the authors also split the sample by metallicity but these bins are quite broad and there is uncertainty associated with metallicity measurements. Hence, a bias due to the selected subpopulations not obeying the same linear relation can be expected.

The discrepancy for the GCS with Schönrich et al. (2010) could be either due to differences in fitting methodologies or differences in the models adopted, with the latter being the most likely cause. The model used here and by Schönrich et al. (2010) is based on the Shu distribution function but still there are some important differences. For example, we have a thick disc while they do not, and the forms of $\sigma_R(L)$ and $\Sigma(L)$ differ. Our form of $\sigma_R(L)$ is the same as that used by Binney (2010) while Schönrich et al. (2010) compute $\sigma_R(L)$ so as to satisfy $\langle v^2_{R,\rm thin}\rangle \propto {\rm e}^{-R/1.5R_d}$. In our case $\langle v^2_{R,\rm thin}\rangle(R)$ depends implicitly upon q and β and both of these parameters are constrained by data. The $\Sigma(L)$ in Schönrich et al. (2010) comes from a numerical simulation involving the processes of accretion, churning and blurring while in our case it comes directly from the constraint that $\Sigma(R) \propto \exp(-R/R_d)$. The prescription for metallicity in Schönrich et al. (2010) is also very different from ours.

6.3. *The circular velocity*

In a recent paper, Bovy et al. (2012a) used data from the APOGEE survey and analyzed stars close to the mid-plane of the disc to find $v_{\rm c} = 218 \pm 6$ km s^{-1} and $V_\odot = 26 \pm 3$

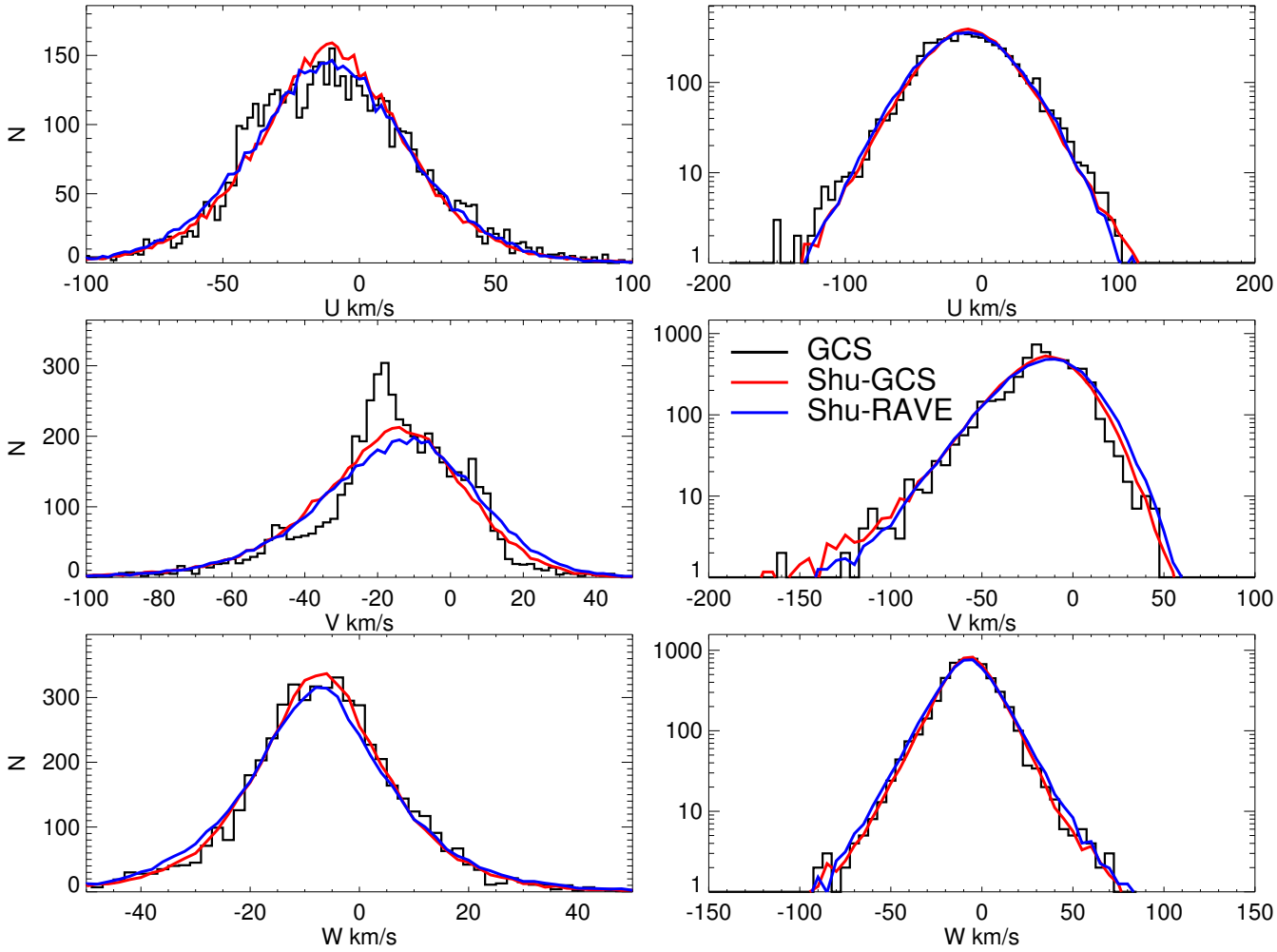

FIG. 7.— Comparison of model velocity distributions with that of GCS data. The models used correspond to columns 2 and 6 of Table 6. These are Shu models that a) best fit the GCS data but with a few parameters fixed and b) best fit the RAVE data. The positive wing of V is slightly overestimated by the RAVE best fit model.

TABLE 7
INVESTIGATION OF SYSTEMATICS.

Model	RAVE SHU	RAVE SHU	RAVE SHU	RAVE SHU	RAVE SHU	RAVE SHU	RAVE SHU
Distance Offset					90%	110%	
Iterations	200000.00	200000.00	200000.00	200000.00	200000.00	188501.00	70101.000
$U_\odot$	$10.96^{+0.14}_{-0.13}$	$11.05^{+0.15}_{-0.16}$	$10.81^{+0.15}_{-0.14}$	$10.98^{+0.14}_{-0.15}$	$11.01^{+0.13}_{-0.14}$	$10.82^{+0.15}_{-0.14}$	$10.73^{+0.13}_{-0.15}$
$V_\odot$	$7.53^{+0.16}_{-0.16}$	$7.62^{+0.13}_{-0.16}$	$7.39^{+0.14}_{-0.14}$	$7.59^{+0.16}_{-0.14}$	$8.26^{+0.15}_{-0.15}$	$6.81^{+0.15}_{-0.16}$	$6.99^{+0.15}_{-0.15}$
$W_\odot$	$7.539^{+0.095}_{-0.09}$	$7.553^{+0.086}_{-0.09}$	$7.52^{+0.085}_{-0.088}$	$7.535^{+0.082}_{-0.089}$	$7.553^{+0.078}_{-0.091}$	$7.53^{+0.09}_{-0.083}$	$7.52^{+0.087}_{-0.091}$
$\sigma_z^{\rm thin}$	$25.73^{+0.21}_{-0.21}$	$25.72^{+0.23}_{-0.25}$	$25.69^{+0.22}_{-0.2}$	$25.67^{+0.23}_{-0.23}$	$25.68^{+0.25}_{-0.21}$	$25.77^{+0.18}_{-0.22}$	$17.57^{+0.12}_{-0.11}$
$\sigma_R^{\rm thin}$	$39.67^{+0.63}_{-0.72}$	$39.56^{+0.66}_{-0.7}$	$39.27^{+0.56}_{-0.62}$	$39.45^{+0.67}_{-0.61}$	$39.23^{+0.74}_{-0.6}$	$40.09^{+0.59}_{-0.49}$	$31.21^{+0.11}_{-0.13}$
$\sigma_z^{\rm thick}$	$34.3^{+0.51}_{-0.57}$	$34.48^{+0.54}_{-0.53}$	$34.48^{+0.58}_{-0.56}$	$34.66^{+0.52}_{-0.55}$	$34.8^{+0.55}_{-0.6}$	$33.8^{+0.55}_{-0.55}$	$37.99^{+0.48}_{-0.46}$
$\sigma_R^{\rm thick}$	$42.43^{+0.95}_{-1}$	$43.23^{+0.96}_{-1.1}$	$42.98^{+0.86}_{-0.73}$	$42.67^{+0.96}_{-0.72}$	$43.51^{+0.85}_{-0.82}$	$41.28^{+0.71}_{-0.94}$	$48.6^{+0.55}_{-0.61}$
β_R	$0.195^{+0.011}_{-0.013}$	$0.192^{+0.012}_{-0.013}$	$0.188^{+0.01}_{-0.011}$	$0.192^{+0.013}_{-0.012}$	$0.188^{+0.013}_{-0.013}$	$0.2018^{+0.01}_{-0.0093}$	0.01
β_z	0.37	0.37	0.37	0.37	0.37	0.37	0.01
$q_{\rm thin}$	$0.1825^{+0.0092}_{-0.0075}$	$0.1809^{+0.0078}_{-0.0078}$	$0.188^{+0.0086}_{-0.0086}$	$0.1804^{+0.0071}_{-0.0066}$	$0.2061^{+0.0094}_{-0.0079}$	$0.1577^{+0.0073}_{-0.0068}$	$0.2454^{+0.0082}_{-0.0059}$
$q_{\rm thick}$	$0.332^{+0.012}_{-0.013}$	$0.325^{+0.014}_{-0.012}$	$0.339^{+0.01}_{-0.011}$	$0.3149^{+0.0087}_{-0.012}$	$0.339^{+0.01}_{-0.011}$	$0.3297^{+0.013}_{-0.0091}$	$0.2532^{+0.0073}_{-0.0072}$
v_0	$231.9^{+1.4}_{-1.5}$	$235.02^{+0.86}_{-0.83}$	$223.3^{+1.3}_{-1.4}$	$242.5^{+1.6}_{-1.5}$	$249.8^{+1.6}_{-1.5}$	$218.9^{+1.5}_{-1.4}$	$237.2^{+1.8}_{-1.5}$
$R_\odot$	8	8	7.5	8.5	8	8	8
α_z	$0.0471^{+0.0016}_{-0.0019}$	$0.0471^{+0.0019}_{-0.0019}$	$0.0532^{+0.0017}_{-0.0017}$	$0.0439^{+0.0016}_{-0.0017}$	$0.0504^{+0.0018}_{-0.0018}$	$0.0462^{+0.0016}_{-0.0018}$	$0.0531^{+0.0018}_{-0.002}$
α_R	0	$0.67^{+0.25}_{-0.26}$	0	0	0	0	0

km s^{-1}. The resulting angular velocity $\Omega_\odot = (v_c + V_\odot)/R_0$ agrees with the value of 30.24 ± 0.11 km s^{-1} kpc^{-1} as estimated by Reid & Brunthaler (2004) using the Sgr A* proper motion or as estimated by McMillan & Binney (2010) using masers ($\Omega_\odot$ in range 29.9-31.6 km s^{-1} kpc^{-1}). However, Bovy et al. (2012a) $V_\odot$ is about 14 km s^{-1} larger than what has been measured in the solar neighborhood by GCS. As a way to reconcile their high $V_\odot$, Bovy et al. (2012a) suggest that the LSR itself is rotating with a velocity of ~ 12 km s^{-1} with respect to the RSR (rotational standard of rest as measured by circular velocity in an axis-symmetric approximation of the full potential of the Milky Way).

For RAVE data, we get $v_c = 232 \pm 1.7$ and $\Omega_\odot = 29.9 \pm 0.3$ km s^{-1} kpc^{-1} which agrees with the Sgr A* proper motion of 30.24 ± 0.1. Hence, the RAVE data suggests that the LSR is on a circular orbit and is consistent with RSR. Our $\alpha_z = 0.047$ is slightly higher than the value of 0.0374 predicted by analytical models of the Milky Way potential. This is expected because in our formalism, the parameter α_z also accounts for the increase in asymmetric drift with height. If we explicitly put a prior on $\Omega_\odot$, then we have the liberty of constraining one more parameter and we use it to constrain the radial gradient of circular velocity α_R. Doing so, we find a small gradient of about 0.67 km s^{-1} kpc^{-1} (column 7 of Table 6) and v_0 increases to 235 km s^{-1}.

We find that the parameter α_z that controls the vertical dependence of circular velocity plays an important role in determining v_0. For models with $\alpha_z = 0$, v_0 is underestimated and we end up with $v_0 = 212 \pm 1.4$. This is in rough agreement with Bovy et al. (2012a) but $V_\odot$ is not. The resulting angular velocity $\Omega_\odot$ is also much lower than the value obtained from the proper motion of Sgr A*. On the contrary, if α_z is kept free then we automatically match the Sgr A* proper motion, and we get a $V_\odot$ which is similar to the local GCS sample.

6.4. *The age-velocity dispersion relation (AVR)*

We now discuss our model predictions for the age velocity dispersion of the thin disc, specifically the parameters $\beta_z, \beta_\phi, \beta_R$, $\sigma_z^{\rm thin}, \sigma_\phi^{\rm thin}$ and $\sigma_R^{\rm thin}$. The $\sigma_{\phi,z,R}^{\rm thin}$ values for both GCS and RAVE are found to be in agreement with each other and are similar for both Gaussian and Shu models. The GCS $\beta_{\phi,R,z}$ values were also similar for both Gaussian and Shu models. The RAVE β_R, using the Shu model, was also in agreement with GCS. The β_z was difficult to determine precisely with RAVE, so, we choose to use the corresponding GCS value for it, while performing the fits. The RAVE $\beta_{R,\phi}$ values using the Gaussian model were systematically lower than GCS values. Since the RAVE Gaussian model did not fit the data well, we give less importance to its β values and ignore them for the present discussion. Overall, results in column 1 of Table 5 provide a good representation of our predictions and are shown alongside other literature values in Table 8.

We now discuss our results by comparing them with previous estimates. In the Besançon model, the age-velocity dispersion relation for the thin disc was based on an analysis of Hipparcos stars by Gomez et al. (1997). Sharma et al. (2011) fitted their tabulated values using analytical functions and the values are given in Table 4. Binney et al. (2000) also analyzed a subset of Hipparcos stars and showed that the value of β (considering all the three velocity components together) is around 0.33. In comparison, Just & Jahreiß (2010) used Hipparcos stars to determine $\beta_z = 0.375$.

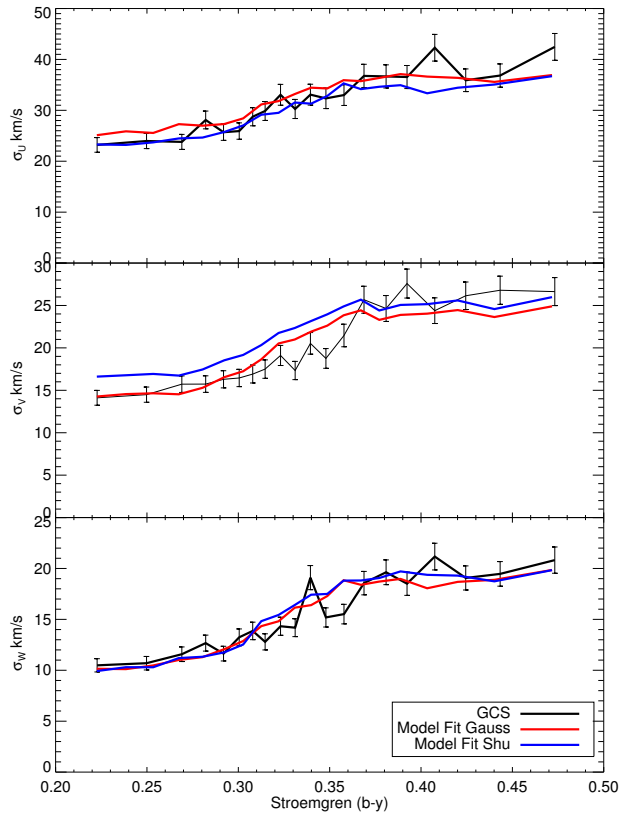

FIG. 8.— Velocity as a function of $b-y$ Strömgren color for GCS stars. The error bars were estimated from Poisson noise. Shown alongside are predictions from various models. Note, the color distribution was not taken into account when fitting models to data.

Aumer & Binney (2009) reanalyzed the Hipparcos data and found $(\beta_R, \beta_\phi, \beta_z) = (0.307, 0.430, 0.445)$. Note, these values were estimated by fitting the velocity dispersion as a function of color and only proper motion was used in the analysis. Their method is sensitive to outliers which must be corrected for prior to estimating the velocity dispersions. Also, the non-Gaussian nature of the v_ϕ distribution is not taken into account. Finally, Nordström et al. (2004) find $(\beta_R, \beta_\phi, \beta_z) = (0.31, 0.34, 0.47)$ in the GCS data. Seabroke & Gilmore (2007) using the same data suggest that error bars are higher and that excluding Hercules stream increases β_z to 0.5. Holmberg et al. (2007) and Holmberg et al. (2009) updated the data with new parallaxes and photometric calibrations and found $(\beta_R, \beta_\phi, \beta_z) = (0.39, 0.40, 0.53)$.

Our values for β are significantly lower than previous studies (see Table 8). While uncertainty in ages remains a big worry in the analysis of Holmberg et al. (2009), the difference between our results with those of Aumer & Binney (2009) is most likely due to different methods. The main difference being that unlike Aumer & Binney (2009) we fit the model to all three velocity components. Also the density laws assumed for distribution of stars in space are different. In Figure 8 we show the velocity dispersion as a function of Stroemgren $b-y$ color. We find that although we do not take color into account, our fitted model correctly reproduces dispersion as a function of color. The Shu model is found to overpredict σ_V for $(b-y) < 0.35$ but only slightly.

We find that the values of β and $\sigma_{R,\phi,z}^{\rm thin}$ (velocity dispersion in the Solar neighborhood for 10 Gyr old stars) varies

TABLE 8
COMPARISON OF VALUES OF β AS ESTIMATED BY DIFFERENT SOURCES

Source	β_R	β_ϕ	β_z
Fit to Robin et al. (2003)	0.33	0.33	0.33
Nordstrom et al. (2004)	0.31 ± 0.05	0.34 ± 0.05	0.47 ± 0.05
Seabroke & Gilmore (2007)			0.48 ± 0.26
Holmberg et al. (2007)	0.38	0.38	0.54
Holmberg et al. (2009)	0.39	0.40	0.53
Aumer & Binney (2009)	0.307	0.430	0.445
Just and Jahreiss (2010)			0.375
Our GCS Thin only	0.27±0.02	0.35±0.02	0.43±0.02
Our GCS Thin+Thick	0.20±0.02	0.27±0.02	0.36±0.02
Our RAVE Thin+Thick	0.19±0.01		0.3-0.4

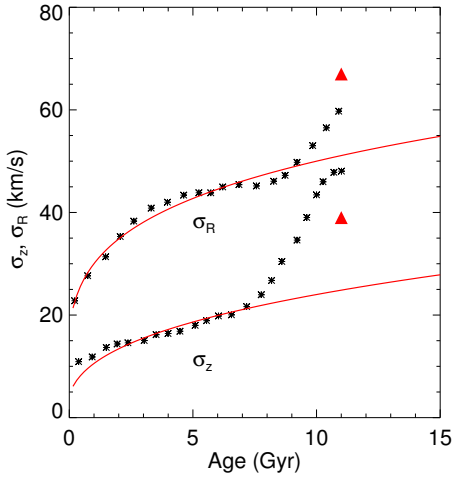

FIG. 9.— Comparison of our age-velocity dispersion relation (solid line) with that of Minchev et al. (2013) (black points). The slopes used are $\beta_z = 0.37$ and $\beta_R = 0.23$ for $\sigma_R = 50.0$ and $\sigma_z = 24.0$. The triangles are for the thick disc in our Gaussian models.

depending upon whether the thick disc is considered as a separate component or not. We find that a single component model (no thick disc) in general gives higher values for these quantities (column 2 of Table 5). For GCS we find $\beta_R < \beta_\phi < \beta_z$. Our β values when the thick disc is excluded are in better agreement with Aumer & Binney (2009) (who do not consider a thick disc) than when it is included. For the velocity dispersions, however, Aumer & Binney (2009) find $(\sigma_R^{\rm thin}, \sigma_\phi^{\rm thin}, \sigma_z^{\rm thin}) = (41.90, 28.82, 23.83)$, which agree better with our results when the thick disc is included.

The ratio of σ_z/σ_R and σ_ϕ/σ_R and the values of β_R, β_ϕ and β_z are useful for understanding the physical processes responsible for heating the disc. Spitzer & Schwarzschild (1953) first showed that scattering of stars by gas clouds can cause an increase in velocity dispersion with age. Hänninen & Flynn (2002) showed that with giant molecular clouds one gets $\beta_R = 0.2$ and $\beta_z = 0.25$. Using $10^7 M_\odot$ massive halo objects (they refer to these as black holes) one gets $\beta_{R,z} \sim 0.5$. The predicted ratio of σ_z/σ_R is between 0.40 and 0.67. The main difference in the two types of perturbers is that the GMCs have a planar distribution while massive halo objects are distributed in a sphere. However, in order to fit the observations, they require a GMC population that is too numerous or too massive (see also Jenkins 1992). Instead, it has been shown that scattering by massive clouds leads to axial ratios that evolve to a steady value of $(\sigma_R : \sigma_z) = (1 : 0.62)$ (Ida et al. 1993; Shiidsuka & Ida 1999; Sellwood 2008).

For RAVE both for Gaussian and Shu models we get $\sigma_z^{\rm thin}/\sigma_R^{\rm thin}$ 0.65 (column 6 of Table 5 and column 6 of Table 6). The corresponding GCS value is 0.58 (column 3 of Table 5 and column 2 of Table 6). Models without a thick disc also give a similar value. These are values for a 10 Gyr old population and we think they agree well with the above predictions. If the different values of β are not same, then the axial ratios of the velocity ellipsoid are also a function of age. If $\beta_z > \beta_R$ then this means σ_z/σ_R increases monotonically with age as $\tau^{\beta_z - \beta_R}$, meaning it is much lower for younger stars. Aumer & Binney (2009) suggest that $\sigma_z/\sigma_R = 0.62$ as predicted by Sellwood (2008) is inconsistent with observations, as much lower values are seen for bluer stars, which are supposed to be younger. We think that their value of 0.57 at 10 Gyr is in reasonable agreement with theoretical predictions, the discrepancy if at all is related to the fact that $\beta_z \neq \beta_R$. For the thick disc we find that the Gaussian model predicts $\sigma_z^{\rm thick}/\sigma_R^{\rm thick} = 0.68$, while the Shu model predicts it be 0.80, which is much higher.

The heating mechanisms discussed so far predict $\beta_R \sim \beta_z$. Another way to heat up the disc is by Lindblad resonances of transient spiral arms that can scatter the stars. This process only increases in-plane dispersions (Carlberg & Sellwood 1985; Sellwood 2013). The predicted values of β_R are between 0.2 for high-velocity stars and 0.5 for low-velocity stars. Multiple spiral density waves (Minchev & Quillen 2006) or a combination of bar and spirals can also heat up the disc (Minchev & Famaey 2010).

Recently, Minchev et al. (2013) investigated the age-velocity dispersion relation for stars seen from simulations of disc galaxies and find it to be in rough agreement with observations. We now compare our results with their findings. In Figure 9, we plot their predictions for σ_R and σ_z which are for stars in a Solar cylinder defined by $7 < R < 9$ kpc. We over plot our AVR given by Equation (8). We adopt $\beta_z = 0.37$ and $\beta_R = 0.23$, a value that fits both the RAVE and GCS data well when using the Shu model (column 2 of Table 6). It can be seen that for ages less than 7 Gyr, the adopted β values correctly reproduce the profiles seen in simulations. The normalization was adjusted to fit the data in the simulations. There is a slight hint that the σ_R flattens beyond 5 Gyr, but it is also consistent with our power law prescription. While the normalization constant $\sigma_z^{\rm thin}$ is roughly in agreement with what we get, the normalization constant $\sigma_R^{\rm thin}$ is higher by about 10 km s^{-1} in simulations. Since their results are for $7 < R < 9$ kpc, and the density of stars and the velocity dispersion increases inwards, the dispersions in simulations are expected to be slightly higher as compared to dispersions at $R = R_0$. In our model the thin disc extends till 10 Gyr (solid line) and stars older than this belong to the thick disc with a constant age of 11 Gyr (shown by red triangles). It is clear from the figure that our simple model is not adequate for old stars. The discrepancy of our model is more acute for the σ_z profile which shows an abrupt increase for ages greater than 7 Gyr. It should be noted that, the power law model that we use is adequate for handling a mild flattening of velocity dispersion with age, but is not suitable to handle a sharp rise.

6.5. *The thick disc*

First, we discuss our results for the Gaussian model. For the thick disc, our values for $(\sigma_R^{\rm thick}, \sigma_\phi^{\rm thick}, \sigma_z^{\rm thick})$ for GCS

as given by the Gaussian model (column 3 of Table 5) are in good agreement with results of Soubiran et al. (2003)($39 \pm 4, 39 \pm 4, 63 \pm 6$) but differ from Robin et al. (2003) regarding $\sigma_\phi^{\rm thick}$. The RAVE $\sigma_R^{\rm thick}$ is lower than GCS by 7 $\rm km\,s^{-1}$ (column 6 of Table 5) but the other dispersions match up with GCS.

In the Gaussian model the thick disc velocity dispersions are much larger than that of the old thin disc. In the Shu models, we find that the thick disc dispersions are very similar to the old thin disc (column 6 of Table 6). However, $q_{\rm thick}$ is much larger than $q_{\rm thin}$. The reason that the Gaussian and Shu models differ in their estimates for the thick disc velocity dispersions is as follows. In the Shu model, the parameter σ_R^2, which controls the velocity dispersion, is a function of age τ and guiding radius R_g and is not equal to the velocity dispersion $\overline{v_R^2}(\tau, R)$. For a positive q, $\overline{v_R^2}(\tau, R) = \int \sigma_R^2(\tau, R_g) P(R_g|R, \tau) dR_g > \sigma_R^2(\tau, R_g = R)$. At a given radius R in general for warm discs there are a significant number of stars with $R_g < R$. Increasing q not only makes stars at small radius hotter but also makes them more likely to be found at higher radius. So increasing q increases $\overline{v_R^2}(\tau, R)$. For the set of parameters given in column 6 of Table 6, we find that at $R = R_\odot$

$$\sqrt{\langle v_{z,\rm thin}^2\rangle(\tau)} = 26.8 \left(\frac{\tau + 0.1}{10.1\ \rm Gyr}\right)^{0.41} \ \rm km\,s^{-1} \quad (29)$$

$$\sqrt{\langle v_{R,\rm thin}^2\rangle(\tau)} = 41.4 \left(\frac{\tau + 0.1}{10.1\ \rm Gyr}\right)^{0.22} \ \rm km\,s^{-1} \quad (30)$$

$$\sqrt{\langle v_{z,\rm thick}^2\rangle} = 40.0\ \rm km\,s^{-1}, \quad (31)$$

$$\sqrt{\langle v_{R,\rm thick}^2\rangle} = 49.4\ \rm km\,s^{-1}, \quad (32)$$

with $0 < \tau < 10$ Gyr. So the total thick disc $\overline{v_{z,R}^2}$ in the solar neighborhood is still much larger than that of the thin-disc.

In the Shu model the thick disc dispersions at $R_g = R_\odot$ are similar to that of the old thin disc. However, $q_{\rm thick}$ is much larger than $q_{\rm thin}$, i.e., the radial scale length of dispersion profile is smaller for the thick disc. While this can be an argument for a distinct thick disc, a smooth increase of q with age cannot be ruled out at this stage. Additionally, our prior on age and distance distribution assumes a distinct thick disc, e.g., in Figure 3 it can be seen that the distance distribution changes suddenly at 10 Gyr. This could be responsible for $q_{\rm thick}$ being larger than $q_{\rm thin}$. An alternative interpretation is as follows. In our kinematic analysis we use the same $R_d = 2.5$ (radial density scale length) for both thin and thick disc. So, we are effectively measuring the quantity q/R_d. This suggests that the thick disc radial scale length is smaller than that of thin disc. This agrees with the findings of Bovy et al. (2012d) who suggest a decrease of scale length with age.

6.6. *The radial gradient of velocity dispersions*

To date, there has been little discussion about the q parameter in the literature. This 'constant' controls the radial dependence of velocity dispersion which is modelled as $\exp(-qR/R_d)$. This choice of the radial dependence is motivated by the desire to produce discs in which the scale height is independent of radius. For example, under the epicyclic approximation, if σ_z/σ_R is assumed to be constant, then the scale height is independent of radius for $q = 0.5$ (van der Kruit & Searle 1982; van der Kruit 1988; van der Kruit & Freeman 2011). Lewis & Freeman (1989) using 600 old disc K giants spanning 1 to 17 kpc in galactocentric radius estimate q to be 0.29 for radial velocity and 0.38 for azimuthal velocity. Ojha et al. (1996) using a survey of UBVR photometry and proper motions in different directions of the Galaxy estimated $q = 0.225 \pm 0.04$. Bovy et al. (2012c) using SDSS/SEGUE data find $q = 0.35$ for vertical velocity dispersions. Bovy et al. (2012a) using APOGEE data find q to be between -0.07 to 0.01, for the radial and azimuthal motion. In our modelling, the radial gradient is assumed to be same for all the three components. Note, most authors quote $R_\sigma = R_d/q$ instead of q. Since we use $R_d = 2.5$ in our analysis so we use this value to convert R_σ reported by other authors to q.

Our results indicate that for GCS, q is positive for both Gaussian and Shu models. For RAVE, the Gaussian model predicts q to be negative, whereas the Shu model predicts $q \sim 0.23$. We also applied the Gaussian model used by Bovy et al. (2012a) to the RAVE data and find $q_{\rm thin} = -0.07$ (see column 1 of Table 9) similar to their result ($-0.075 < q < 0.01$). Since the Shu model also fits the data better, we think that a negative q peculiar to the Gaussian model is spurious. *The reason the Gaussian model does not fit the data well is because the v_ϕ distribution is not symmetric for warm discs, and the Shu DF correctly handles the asymmetry.* Moreover, the $q_{\rm thin}$ estimate from the Shu model agrees for both GCS and RAVE, lending further support to the idea that the problem is related to the use of the Gaussian model.

A positive q agrees with findings of Lewis & Freeman (1989). It should be noted that in the analysis done by us and Bovy et al. (2012a), the q values are strongly dominated by how we model the asymmetric drift and hence are model dependent. On the other hand, the values reported by Lewis & Freeman (1989) are a direct measure of the radial gradient of velocity dispersion. The q value for the thick disc is in general higher than for the thin disc in RAVE.

6.7. *Comparison with Bovy's kinematic model*

We carried out a more detailed analysis of the kinematic model used by Bovy et al. (2012a). We stress that there are significant differences between the analysis done by us and by Bovy et al. (2012a), both related to data and methodology, which should be kept in mind when comparing the results. Their sample is close to the plane $|b| < 1.5^\circ$ and lies in the range $30^\circ < \ell < 330^\circ$. Being close to the plane they cannot measure the vertical motion, but the advantage is they do not have to worry about the dependence of asymmetric drift with vertical height z. The ℓ and b range being different means that their data and ours probe spatially different regions of the Milky Way. If the disc is axissymmetric, we hope to get similar answers, but not otherwise.

Their main analysis is using a single population model and without any age-velocity dispersion relation. We approximate this by assuming $\beta_R \sim \beta_\phi \sim \beta_z \sim 0$, and the results are shown in column 1 of Table 9. They use a Gaussian model, but with a modified formula for the asymmetric drift. To model their asymmetric drift formula, we assume the parameter X used by them in Equation 5 to be 0.85 (we rename X as $k_{\rm ad}$). As mentioned earlier, using RAVE data and a Gaussian model we find $q_{\rm thin} = -0.07$ in agreement with them, a consequence of the Gaussian model assumption. Our value of $v_{\rm c}$ is also in agreement but our σ_R is much larger than their value 31.4 km $\rm s^{-1}$. Their sample could be dominated by cold stars because of its proximity to the plane. They find

TABLE 9
CONSTRAINTS ON MODEL PARAMETERS WITH BOVY ET AL. (2012A) GAUSSIAN MODEL. SEE TABLE 5 FOR FURTHER DESCRIPTION.

Model	RAVE BOVY	RAVE BOVY	RAVE BOVY	RAVE BOVY
$U_\odot$	$10.16^{+0.15}_{-0.15}$	$11.78^{+0.15}_{-0.15}$	$11.59^{+0.15}_{-0.14}$	$10.96^{+0.14}_{-0.14}$
$V_\odot$	$13.36^{+0.25}_{-0.22}$	$6.2^{+0.18}_{-0.18}$	$8.77^{+0.28}_{-0.28}$	$0.032^{+0.052}_{-0.024}$
$W_\odot$	$7.364^{+0.098}_{-0.098}$	$7.688^{+0.09}_{-0.09}$	$7.694^{+0.097}_{-0.089}$	$7.622^{+0.094}_{-0.087}$
$\sigma_\phi^{\rm thin}$	$26.455^{+0.095}_{-0.096}$	$33.11^{+0.29}_{-0.27}$	$25.83^{+0.36}_{-0.38}$	$33.84^{+0.28}_{-0.27}$
$\sigma_z^{\rm thin}$	$22.99^{+0.12}_{-0.12}$	$31.55^{+0.3}_{-0.3}$	$23.29^{+0.6}_{-0.62}$	$33.6^{+0.29}_{-0.29}$
$\sigma_R^{\rm thin}$	$41.39^{+0.17}_{-0.16}$	$57.58^{+0.31}_{-0.31}$	$41.55^{+0.57}_{-0.62}$	$51.44^{+0.29}_{-0.27}$
$\sigma_\phi^{\rm thick}$			$37.45^{+0.5}_{-0.56}$	
$\sigma_z^{\rm thick}$			$38.84^{+0.47}_{-0.52}$	
$\sigma_R^{\rm thick}$			$65.69^{+0.55}_{-0.64}$	
β_R	0.01	$0.4584^{+0.0071}_{-0.0066}$	$0.193^{+0.013}_{-0.012}$	$0.3568^{+0.0045}_{-0.0049}$
β_z	0.01	$0.514^{+0.016}_{-0.013}$	$0.263^{+0.025}_{-0.027}$	$0.588^{+0.013}_{-0.016}$
β_ϕ	0.01	$0.3747^{+0.0093}_{-0.009}$	$0.166^{+0.013}_{-0.013}$	$0.4151^{+0.0081}_{-0.0093}$
$q_{\rm thin}$	$-0.0725^{+0.0075}_{-0.0088}$	$0.1232^{+0.0064}_{-0.006}$	$0.029^{+0.013}_{-0.013}$	$0.1045^{+0.0069}_{-0.0077}$
$q_{\rm thick}$			$0.1726^{+0.0095}_{-0.0097}$	
v_0	$210.8^{+1.5}_{-1.5}$	$205.5^{+1.5}_{-1.5}$	$213.9^{+1.6}_{-1.6}$	$239.1^{+1.7}_{-1.9}$
$k_{\rm ad}$	0.85	0.85	0.85	$1.968^{+0.022}_{-0.021}$

$\sigma_\phi/\sigma_R = 0.83$ which is higher by about 0.1 than what we get for either RAVE or GCS using any type of model.

They also explored multiple populations with a prior on age as given by an exponentially declining star formation history. However, they only quote $v_{\rm c}, R_0$ and σ_R for it. For multiple populations, their prior on age for the selected stars ignores the fact that scale height increases with age. This will probably have little impact on $v_{\rm c}$, but their σ_R values cannot be compared with ours. Also, they assume a priori that $\beta_R = \beta_\phi = 0.38$, but we have shown that σ_R depends upon the choice of β_R. When we leave β free, we see they differ from the value of 0.38 (column 2 of Table 9). If the thick disc is included, the β values are significantly reduced (column 3). In agreement with Bovy et al. (2012a) we find that the value of $v_{\rm c}$ is not affected much by the choice of age-velocity dispersion relation. Including the thick disc leads to an increase in $v_{\rm c}$ by only 8 km s^{-1}. Interestingly, when $k_{\rm ad}$ is left free we find it favors very high values (column 4). This suggests that we are underestimating the asymmetric drift, most probably due to our neglect of the vertical dependence.

6.8. *Systematics*

Although we get quite precise values for most model parameters, there are additional systematic uncertainties that we have neglected. We performed some additional MCMC runs to investigate these systematics and these results are summarized in Table 7. The first set of systematics is due to two parameters that were kept fixed in our analysis, while the second set is related to our choice of priors on the age and distance distribution of stars.

The distance of the sun from the Galactic center $R_\odot$ and the radial gradient of circular velocity α_R were kept fixed at 8.0 kpc and zero for most of our analysis. This is because these are strongly correlated with v_0. Using just the angular position and radial velocity of RAVE stars, it is not possible to constrain them. The effect of changing R_0 from 7.5 to 8.5 kpc can be seen in column 3 and 4 of Table 7, while the effect of changing α_R from zero to 0.65 km s^{-1} kpc^{-1} can be gauged by comparing columns 1 and 2 in the same table. Using these tables, if needed one can obtain values for any given R_0 and α_R by linearly interpolating between the respective columns. Increasing α_R increases v_0, while the other parameters are relatively unaffected. Increasing $R_\odot$ increases α_z as well as v_0. Again, there is little change in other parameters. The $\Omega_\odot$ was found to decrease from 30.8 km s^{-1} kpc^{-1} at $R_\odot = 7.5$ kpc to 29.4 km s^{-1} kpc^{-1} at $R_\odot = 8.5$ kpc. The above relationship tentatively suggests that at $R_\odot \sim 7.92$ one can match the Sgr A* proper motion. We also checked the effect of setting $\alpha_z = 0.0374$, the value we expect from analytical models. We found that this makes $v_0 \sim 229.2$ km s^{-1} and $V_\odot \sim 8.0$, which is not significantly far from the value we get when α_z is free.

We now discuss systematics related to our choice of priors. Our main prior is that the age and distance distribution of stars along a particular line of sight is in accordance with Besançon model of the Galaxy. Additionally, the distance distribution for a given $I_{\rm DENIS}$ magnitude of a star depends upon the isochrones that are used in the model. As a crude way to gauge the sensitivity to our priors in age, we run a model with $\beta_z = \beta_R = 0.01$ (column 7) which makes the kinematics of the thin disc independent of age. As expected, the thin and thick disc parameters change. Other than this, the α_z and v_0 are found to increase by 12% and 2% respectively.

Next, we test the effect of changing the distance prior. For this we alternately increase and decrease our prior distance distribution by multiplying the distances by a factor of 1.1 and 0.9 respectively. The $V_\odot$, v_0 and $q_{\rm thin}$ show significant changes. It should be noted that this is only an approximate way to check the sensitivity of our results on the priors. In reality, if magnitudes are systematically wrong then the spatial density model that we use will not match the number count of stars obtained from photometric surveys. So, the mass density laws of the model will have to be modified as well. The proper way to do this is to do a dynamical modelling in which the kinematics and the spatial distribution of stars are fitted jointly to the observational data.

Finally, systematics due to inaccuracy of the model to self-consistently describe the system remains a worry. Flattened axisymmetric galactic potential admit three integral of motions, so ideally one should construct distribution functions that are functions of these three integrals of motion. If the potential is separable in R and z such that $\Phi(R,z) = \Phi(R) + \Phi(z)$, then the vertical and planar motion are independent of each other. The Shu DF is a planar

distribution function that takes two integrals of motion as E_R and L_z. The vertical energy in such cases can be used as the third integral of motion. In reality the $\partial\Phi/\partial R$ has a z dependence and in general the potential is not separable in R and z as discussed above. In such cases a better approach would be to use distribution functions that are functions of action integrals (Binney 2010, 2012b). However, converting phase space coordinates to $(\mathbf{x}, \mathbf{v})$ is not easy and methods to do this are being developed (Binney 2012a).

Another problem with our current model is that the scale height only depends upon age. A mono age population in reality is a superposition of multiple isothermal populations, e.g., isothermal populations characterized by their guiding radius R_g. Populations with smaller R_g are in general hotter and will have larger scale height at a given radius R. This means that even for a mono age population the asymmetric drift has a z dependence which was ignored in our present analysis. Allowing the circular velocity to fall with z was an attempt to also take this into account. Use of an action based model will also rectify this problem. An approximate formula to take this into account has been given by Schönrich & Binney (2012). Their treatment in general extends the Shu DF by making use of the adiabatic approximation, i.e., conservation of vertical action, and allows modelling the kinematics as a function of distance from the plane. Recently, it has been shown by Binney (2012a); Binney & McMillan (2011) that adiabatic approximation is accurate only close to the midplane and that much better results are obtained by the use of the Stackel approximation. In future, we need to compare our findings with action based models.

7. SUMMARY AND CONCLUSIONS

In this paper, we have constrained the kinematic parameters of the Milky Way disc using stars from the RAVE and the GCS surveys. To constrain kinematic parameters, we use analytic models based on the Gaussian and Shu distribution functions. We investigated the Solar peculiar motion ($U_\odot, V_\odot, W_\odot$), the circular velocity at the Sun v_0 and its vertical profile α_z, the age velocity dispersion relations ($\beta_{R,\phi,z}$, $\sigma^{\rm thin}_{R,\phi,z}$ and $\sigma^{\rm thick}_{R,\phi,z}$), and the radial gradient of dispersions ($q_{\rm thin}$ and $q_{\rm thick}$). We explored the full posterior distribution of parameters using the Markov Chain Monte Carlo technique. The main assumption we make is that we assume a SFR, IMF and density laws that describe the spatial distribution of stars in accordance with the Besançon model of Robin et al. (2003), but with slight modifications as described in Sharma et al. (2011). Isochrones from the Padova database are used to compute photometric magnitude of the model stars (Marigo et al. 2008; Bertelli et al. 1994). This model provides a good fit to the photometric star counts of Milky Way. For GCS data, which has full phase space information for the stars, we compute the likelihood in $(\mathbf{x}, \mathbf{v})$ phase space. For RAVE data, we do not have accurate distances and proper motions can have systematics, hence we measure the likelihood only in $(\ell, b, v_{\rm los})$ space. In other words, we limit our analysis to quantities that are accurately and unambiguously known. Our best fit parameters for both Gaussian and Shu models using the RAVE data are summarized in column 6 of Table 6.

One could in principle combine both the surveys and give joint constraints. However, for some parameters like solar motion it is not obvious if two surveys that probe different volumes of the Galaxy should give the same answer, e.g., if the local standard of rest is moving or there are kinematic substructures like in GCS sample. In such cases one cannot do a joint constraint. In general, if there are systematic differences between the two surveys, then the interpretation of the results coming from joint constraints becomes difficult. Hence, in this paper, to begin with we analyze the surveys separately and make an effort to understand the systematics. Having done that, in an attempt to improve the accuracy we follow the approach of judiciously fixing some parameters in one survey to that from the other, where they are found to be better constrained. We do this only for those parameters which we expect to be same in both surveys.

We find the Gaussian model to be unsuitable for estimating disc parameters like $q_{\rm thin}$ and $V_\odot$ as they are strongly degenerate. The Gaussian model gives different $q_{\rm thin}$ for RAVE and GCS. For RAVE it predicts a negative $q_{\rm thin}$, i.e., a positive radial gradient for σ_R. This is not expected if the disc height and the ratio σ_z/σ_R do not vary with radius. A negative $q_{\rm thin}$ also disagrees with the findings of Lewis & Freeman (1989). The Shu model has three fewer parameters than the Gaussian model and this helps it to break the degeneracy between $q_{\rm thin}$ and $V_\odot$. It gives positive and consistent values for $q_{\rm thin}$ for both RAVE and GCS. The Shu models also fits the RAVE data better than the Gaussian model, especially with regard to the azimuthal component.

The RAVE data allows us to constrain the Solar peculiar motion and the local circular velocity quite precisely. Our $U_\odot$ and $W_\odot$ are in good agreement with the results of Schönrich et al. (2010), but our $V_\odot$ is lower by 5 km s^{-1}. The RAVE $U_\odot$ and $W_\odot$ are within 2-σ range of GCS values, but $V_\odot$ is lower by 2 km s^{-1}. Using $R_0 = 8.0$ kpc and assuming $\partial v_{\rm circ}/\partial R = 0$ we get $v_0 = 232 \pm 2$ km s^{-1}. Combining the estimate of v_0 and $V_\odot$, we find the Solar angular velocity with respect to the Galactic Center to be in good agreement with the measured Sgr A* proper motion. We find that for any extended sample in the vertical direction, a proper treatment of vertical dependence of the effective circular velocity is needed to measure v_0. The v_0 and $V_\odot$ are sensitive to the priors on age and distance distribution of stars. So systematic errors of the order of the uncertainty in the priors is expected. **Also, systematics due to inaccurate handling of the third integral of motion need to be investigated with better models, e.g., models based on action integrals.**

When using the Shu model, except for $V_\odot$ and thick disc parameters, all parameters show similar values for RAVE and GCS. Since there are very few thick disc stars in GCS, we deem the RAVE thick disc parameters to be more reliable. Also, the uncertainty on $q_{\rm thin}$ and $q_{\rm thick}$ is substantially less for RAVE than for GCS. The only parameter that is constrained better by GCS than RAVE is β_z and this is partly due to the fact that we only use radial velocities in RAVE. In an attempt to build a concordance model, and to enable better comparison between the two data sets, we fix β_z in RAVE to GCS values and then fix $q_{\rm thin}$ and $q_{\rm thick}$ in GCS to RAVE values. Doing so we find that RAVE results are within 3σ of GCS results. The most significant difference between the two is the value of $V_\odot$ which is lower for RAVE by about 2 km s^{-1}. The presence of prominent kinematic substructures in GCS could be responsible for this.

We find that the age-velocity dispersion relations in general obey $\beta_R < \beta_\phi < \beta_z$. Unlike Aumer & Binney (2009), β_ϕ is closer to β_R than β_z. This suggests that the heating mechanisms for the radial and the azimuthal velocities as expected

from physics are similar but different from that of vertical motion. The β values vary depending if the thick disc is included or not. If the thick disc is not included, then the β values are higher compared to when it is included. The axial ratio σ_z/σ_R of the thin disc velocity ellipsoid for the 10 Gyr population is consistent with those predicted for scattering from clouds by Sellwood (2008). The values of β_R and β_z that we find agree well with age-velocity profiles seen in simulations of disc galaxies by Minchev et al. (2013) at least till 7 Gyr. For stars older than 7 Gyr, we think the current model used by us, a power law AVR for thin disc and a constant age thick disc, is too restrictive to model the profile found in simulations. In future, we think analytical models motivated by properties of discs seen in simulations should be useful for constraining the kinematic parameters of the Milky Way. Also it would be interesting to explore dynamical models that are self-consistent rather than pure kinematic models as studied here. The question of how to model the thick disc is also something that needs further attention.

The thick disc velocity dispersions for $R_g = R_\odot$ in the Shu model were found to be very similar to that of the old thin disc. However, the q parameter controlling the radial dispersion profile for the thick disc was found to be much larger than that of the thin disc. This could mean that for thick disc, the radial scale length of the velocity dispersion profile is smaller or the radial scale length of the density profile is smaller, or both. This agrees with the findings of Bovy et al. (2012d) who suggest a decrease of radial density scale length with age. In this regard, the role of our adopted priors on age and distance distribution of stars needs to be investigated further.

Finally, in most models, the properties of the disc are assumed to be a function of age. However, age is a quantity that is most difficult to measure. Stellar astroseismology with missions like CoRoT and KEPLER provides an opportunity to measure ages with more accuracy than before (Chaplin et al. 2010, 2011; Appourchaux et al. 2008), but there is still a long way to go before this can be done for a large number of stars. Alternatively, the chemical abundance, e.g., alpha abundance, provides a fair proxy for age at a given metallicity. Hence, studying the relationship of kinematic properties with abundance would be crucial. It was shown by Bovy et al. (2012d) that decomposing a disc based on geometry versus abundance can give different results. In the same way, age and abundance decomposition can also give different results. In future, more effort will be required to determine as to how accurate are alpha abundances in tracking the age and how to incorporate them in modelling.

ACKNOWLEDGMENTS

SS is funded through ARC DP grant 120104562 (PI Bland-Hawthorn) which supports the HERMES project. JBH is funded through a Federation Fellowship from the Australian Research Council (ARC).

Funding for RAVE has been provided by: the Australian Astronomical Observatory; the Leibniz-Institut fuer Astrophysik Potsdam (AIP); the Australian National University; the Australian Research Council; the French National Research Agency; the German Research Foundation (SPP 1177 and SFB 881); the European Research Council (ERC-StG 240271 Galactica); the Istituto Nazionale di Astrofisica at Padova; The Johns Hopkins University; the National Science Foundation of the USA (AST-0908326); the W. M. Keck foundation; the Macquarie University; the Netherlands Research School for Astronomy; the Natural Sciences and Engineering Research Council of Canada; the Slovenian Research Agency; the Swiss National Science Foundation; the Science & Technology Facilities Council of the UK; Opticon; Strasbourg Observatory; and the Universities of Groningen, Heidelberg and Sydney. The RAVE web site is at http://www.rave-survey.org.

APPENDIX

REFERENCES

Abadi, M. G., Navarro, J. F., Steinmetz, M., & Eke, V. R. 2003, ApJ, 591, 499, arXiv:astro-ph/0211331
Abazajian, K. N. et al. 2009, ApJS, 182, 543, 0812.0649
Andrieu, C., & Thoms, J. 2008, Statistics and Computing, 18, 343
Appourchaux, T. et al. 2008, A&A, 488, 705
Arce, H. G., & Goodman, A. A. 1999, ApJ, 512, L135, arXiv:astro-ph/9902109
Aumer, M., & Binney, J. J. 2009, MNRAS, 397, 1286, 0905.2512
Aumer, M., & White, S. D. M. 2013, MNRAS, 428, 1055, 1203.1190
Bahcall, J. N., & Soneira, R. M. 1980a, ApJ, 238, L17
——. 1980b, ApJS, 44, 73
——. 1984, ApJS, 55, 67
Bensby, T., Feltzing, S., & Lundström, I. 2003, A&A, 410, 527
Bensby, T., Feltzing, S., Lundström, I., & Ilyin, I. 2005, A&A, 433, 185, arXiv:astro-ph/0412132
Bertelli, G., Bressan, A., Chiosi, C., Fagotto, F., & Nasi, E. 1994, A&AS, 106, 275
Bienayme, O., Robin, A. C., & Creze, M. 1987, A&A, 180, 94
Binney, J. 2010, MNRAS, 401, 2318, 0910.1512
——. 2012a, MNRAS, 426, 1324, 1207.4910
——. 2012b, MNRAS, 426, 1328, 1207.4917
Binney, J., Dehnen, W., & Bertelli, G. 2000, MNRAS, 318, 658, arXiv:astro-ph/0003479
Binney, J., & McMillan, P. 2011, MNRAS, 413, 1889, 1101.0747
Binney, J., & Tremaine, S. 2008, Galactic Dynamics: Second Edition (Princeton University Press)
Boeche, C. et al. 2013, A&A, 553, A19, 1302.3466
Bond, N. A. et al. 2010, ApJ, 716, 1, 0909.0013
Bournaud, F., Elmegreen, B. G., & Martig, M. 2009, ApJ, 707, L1, 0910.3677
Bovy, J. et al. 2012a, ApJ, 759, 131, 1209.0759
Bovy, J., Rix, H.-W., & Hogg, D. W. 2012b, ApJ, 751, 131, 1111.6585
Bovy, J., Rix, H.-W., Hogg, D. W., Beers, T. C., Lee, Y. S., & Zhang, L. 2012c, ApJ, 755, 115, 1202.2819
Bovy, J., Rix, H.-W., Liu, C., Hogg, D. W., Beers, T. C., & Lee, Y. S. 2012d, ApJ, 753, 148, 1111.1724
Brook, C. B., Kawata, D., Gibson, B. K., & Freeman, K. C. 2004, ApJ, 612, 894, arXiv:astro-ph/0405306
Cambrésy, L., Jarrett, T. H., & Beichman, C. A. 2005, A&A, 435, 131, arXiv:astro-ph/0501444
Carlberg, R. G., & Sellwood, J. A. 1985, ApJ, 292, 79
Casagrande, L., Schönrich, R., Asplund, M., Cassisi, S., Ramírez, I., Meléndez, J., Bensby, T., & Feltzing, S. 2011, A&A, 530, A138, 1103.4651
Chaplin, W. J. et al. 2010, ApJ, 713, L169, 1001.0506
——. 2011, Science, 332, 213, 1109.4723
Chiappini, C., Matteucci, F., & Gratton, R. 1997, ApJ, 477, 765, astro-ph/9609199
Chiba, M., & Beers, T. C. 2000, AJ, 119, 2843, arXiv:astro-ph/0003087
Dehnen, W., & Binney, J. 1998a, MNRAS, 294, 429, arXiv:astro-ph/9612059
Dehnen, W., & Binney, J. J. 1998b, MNRAS, 298, 387, arXiv:astro-ph/9710077
Delhaye, J. 1965, in Galactic Structure, ed. A. Blaauw & M. Schmidt, 61
Di Matteo, P., Lehnert, M. D., Qu, Y., & van Driel, W. 2011, A&A, 525, L3, 1011.3825

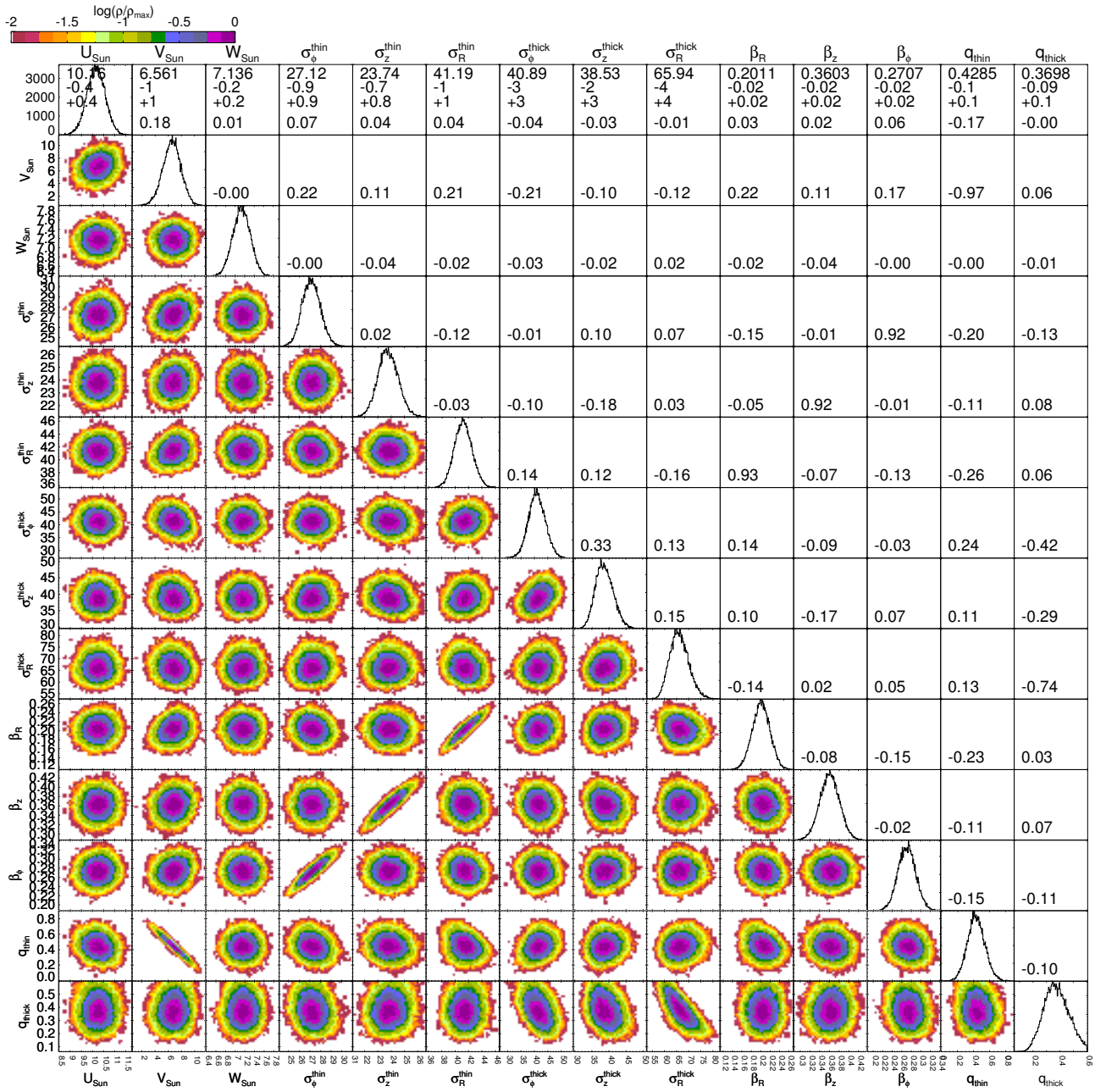

FIG. 10.— Marginalized posterior distribution of model parameters. The numbers are the linear Pearson correlation coefficient. Shown is the case of Gaussian model for GCS data (column 1 of Table 5). Strong dependency can be seen between β and $\sigma^{\rm thin}$ values. Additionally, $(q_{\rm thin}, V_\odot)$, $(q_{\rm thin}, \sigma_R^{\rm thin})$, $(q_{\rm thin}, \beta_R)$ and $(q_{\rm thick}, \sigma_R^{\rm thick})$ also show dependency.

Edvardsson, B., Andersen, J., Gustafsson, B., Lambert, D. L., Nissen, P. E., & Tomkin, J. 1993, A&A, 275, 101

Epchtein, N. et al. 1999, A&A, 349, 236

Forbes, J., Krumholz, M., & Burkert, A. 2012, ApJ, 754, 48, 1112.1410

Freeman, K., & Bland-Hawthorn, J. 2008, in Astronomical Society of the Pacific Conference Series, Vol. 399, Astronomical Society of the Pacific Conference Series, ed. T. Kodama, T. Yamada, & K. Aoki, 439–+

Fuhrmann, K. 1998, A&A, 338, 161

Gelfand, A. E., & Smith, A. F. 1990, Journal of the American statistical association, 85, 398

Gelman, A., Roberts, G., & Gilks, W. 1996, Bayesian statistics, 5, 599

Gillessen, S., Eisenhauer, F., Trippe, S., Alexander, T., Genzel, R., Martins, F., & Ott, T. 2009, ApJ, 692, 1075, 0810.4674

Gilmore, G., & Reid, N. 1983, MNRAS, 202, 1025

Girard, T. M. et al. 2011, AJ, 142, 15, 1104.5708

Girardi, L., Groenewegen, M. A. T., Hatziminaoglou, E., & da Costa, L. 2005, A&A, 436, 895, arXiv:astro-ph/0504047

Golubov, O. et al. 2013, A&A, 557, A92, 1307.6073

Gomez, A. E., Grenier, S., Udry, S., Haywood, M., Meillon, L., Sabas, V., Sellier, A., & Morin, D. 1997, in ESA Special Publication, Vol. 402, Hipparcos - Venice '97, ed. R. M. Bonnet, E. Høg, P. L. Bernacca, L. Emiliani, A. Blaauw, C. Turon, J. Kovalevsky, L. Lindegren, H. Hassan, M. Bouffard, B. Strim, D. Heger, M. A. C. Perryman, & L. Woltjer, 621–624

Górski, K. M., Hivon, E., Banday, A. J., Wandelt, B. D., Hansen, F. K., Reinecke, M., & Bartelmann, M. 2005, ApJ, 622, 759, arXiv:astro-ph/0409513

Hänninen, J., & Flynn, C. 2002, MNRAS, 337, 731, arXiv:astro-ph/0208426

Haywood, M., Robin, A. C., & Creze, M. 1997a, A&A, 320, 428

——. 1997b, A&A, 320, 440

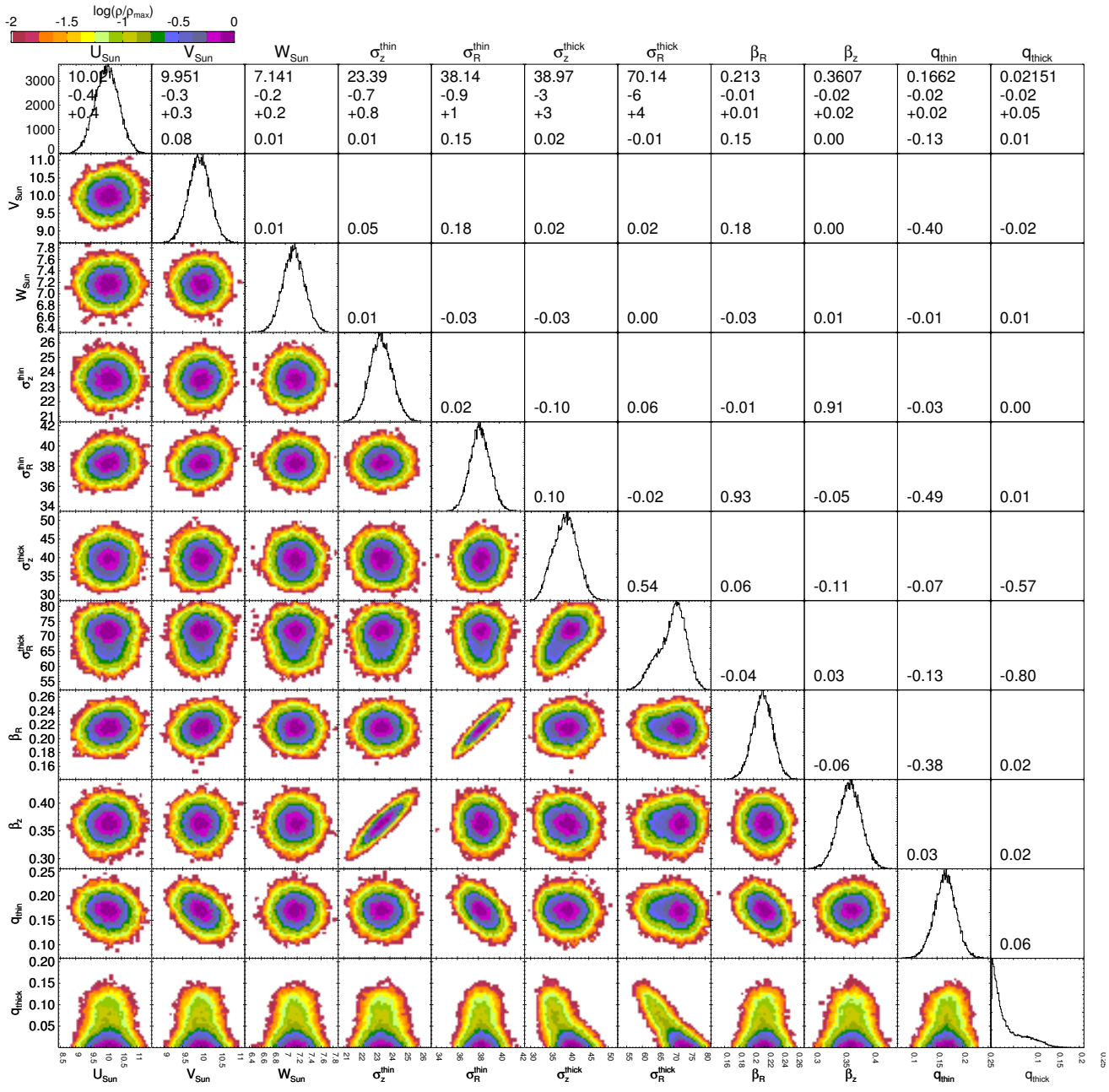

FIG. 11.— Marginalized posterior distribution of model parameters. The numbers are the linear Pearson correlation coefficient. Shown is the case of Shu model for GCS data (column 1 of Table 6). Same dependencies as in Figure 10 can be seen. Dependency of $(q_{\rm thin}, V_\odot)$ has got weaker while that of $(q_{\rm thick}, \sigma_R^{\rm thick})$ has become stronger.

Høg, E. et al. 2000, A&A, 355, L27
Holmberg, J., Nordström, B., & Andersen, J. 2007, A&A, 475, 519, 0707.1891
——. 2009, A&A, 501, 941, 0811.3982
House, E. L. et al. 2011, MNRAS, 415, 2652, 1104.2037
Ida, S., Kokubo, E., & Makino, J. 1993, MNRAS, 263, 875
Ivezić, Ž. et al. 2008, ApJ, 684, 287, 0804.3850
Jenkins, A. 1992, MNRAS, 257, 620
Jurić, M. et al. 2008, ApJ, 673, 864, arXiv:astro-ph/0510520
Just, A., & Jahreiß, H. 2010, MNRAS, 402, 461, 0910.3481
Kazantzidis, S., Bullock, J. S., Zentner, A. R., Kravtsov, A. V., & Moustakas, L. A. 2008, ApJ, 688, 254, 0708.1949
Kazantzidis, S., Zentner, A. R., Kravtsov, A. V., Bullock, J. S., & Debattista, V. P. 2009, ApJ, 700, 1896, 0902.1983
Kordopatis, G. et al. 2013, AJ, 146, 134, 1309.4284
Law, D. R., & Majewski, S. R. 2010, ApJ, 714, 229, 1003.1132
Lewis, J. R., & Freeman, K. C. 1989, AJ, 97, 139
Loebman, S. R., Roškar, R., Debattista, V. P., Ivezić, Ž., Quinn, T. R., & Wadsley, J. 2011, ApJ, 737, 8, 1009.5997
Marigo, P., Girardi, L., Bressan, A., Groenewegen, M. A. T., Silva, L., & Granato, G. L. 2008, A&A, 482, 883, 0711.4922
McMillan, P. J., & Binney, J. 2012, MNRAS, 419, 2251, 1108.1749
McMillan, P. J., & Binney, J. J. 2008, MNRAS, 390, 429, 0806.0319
——. 2010, MNRAS, 402, 934, 0907.4685
Minchev, I., Chiappini, C., & Martig, M. 2013, A&A, 558, A9, 1208.1506
Minchev, I., & Famaey, B. 2010, ApJ, 722, 112, 0911.1794
Minchev, I., & Quillen, A. C. 2006, MNRAS, 368, 623, arXiv:astro-ph/0511037
Nordström, B. et al. 2004, A&A, 418, 989, arXiv:astro-ph/0405198
Ojha, D. K., Bienayme, O., Robin, A. C., Creze, M., & Mohan, V. 1996, A&A, 311, 456, arXiv:astro-ph/9511049
Pasetto, S. et al. 2012a, A&A, 547, A70, 1209.0456

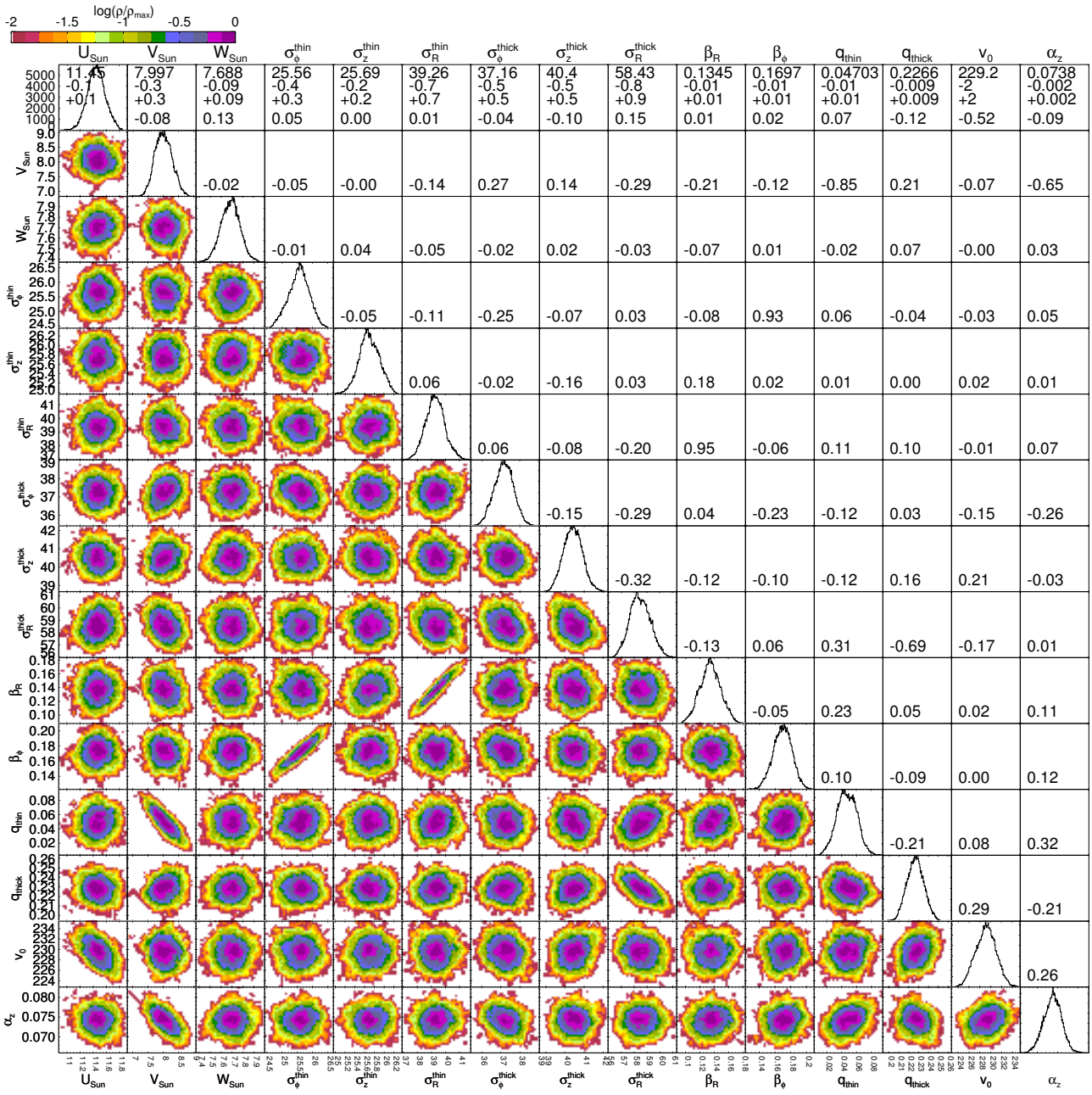

FIG. 12.— Marginalized posterior distribution of model parameters. The numbers are the linear Pearson correlation coefficient. Shown is the case of Gaussian model for RAVE data (column 5 of Table 5). Strong dependency can be seen between β and $\sigma^{\rm thin}$ values. Additionally, $(q_{\rm thin}, V_\odot)$ and $(q_{\rm thick}, \sigma_R^{\rm thick})$ also show dependency. Finally, the v_c is anti-correlated to $U_\odot$ and α_z to $V_\odot$.

——. 2012b, A&A, 547, A71, 1209.0460
Perryman, M. A. C. et al. 1997, A&A, 323, L49
Quillen, A. C., & Garnett, D. R. 2001, in Astronomical Society of the Pacific Conference Series, Vol. 230, Galaxy Disks and Disk Galaxies, ed. J. G. Funes & E. M. Corsini, 87–88
Quinn, P. J., Hernquist, L., & Fullagar, D. P. 1993, ApJ, 403, 74
Reid, M. J. 1993, ARA&A, 31, 345
Reid, M. J., & Brunthaler, A. 2004, ApJ, 616, 872, arXiv:astro-ph/0408107
Reid, N., & Majewski, S. R. 1993, ApJ, 409, 635
Robin, A., & Creze, M. 1986, A&A, 157, 71
Robin, A. C., Reylé, C., Derrière, S., & Picaud, S. 2003, A&A, 409, 523
Röser, S., Schilbach, E., Schwan, H., Kharchenko, N. V., Piskunov, A. E., & Scholz, R.-D. 2008, A&A, 488, 401, 0806.1009
Roškar, R., Debattista, V. P., Brooks, A. M., Quinn, T. R., Brook, C. B., Governato, F., Dalcanton, J. J., & Wadsley, J. 2010, MNRAS, 408, 783, 1006.1659
Sale, S. E. 2012, MNRAS, 427, 2119, 1208.4946
Sales, L. V., Navarro, J. F., Theuns, T., Schaye, J., White, S. D. M., Frenk, C. S., Crain, R. A., & Dalla Vecchia, C. 2012, MNRAS, 423, 1544, 1112.2220
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525, arXiv:astro-ph/9710327
Schönrich, R. 2012, MNRAS, 427, 274, 1207.3079
Schönrich, R., & Binney, J. 2009a, MNRAS, 396, 203, 0809.3006
——. 2009b, MNRAS, 399, 1145, 0907.1899
——. 2012, MNRAS, 419, 1546, 1109.4417
Schönrich, R., Binney, J., & Dehnen, W. 2010, MNRAS, 403, 1829, 0912.3693
Seabroke, G. 2008, The Observatory, 128, 520
Seabroke, G. M., & Gilmore, G. 2007, MNRAS, 380, 1348, 0707.1027
Sellwood, J. A. 2008, ASPCS, 396, 241, 0803.1574
——. 2013, ArXiv e-prints, 1303.4919

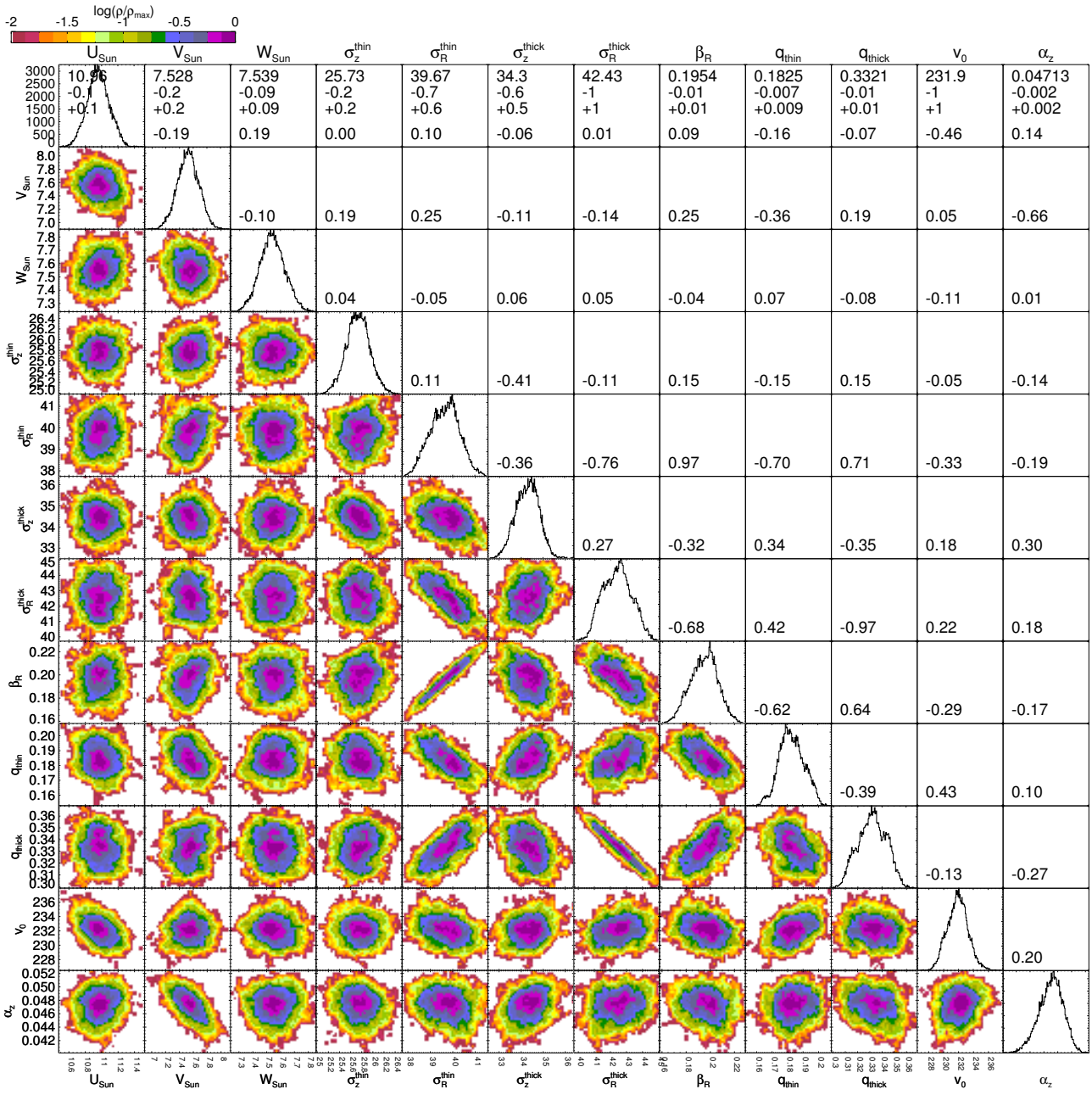

FIG. 13.— Marginalized posterior distribution of model parameters. The numbers are the linear Pearson correlation coefficient. Shown is the case of Shu model for RAVE data (column 6 of Table 6). Strong dependency can be seen between β and $\sigma^{\rm thin}$ values. Additionally, $(q_{\rm thin}, V_\odot)$, $(q_{\rm thin}, \sigma_R^{\rm thin})$, $(q_{\rm thin}, \beta_R)$ and $(q_{\rm thick}, \sigma_R^{\rm thick})$ also show dependency. Unlike GCS a dependency of $(\sigma_R^{\rm thin}, \sigma_R^{\rm thick})$ and (β_z, β_R) can be seen. Finally, the $v_{\rm c}$ is anti-correlated to $U_\odot$ and α_z to $V_\odot$.

Sharma, S., & Bland-Hawthorn, J. 2013, ApJ, 773, 183, 1306.6636

Sharma, S., Bland-Hawthorn, J., Johnston, K. V., & Binney, J. 2011, ApJ, 730, 3, 1101.3561

Sharma, S., Steinmetz, M., & Bland-Hawthorn, J. 2012, ApJ, 750, 107, 1203.0315

Shiidsuka, K., & Ida, S. 1999, MNRAS, 307, 737, arXiv:astro-ph/9907243

Shu, F. H. 1969, ApJ, 158, 505

Siebert, A. et al. 2008, MNRAS, 391, 793, 0809.0615

——. 2011, AJ, 141, 187, 1104.3576

Skrutskie, M. F. et al. 2006, AJ, 131, 1163

Soubiran, C., Bienaymé, O., & Siebert, A. 2003, A&A, 398, 141, arXiv:astro-ph/0210628

Spitzer, Jr., L., & Schwarzschild, M. 1953, ApJ, 118, 106

Steinmetz, M. et al. 2006, AJ, 132, 1645, arXiv:astro-ph/0606211

Tanner, M. A., & Wong, W. H. 1987, Journal of the American statistical Association, 82, 528

Tierney, L. 1994, the Annals of Statistics, 1701

van der Kruit, P. C. 1988, A&A, 192, 117

van der Kruit, P. C., & Freeman, K. C. 2011, ARA&A, 49, 301, 1101.1771

van der Kruit, P. C., & Searle, L. 1982, A&A, 110, 61

Villalobos, Á., & Helmi, A. 2008, MNRAS, 391, 1806

Williams, M. E. K. et al. 2013, MNRAS, 436, 101, 1302.2468

Yanny, B. et al. 2009, AJ, 137, 4377, 0902.1781

Yoachim, P., & Dalcanton, J. J. 2006, AJ, 131, 226, arXiv:astro-ph/0508460

Zacharias, N. et al. 2010, AJ, 139, 2184, 1003.2136

Zacharias, N., Urban, S. E., Zacharias, M. I., Wycoff, G. L., Hall, D. M., Monet, D. G., & Rafferty, T. J. 2004, AJ, 127, 3043, arXiv:astro-ph/0403060

Zwitter, T. et al. 2008, AJ, 136, 421, 0806.0546