abstract
A new class of models of stellar discs is introduced and used to build a self-consistent model of our Galaxy. The model is defined by the parameters that specify the action-based distribution functions (DFs) \( f(J) \) of four stellar discs (three thin-disc age cohorts and a thick disc), spheroidal bulge and spheroidal stellar and dark haloes. From these DFs plus a specified distribution of gas, we solve for the densities of stars and dark matter and the potential they generate. The principal observational constraints are the kinematics of stars with Gaia RVS data and the density of stars in the column above the Sun. The model predicts the density and kinematics of stars and dark matter throughout the Galaxy. We determine the structure of the dark halo prior to the infall of baryons. A simple extension of the DFs of stellar components to include chemistry allows the model to reproduce the way the Galaxy’s chemistry is observed to vary in the \((R, z)\) plane. Surprisingly, the data indicate that high-\(\alpha\) stars are confined to orbits with \(J_z \gtrsim 50\) kpc km s\(^{-1}\). The code used to create the model is available on GITHUB.

key words: Galaxy: kinematics and dynamics – galaxies: kinematics and dynamics – methods: numerical

1 introduction
The volume and quality of the observational data that are available for our Galaxy have increased spectacularly over the last decade. The spectra of millions of stars have been taken from the ground, while ESA’s satellite Gaia has been tracking the motion across the sky of over a billion stars, deriving for them photometry of unprecedented precision, and measuring the line-of-sight velocities of millions of the brighter stars. The astrophysical community now faces the challenge of synthesising these data into a coherent physical picture of how our archetypal Galaxy is structured, how it functions as a machine, and how it arrived at this state.

While ideally one would exploit the data for all the \(> 10^6\) stars that Gaia has been tracking, many studies have focused on the \(\sim 7 \times 10^5\) stars for which Gaia has measured line-of-sight velocities (Helmi et al. 2018; Antoja et al. 2018; Hunt et al. 2019; Sellwood et al. 2019; Trick et al. 2019; Gaia Collaboration et al. 2021). We shall see that this sub-sample of the Gaia data is large enough for Poisson noise to be insignificant, and it is relatively easy to model because when the line-of-sight velocities are missing, one has to marginalise over them, which is computationally expensive (e.g. Li & Binney 2022a). The potential strengths of the much larger sample of all Gaia stars are (i) that it extends to much fainter magnitudes, and thus covers a more significant slice of our Galaxy, and (ii) that it should be possible to determine its selection function. Sadly, however, at the current time it is hard to determine the probability that the data for a star of given magnitude, and sky-coordinates will pass a given quality threshold (Boubert & Everall 2020; Everall et al. 2021), and impossible to predict the apparent magnitudes of distant stars of given absolute magnitudes on account of the poorly known distribution of dust in the Galaxy (e.g. Li & Binney 2022b). These problems can be to a large extent side-stepped by modelling the distribution of velocities at given locations because the probability of a star entering Gaia’s radial-velocity sample (RVS) is independent of its velocity. In this paper we fit to Gaia RVS data self-consistent, axisymmetric models of our Galaxy that are defined by distribution functions (DFs) that are analytic functions of the action integrals \(J_i\).

Most previous fits of kinematic data have adopted a simple functional form for the density of dark matter (DM) (Robin et al. 2003; Binney 2010; Bovy & Rix 2013; Piffl et al. 2014). This procedure is unsatisfactory both because DM will actually respond in a non-trivial way to the gravitational field of the stars, which is strongly flattened, and because it yields no information regarding the velocity distribution of local DM, which is important from the perspective of experiments to detect DM on Earth. Here our approach is basically that pioneered by Piffl et al. (2015) in which a distribution function (DF) \(f(J)\) that depends on the action integrals \(J_i\) is assigned to both DM and various stellar pop-
ulations, and then the gravitational field that stars and DM jointly generate is determined iteratively.

Whereas Piffl et al. (2015) simply computed one not very satisfactory model, Binney & Piffl (2015) and Cole & Binney (2017) searched the space of candidate DFs for ones that were consistent with data. The data they employed were rather heterogeneous: astrometry of some stellar masers (Reid & Brunthaler 2004), terminal velocities of neutral hydrogen and carbon monoxide (Malhotra 1995), star counts from the Sloan Digital Sky Survey (Abazajian et al. 2003) and the kinematics of giant stars measured by the RAdial Velocity Experiment (RAVE; Steinmetz & et al. 2006). The heterogeneity of the data combined with sub-optimal numerical methods made the search for an acceptable model computationally expensive.

Here we use the AGAMA software library (Vasiliev 2019) to show that the data for stars with Gaia line-of-sight velocities is so extensive that they strongly constrain DFs for stars and DM on their own. We do this by constructing fully self-consistent dynamical models of the Galaxy. These models are, for the first time, completely specified only by their DF: all other quantities – their density distributions, gravitational potential and kinematics – follow from the DF. In Section 2 we describe a new family of DFs for stellar discs, and in Section 3 we describe the DFs we use to represent the bulge and the dark and stellar haloes. In Section 4 we compare one of our models with Gaia kinematics and a variety of older data sources. This comparison confirms that the Galaxy’s circular speed declines outwards at the solar radius and yields values for the local densities of stars and dark matter. We infer the structure of the dark halo before the baryons fell in under both the assumption of adiabatic invariance and the assumption that the dark halo originally had a central density cusp that was eliminated by baryons upscattering dark-matter particles. In Section 4.6 we show that reasonable hypotheses regarding the chemical compositions of the model’s stellar components provides a good fit to the distribution of stars in the \([\alpha/Fe], [Fe/H]\) plane at various locations \((R, z)\) that were reported by Hayden et al. (2015).

## 2 A NEW FAMILY OF DFS FOR DISCS

Binney (2012), Bovy & Rix (2013), Piffl et al. (2014) and several others modelled the Galaxy’s discs with the quasi-isothermal DF that was introduced by Binney (2010) and modified to its current form by Binney & McMillan (2011). This DF was introduced in the context of solar-neighbourhood kinematics using an analytic model of the Galactic potential \(\Phi\). Weaknesses in this DF emerged when Piffl et al. (2015), Binney & Piffl (2015) and Cole & Binney (2017) used it to model the computing-so-consistently generated potential. They employed ad-hoc work-arounds for these problems, but it is now time to address these problems clinically and resolve them in a satisfying way. Our discussion extends that in Section 4.4 of Vasiliev (2019).

The root problem with the quasi-isothermal DF is that it references the circular radius \(R_c(J_0)\) and the radial and vertical epicycle frequencies \(\kappa(J_0)\) and \(\nu(J_0)\). The way these quantities vary with \(J_0\) depends on the potential, so when they appear in the DF, the latter is no longer a function of the actions alone. This dependence of the DF on \(\Phi\) endangers the convergence of the algorithm Binney (2014) introduced for finding the potential. The problem can be evaded by using the functions \(R_c(J_0)\), etc., associated with a fixed but suitable potential instead of the real potential, but this fix is inelegant and means that a model is not uniquely specified by a set of DFs.

Another problem with the quasi-isothermal DF is that it ceases to make sense physically for orbits that are highly eccentric or strongly inclined to the plane. Therefore we now write down an extension of the Exponential DF defined by Vasiliev (2019) that does not reference \(R_c\), \(\kappa\) or \(\nu\) and is physically reasonable throughout action space.

Let \(J_{\phi}, J_r\) and \(J_z\) be three fixed actions, the first much larger than the other two. Then we define

\[
J_\phi \equiv |J_\phi| + J_\phi^0 \quad \text{and} \quad J_d \equiv |J_\phi| + J_d^0
\]

and take the DF of a disc component to be

\[
f(J) = f_\phi(J_\phi)f_r(J_r,J_d)f_z(J_r,J_z),
\]

where the functions \(f_\phi, f_r, f_z\) have the common form

\[
f_i(J_\phi,J_i) = \left(\frac{J_i}{J_i^0}\right)^{p_i} \exp\left[-\left(\frac{J_i}{J_i^0}\right)^{p_i}\right] \quad (i = r, z).
\]

Here the prefactor ensures that \(1 \simeq \int_0^{\infty} dJ_i f_i\). The factor \((J_\phi/J_\phi^0)^{p_\phi}\) in the exponential controls the radial gradient in the velocity dispersion \(\sigma_r\).

The function \(f_\phi(J_\phi)\) in equation (2) determines the radial structure of the disc. We adopt

\[
f_\phi(J_\phi) \equiv \begin{cases} 0 & \text{when } J_\phi < 0 \\ \frac{M}{(2\pi)^2} \frac{J_\phi}{J_\phi^0} e^{-J_d/J_\phi^0} & \text{if } J_\phi > 0. \end{cases}
\]

The vanishing of \(f\) for \(J_\phi < 0\) sets this DF apart from several earlier DFs for discs, including that of Vasiliev (2019), who proposed using for \(J_\phi < 0\) the same formula as for \(J_\phi > 0\) multiplied by an additional factor

\[
\exp\left(\frac{J_d}{J_d^0}\right).
\]

When the DF at \(J_\phi < 0\) is taken to be a multiple of the DF at \(J_\phi > 0\) in this way, one is implicitly assuming that there is a retrograde disc like the prograde one but with reduced density. This is not the natural assumption: rather we suppose that any disc stars that are now on retrograde orbits have been scattered onto such orbits after being born on more-or-less eccentric co-rotating orbits. Thus most retrograde stars should be on highly inclined or eccentric orbits and an implicit assumption that the co-rotating disc is shadowed by a counter-rotating counterpart is unnatural. More natural is to deem any counter-rotating stars to belong to a hot component, such as the bulge or the stellar halo. Since the DF should be a continuous function on action space, we arrange for the DF to go to zero as \(J_\phi \rightarrow 0^+\) by replacing \(J_\phi^0\) in the first line of equation (21) in Vasiliev (2019) by \(J_\phi^0\).

| \(J_{\phi0}\) | \(J_{r0}\) | \(J_{z0}\) | \(p_r\) | \(p_z\) | \(J_{\phi0}\) | \(J_{d0}\) |
|---|---|---|---|---|---|---|
| 18 | 6 | 700 | 0.5 | 0.5 | 200 | 150 |
Here we present some examples of how the characteristics of a disc component vary with the parameters that specify its DF through equations (1) to (4). Table 1 lists the default values taken by parameters that are not explicitly given in each example.

### 2.1 How the parameters affect the DF’s values

The scale actions $J_{0}$ and $J_{0}$ set the scale on which the DF declines as $J_{r}$ and $J_{z}$ increase: larger $J_{0}$ implies larger vertical velocity dispersion $\sigma_{z}$, while increasing $J_{0}$ increase both $\sigma_{R}$ and $\sigma_{z}$. In the bottom panel of Fig. 1 we plot in black $f$ versus $J_{r}$ with $J_{r} = J_{z} = 0$ for two values of $J_{0}$. Corresponding values for $(J_{r}, J_{z}) = (10, 5)$ kpc km s$^{-1}$ are plotted in red. $J_{0}$ sets the scale over which the DF decreases as $J_{0}$ grows. The steeper the decline, the smaller will be the scale length of the resulting exponential disc, so $J_{0}$ is a surrogate for $R_{c}$.

In the middle panel of Fig. 1, the black and red curves are again for $(J_{r}, J_{z}) = (0, 0)$ and $(10, 5)$ kpc km s$^{-1}$ but the curves in each colour are now for $J_{0} = 10$ and $100$ kpc km s$^{-1}$. We see that even a ten-fold increase in $J_{0}$ has only a small effect on the value of the DF, so $J_{0}$ is not an important parameter: its role is to prevent the velocity dispersions diverging (for $p_{i} > 0$) or vanishing (if $p_{i} < 0$) at the centre.

The top panel of Fig. 1 shows how the DF varies with $J_{0}$ for two pairs of values, $(p_{r}, p_{z}) = (0, 0)$ and $(−0.25, −0.25)$, of the power-law exponents $p_{r}$ and $p_{z}$. Reducing the $p_{i}$ reduces the difference between the slopes of the black curves for circular orbits ($J_{r} = J_{z} = 0$) and the red curves for non-zero eccentricity and inclination. A large separation between the red and black curves implies a DF that falls rapidly with increasing eccentricity/inclination and thus small values of the velocity dispersions $\sigma_{R}$ and $\sigma_{z}$. Hence reducing the $p_{i}$ increases the velocity dispersions at large radii relative to their values at small radii.

### 2.2 Observational significance of parameters

The real-space structure of a component depends on the gravitational potential in which it resides. In general that potential depends on all the model’s components, so it will change with any component’s parameters. Here in the interests of clarity we fix the potential, choosing the potential of a model that provides a good fit to observational data. Thus we adopt a very realistic potential and explore how the structure of a single component depends on the parameters in its DF.

#### 2.2.1 Roles of the characteristic actions $J_{0}$, $J_{0}$ and $J_{0}$

The most important parameters are the three characteristic actions $J_{0}$, where $i = r, z, \phi$. While $J_{0}$ sets the component’s scale length, $J_{0}$ and $J_{0}$ set the in-plane and vertical dispersions, respectively. By setting these dispersions through $J_{0}$ and $J_{0}$, we simultaneously determine a component’s vertical density profile and asymmetric drift through standard dynamics.

The upper panel of Fig. 2 shows the relationship between $J_{0}$ and $J_{0}$ and the velocity dispersions in the plane...
at \( R_0 = 8.27 \) kpc. The lower panel shows that both dispersion satisfy quite accurately the relationship
\[
\sigma \simeq \sqrt{\langle \Omega_z \rangle J_{z0}},
\]
where \( \langle \Omega_z \rangle \) is the mean value of \( \Omega_z \) of the given disc’s stars sampled at \( (R, z) = (R_0, 0) \). A relationship of this type is suggested by the fact that
\[
J_z d\theta_z = v_z dz + dh, \quad h(x, y, z) = \langle v_z^2 \rangle.
\]
This argument indicates that in the case \( p_1 = 0 \), when the exponentials in the DF make \( J_z \) the same throughout action space, velocity dispersions should fall in parallel with frequencies. In particular we expect \( \sigma_z \) to fall with increasing distance from the plane, and both \( \sigma_R \) ad \( \sigma_z \) to fall roughly as \( 1/R \) with distance from the Galactic centre.

The upper panel of Fig. 3 shows the vertical density profiles of the components plotted in Fig. 2. The profiles steepen rapidly from zero in the plane to a nearly perfect exponential once the density has fallen to less than a tenth of its central values. The lower panel of Fig. 3 plots the resulting asymptotic scale-height against \( J_{z0} \). The relationship is almost exactly linear:
\[
\frac{z_0}{\text{kpc}} \simeq 0.012 \frac{J_{z0}}{\text{kpc} \text{km} \text{s}^{-1}}.
\]

Fig. 4 shows how the vertical velocity dispersion in a disc component varies with height above the plane. The dispersion falls slowly as expected from the argument above relating \( \sigma_z \) to the mean frequency \( \Omega_z \), while the density declines by orders of magnitude. The decline in \( \sigma_z \) is steepest near the mid-plane.

Fig. 5 shows the mean rotation speed (upper panel) and the radial velocity dispersion (lower panel) in increasingly hot disc components: from top to bottom in the upper panel the characteristic actions are \( (J_R, J_{z0}) = (8, 1), (12, 3), (18, 6), (24, 12), (30, 18), (35, 24) \) and \( (40, 30) \) kpc km s\(^{-1}\). The curves are terminated when \( \rho(z)/\rho(0) \) falls to \( 10^{-10} \).
Figure 6. Upper panel: the variation with $R$ of the density at $z = 0$ of disc components with $(J_{r0}, J_{z0}) = (8, 1), (18, 6)$ and $(30, 18)$ kpc km s$^{-1}$. The dashed red line shows an exponential decline with scale-length 2.72 kpc. Lower panel: the mid-plane velocity dispersions of these components in the radial (black) and vertical (red) directions.

Beyond $R_0$ these profiles are quite accurately exponential with the common scale length $R_d = 2.72$ kpc that is indicated by the red dashed line. This scale-length is set by the value $J_{v0} = 700$ kpc km s$^{-1}$ assigned to all components. Interior to $R_0$ the profiles acquire structure from the bulge via the self-consistently generated potential, as will be described below. Despite this structure, the profiles approximate exponentials at all $R$.

The lower panel of Fig. 6 shows the dependence on radius of the velocity dispersions of these components in the radial (black) and vertical (red) directions. The ratio $\sigma_z/\sigma_R$ increases rapidly with disc thickness, approaching unity for the hottest disc plotted [which has $(J_{r0}, J_{z0}) = (35, 24)$ kpc km s$^{-1}$]. This result is a natural consequence of the decision to have $J_{v0}$ span a smaller range ($8 - 40$ kpc km s$^{-1}$) than $J_{z0}$ ($1 - 30$ kpc km s$^{-1}$). In the thinnest disc (full curves) the vertical dispersion falls almost exponentially with $R$, while the radial dispersion falls more slowly. This result reflects the fact that $\Omega_z$ is much more sensitive than $\Omega_r$ to the contribution of the disc to the overall gravitational field: around $R_0$ the density of the disc that generates the gravitational field used for these figures is declining nearly exponentially with $R$.

2.2.2 Role of the exponents $p_r$ and $p_z$

The exponents $p_r$ and $p_z$ that occur in equation (3) moderate the rates at which the $\sigma_R$ and $\sigma_z$ decrease outwards. The results plotted above are for $p_i = 0$. We will find that $\sigma_R$ and $\sigma_z$ are reduced for slightly negative values of $p_i$, which cause $\sigma_R$ and $\sigma_z$ to decrease outwards more slowly.

In Fig. 7 we show the mid-plane density and velocity dispersions for $p_i = -0.25$ (full curves) and $p_i = 0$ in the case of a cold disc and a hot disc (lower pair of panels). Larger values of $p_i$ correspond to steeper outward gradients of the dispersions, as expected.

2.2.3 Role of the characteristic action $J_{v0}$

Fig. 8 shows the effect on a cold disc of increasing the characteristic action $J_{v0}$ from 10 to 100 kpc km s$^{-1}$. This parameter $J_{v0}$ affects the velocity dispersions $\sigma_R$ and $\sigma_z$ in the same way as the exponents $p_r$ and $p_z$.

The results plotted above are for $p_i = -0.25$ (full curves) and $p_i = 0$ in the case of a cold disc and a hot disc (lower pair of panels). Larger values of $p_i$ correspond to steeper outward gradients of the dispersions, as expected.
rameter is active only for non-zero \( p_i \), so Fig. 8 compares profiles for a disc with \( p_i = -0.5 \). As expected, \( J_0 \) has a perceptible effect only near the centre and in the sense that larger values lower the central velocity dispersions, and by hydrostatic balance, correspondingly increase the central mid-plane density. However, the effect on the plotted cold disc of even ten-fold decrease in \( J_0 \) is extremely small. The effect is no larger in the case of a hot disc.

3 HALO DFs

To build a self-consistent model Galaxy we require DFs for at least two hot components: the dark halo and the bulge / stellar halo. On account of the Galactic bar, a satisfactory model of the bulge would be non-axisymmetric so cannot be provided with our technology and we are obliged to model the bulge as an axisymmetric structure. Since few bulge stars reach the volume around the Sun in which the Gaia data are most precise, and the non-axisymmetric components of the bulge’s gravitational field decay strongly with increasing Galactocentric radius, we anticipate obtaining useful results even with a completely axisymmetric model.

For all three hot components of our Galaxy we employ a modified double power-law DF that is related to, but different from, those introduced by Posti et al. (2015) – hereafter P15. Our modifications of the P15 proposal are two-fold. First the phase-space density at the origin of action space is made finite by the mechanism introduced by Cole & Binney (2017). Second, the linear combination of actions \( h = J_1 + \eta J_2 + \eta_0 J_0 \) that forms the basis of the P15 proposal is replaced by a more complex function of the actions. A full explanation of this replacement lies beyond the scope of this paper and will be presented elsewhere (Binney 2022) but it addresses a weakness of the P15 proposal that was mentioned already by Binney & Piffl (2015), namely unphysical behaviour of velocity distributions at small \( |V_o| \).

The P15 halo DF is

\[
f = \frac{M}{(2\pi J_0)^3} \left( 1 + \frac{J_0}{h} \right)^\alpha \exp(-|g/J_{out}|^{\delta}).
\]

(8)

Instead of the linear combinations \( g(J) \) and \( h(J) \), we use

\[
g(J) = h(J) = J_1 e^{-\beta \sin(\pi/2)} + \frac{1}{2}(1 + c\xi)e^{\beta \sin(\pi/2)} \mathcal{L}.
\]

(9)

Here

\[
c \equiv \frac{L}{L + J_1}.
\]

(10)

with \( L \equiv J_z + |J_0| \) is a measure of an orbit’s circularity and

\[
\mathcal{L}(J_z, J_0) \equiv aL_z + bJ_0/J_1 + |J_0|
\]

(11)

is a generalisation of the total angular momentum. The constants \( a \) and \( b \) in the definition of \( \mathcal{L} \) are chosen such that \( \partial \mathcal{L}/\partial V_z \) vanishes as \( V_z \to 0 \). Specifically

\[
a = \frac{1}{2}(k + 1) \quad \text{and} \quad b = \frac{1}{2}(k - 1),
\]

(12)

where \( k \) is a number grater or equal to unity. We allow \( k \) to depend on energy by writing

\[
k(\xi) = (1 - \xi) F_{in} + \xi F_{out}.
\]

(13)

where \( F_{in} \) and \( F_{out} \) are constants and \( \xi \) is a dimensionless surrogate for energy that increases from zero at the origin of phase space to unity for marginally bound orbits through the formula

\[
\xi = \frac{j^2}{J^2} + 1
\]

(14)

where \( \alpha \approx 0.6 \) and

\[
j \equiv \frac{1.5J_1 + L}{L_0},
\]

(15)

where \( L_0 \approx 6J_0 \) is a scale action.

When \( F_{in} = F_{out} = 1 \), \( \mathcal{L} = L \) and in a spherical potential the distributions of \( V_0 \) and \( V_z \) are identical. If \( \alpha \) and \( L_0 \) are then set to the values suggested above, the velocity distribution becomes isotropic when \( \beta = 0 \) and radially biased as \( \beta > 0 \). To flatten the component in a spherical potential, one sets \( F_{in} \) and \( F_{out} \) above unity. If the potential is flattened, setting \( \beta = 0 \), and \( F_{in} \) and \( F_{out} \) to the ratio \( \nu/\Omega \) of vertical to azimuthal epicycle frequencies at small and large radii, respectively causes the velocity distribution to be nearly isotropic.

4 MODELS OF OUR GALAXY

In this section we use the DFs for disc components defined above to fit multi-component, self-consistent models to data for stars in the Gaia DR2 release (Gaia Collaboration & Brown 2018) for which the Radial Velocity Spectrometer (RVS) measured line-of-sight velocities (Gaia Collaboration & Katz 2018).

4.1 The data

We used the stellar locations and velocities computed by Schönrich et al. (2019) for the RVS stars. These authors adopted \( R_0 = 8.27 \) kpc as the distance to the Galactic centre and \( (U_0, V_0, W_0) = (11.1, 250, 7.47) \) km s\(^{-1}\) as the Sun’s Galactocentric velocity. According to this assumption, the Sun is moving towards the Galactic centre, ahead of the local circular speed \( \Theta_0 = 239 \) km s\(^{-1}\) and up out of the plane. They determined a distance-dependent selection function of the RVS sample and used the kinematic method of Schönrich et al. (2012) to determine the offset of Gaia DR2 parallaxes. After determining this offset, they obtained moments for the RVS sample and used the kinematic method of Schönrich et al. (2012) to determine the offset of Gaia DR2 parallaxes. After determining this offset, they obtained moments for the RVS sample and used the kinematic method of Schönrich & Aumer (2017).

We restricted the sample to stars with quoted uncertainties \( \sigma_\pi < \pi/5 \) smaller than a fifth of their parallax \( \pi \). The selection function of the RVS sample is complex (Everall et al. 2021) and a decision was taken not to engage with it in this preliminary work. From this decision it follows that we cannot attach significance to the number of stars listed in any spatial bin; only the distribution of these stars in velocity is significant.

We used the data given by Schönrich et al. (2019) to

| \( R - R_0 \) | \( -3 \) | \( -1.5 \) | \( -0.5 \) | \( 0.5 \) | \( 1 \) | 1.5 | 3 |
|---|---|---|---|---|---|---|
| \( |z| \) | 0 | 0.1 | 0.2 | 0.35 | 0.5 | 1 | 1.75 | 3 |
compute the histograms of Galactocentric velocity components $V_R$, $V_z$ and $V_\phi$ for each of 35 spatial bins. These bins are defined by a rectangular grid in the $(R, |z|)$ plane that stretches from $R_0 - 3\,\text{kpc}$ to $R_0 + 3\,\text{kpc}$ and from the plane up to $|z| = 3\,\text{kpc}$. Table 2 defines this grid.

4.2 Observational uncertainties

We investigated the impact of observational uncertainties on the histograms by drawing a sample from a realistic model using the approximation to the RVS selection function given by Schönrich et al. (2019) and comparing the histograms one obtains from the raw sample and from the sample after the stars have been scattered by the uncertainties in distance, proper motion and line-of-sight velocity. The uncertainties in each observable were averages of the DR2 uncertainties given for stars that lay at essentially the same distance from us – we used 50 equal-width bins in distance out to 3 kpc. We checked the correctness of the code by computing the differences between the observables before and after scattering and showing that the standard deviations of these differences agree with the uncertainties used to scatter the stars.

Errors in velocities will be most important for the most distant stars, so Fig. 9 shows their effect on the most distant bins, looking inward in the lower row of panels and outward in the upper row. The red histograms computed after scattering stars essentially obliterate the black histograms computed before scattering the raw sample. This perhaps surprising result allows us to neglect observational uncertainties even when using DR2 data, and in the following we compare histograms from Gaia with the model’s distribution of velocities at the barycentre of the Gaia stars in each bin. EDR3 data is superior to DR2 data in that its astrometric uncertainties are smaller, so if we upgraded to EDR3 data, more stars would pass our quality cuts so more stars would contribute to each histogram, and observational uncertainties would be even less significant. But given that our data/model comparisons are limited neither by Poisson noise nor by observational uncertainties, we continue to use the DR2 data.

4.3 The models

Our Galaxy models comprise eight components of which seven are defined by DFs and the eighth, the gas disc, is defined as a density distribution that contributes to the total potential. Following Dehnen & Binney (1998) the gas disc has density

$$\rho_g(R) = \frac{\sum_0}{4\pi z_0^3} R^{-2} R_3^{-2} R_0^{-2} \sech^2 \left( \frac{z}{2z_0} \right),$$ (16)

with $\sum_0 = 1.3 \times 10^9 \,\text{M}_\odot /\text{kpc}^{-2}$, $R_3 = 5\,\text{kpc}$, $R_0 = 5\,\text{kpc}$ and $z_0 = 60\,\text{pc}$. This gas disc has total mass $1.04 \times 10^{10} \,\text{M}_\odot$, and at the solar radius $R_0 = 8.27\,\text{kpc}$ has surface density

$$\rho_g(R_0) = 13.6 \,\text{M}_\odot /\text{pc}^{-2}$$

and central density

$$\rho_g(R_0, 0) = 0.057 \,\text{M}_\odot /\text{pc}^{-3}.$$

We include three spheroidal components – a dark halo, a bulge and a stellar halo – and four disc components – a thick disc and old ($\tau > 6\,\text{Gyr}$), middle-aged ($6 > \tau > 2\,\text{Gyr}$), and young components ($\tau < 2\,\text{Gyr}$) of the thin disc. Each of these discs has a DF of the type defined in Section 2, while the spheroidal components have DFs of the type defined in Section 3.

4.4 Fitting Gaia data

AGAMA computes the density of each component on an appropriate grid (spherical for the dark halo and cylindrical for the stellar components) using an assumed potential, then it solves Poisson’s equation for the resulting potential and re-determines the density in this new potential. This sequence of potential determinations is quite rapidly convergent. An excellent first guess at the potential is generally available from the potential of the last model created, so on a good laptop a new model can be computed in only ~ 5 minutes. After computing the self-consistent potential, AGAMA computes one-dimensional velocity distributions at 35 locations $(R, z)$ that are the barycentres of spatial bins into which the Gaia stars have been grouped. In Figs. 10 – 12 these locations $(R, z)$ are shown at the bottom of the $V_z$ panels – Table 2 gives the bin boundaries. The black histograms in Figs. 10 – 12 show the velocity distribution of RVS sample stars. We show histograms for only 12 of the 35 bins, being alternate bins in both $R$ and $|z|$. The top left corner of each panel shows the mean and standard deviation of the histogram. The small, varying non-zero means of $V_R$ and $V_z$ must arise from a combination of the Galactic bar and non-equilibrium dynamics. The red curves in Figs. 10 – 12 show the velocity distributions computed by AGAMA, normalised to the same number of stars as the corresponding black histogram.

Experiments with automatic optimisation of the parameters by the Nelder-Mead downhill simplex algorithm (e.g. Press et al. 1986) were unsuccessful in that the machine proved unable to find a convincing DF in an acceptable number of iterations. It seems that the high dimension of the relevant parameter space combined with the cost of each model evaluation (which arises more from computation of diagnostics than from model construction) sets a requirement for a more sophisticated machine-learning algorithm than naive least-squares minimisation. Therefore we now present a model that has been fitted by hand to the Gaia data and the vertical stellar density profile of Gilmore & Reid (1983). Table 3 lists the values of the parameters that define the model. Fig. 13 shows the density of the stars and dark halo in the meridional plane.

The lower block of panels in Fig. 11 breaks the model velocity histograms for the solar cylinder into the contributions of each stellar component. The full blue curve of the young disc only dominates at the core of the $V_z$ distribution, while the thick disc dominates in the wings and furthest from the plane everywhere except in the $V_z$ histogram, where at small $V_z$ the stellar halo dominates. The values of $J_{10}$ listed in Table 3 show a six-fold increase in $J_{10}$ between the thick disc and the old disc, whereas the characteristic actions $J_{10}$ that control the in-plane dispersions increase by a factor of six only between the young disc and the thick disc. Another noteworthy trend among the parameters are the steady increases in the values of $p_x$ and $p_z$ as one passes from the young disc to the thick disc. That is, the decrease outwards of the velocity dispersions is steeper in older components.

When fitting models by hand, velocity histograms like those shown in Figs. 10 to 12 yield numerous clues for pa-
Figure 9. The impact of observational uncertainty on velocity histograms. The black histograms show the result of plotting a sample drawn from a realistic model without allowing for errors, while the red histograms are obtained from the same sample after scattering by typical Gaia DR2 errors. Both rows are for bins with $|z| < 0.1$ kpc but the upper row is for a bin at $R - R_0 \sim 2$ kpc, while the lower row is for the bin at $R - R_0 \sim -2$ kpc.

Table 3. DFs fitted to Gaia data. Normalisations in $10^{10} M_\odot$, actions in kpc km s$^{-1}$

| Spheroids          | Mass | $J_c$  | $J_0$ | $J_{out}$ | $\alpha$ | $\beta$ | $\delta$ | $F_{in}$ | $F_{out}$ | $L_0$  |
|--------------------|------|--------|-------|-----------|-----------|---------|----------|----------|----------|--------|
| Dark halo          | 98   | 100    | 10000 | 20000     | 1.6       | 2.7     | 2        | 1.4      | 1.2      | 100000 |
| Stellar halo       | 0.1  | 5      | 600   | 100000    | 1         | 3.5     | 2        | 1.8      | 1.2      | 100000 |
| Bulge              | 1.1  | 5      | 19.5  | 200       | 0.5       | 1.8     | 2        | 3        | 3        | 100000 |

| Discs              | Mass | $J_{d0}$ | $J_{d0}$ | $J_0$ | $p_r$ | $p_z$ | $J_{d0}$ | $J_0$ |
|--------------------|------|----------|----------|-------|-------|-------|----------|-------|
| Young disc         | 0.175| 650      | 10       | 0.65  | -0.7  | -0.3  | 10       | 100   |
| Middle disc        | 0.75 | 600      | 17       | 2.8   | -0.35 | -0.1  | 10       | 700   |
| Old disc           | 1.0  | 550      | 22       | 5     | -0.25 | -0.1  | 10       | 700   |
| Thick disc         | 0.95 | 400      | 63       | 30    | 0.13  | 0.05  | 20       | 40    |

The red points are from Gilmore & Reid (1983), while the black squares show the double-exponential fit of Jurić et al. (2008). Both sets of points can be shifted vertically at will. To make the full black curve pass through the red data points, one has to adjust the normalisations of the dark halo and the discs, and also the disc parameters $J_{d0}$, in the choice of which Fig. 11 provides additional guidance. Crucially, the middle panels of Figs. 10 to 12 effectively fix the $J_{d0}$ parameters, so the only way to address the model density falling off too steeply with $z$ is to weaken the gravitational field near the plane by shifting mass from the disc to the dark halo. Hence, if the black squares from the SDSS survey were trusted more than the work of Gilmore & Reid (1983), one would make the dark halo more massive and the disc less massive.

Further clues to parameter choice can be drawn from Fig. 14, which shows, in blue, the densities contributed at different distances from the plane by the disc components (full curve: young disc; dotted: middle-aged disc; short-dashed: old thin disc; long-dashed: thick disc). The dashed black and dotted red curves, show the contributions of the dark and stellar halos, respectively. The full black curve shows the sum of the stellar components, which may be compared with the data shown by red dots and black squares.
problem might be addressed by reducing the thick disc’s value of \( p_z \).

In the bottom left panel of Fig. 10 the red model curve lies above the black data curve. The thick, middle and old discs all make significant contributions to the wings of these curves. The problem could be addressed by lowering \( J_{r0} \) for the middle disc at the risk of spoiling agreement in other panels.

In Figs. 10 to 12 the red curves for \( V_\phi \) have a tendency to be too high at \( V_\phi < 0 \). It would be unwise to take the model’s predictions for stars of low \( V_\phi \) too seriously because stars that approach the Galactic centre must be influenced by the Galactic bar, which is not included in the model. However, the model’s prediction of a surfeit of stars at \( V_\phi < 0 \) is least apparent in Fig. 10 for the bins closest to the centre, where the bar should be most influential. The red curves for \( V_\phi \) in the top-right panels of Figs. 10 to 12 are dominated by the stellar halo, which we have modelled as a non-rotating component because many studies (e.g. Schönrich et al. 2014, and references therein) have concluded that this is so. So perhaps one should revisit the issue of halo rotation in light of these Gaia data: the outer halo may be non-rotating (or even counter-rotating (Carollo et al. 2007)) but the inner halo may rotate detectably. In any event, ours is a very basic model of the stellar halo, that cannot address the empirical chemodynamical dichotomy into fairly isotropic, metal-poor and radially biased, metal-richer components (Helmi et al. 2018; Belokurov et al. 2018). Such a model must await resolution of the issue regarding choice of DFs discussed by Binney (2022).

### 4.4.2 Predictions of the models

The full black curve in Fig. 15 shows the circular-speed curve of the gravitational potential that emerges from the choices made for the DF by the self-consistency condition. After optimising agreement between model and data in the earlier figures, the model does not have freedom to adjust this curve, so it is a significant result that it agrees remarkably well with the data points from other recent work, including the Wegg & Gerhard (2013) dynamical model of the bar/bulge, modelling of Cepheid variables by Mróz et al. (2019) and Ablimit et al. (2020) and a study of 23 000 red giants by Eilers et al. (2019). In Fig. 15 the model curve is set by the kinematics of all stars in the restricted radial range \( 6 \lesssim R \lesssim 12 \) kpc rather than by objects presumed to be on near circular orbits over a wider range of radii. The curve extends in to the centre and out to large radii by

| \( \Sigma(R_0, 1.1 \text{kpc}) \) | gas | stars | DM | Total |
|---|---|---|---|---|
| \( \rho(R_0, 0) \) | 0.057 | 0.0317 | 0.0121 | 0.0436 |
Figure 11. Upper block of panels: the same as Fig. 10 but for $R - R_0 \in (-0.5, +0.5)$ kpc. Lower block of panels: how the model histograms above are assembled from the contributions of different components. Blue curves show the disc’s contributions (full, dotted and dashed as age increases and long-dashed for the thick disc). The contributions from the bulge and the stellar halo are plotted in heavy black and red, respectively.
Self-consistent models of our Galaxy

Figure 12. The same as Fig. 10 but for \( R - R_0 \in (1.5, 3) \) kpc.

Figure 13. The density in the meridional \((R, z)\) plane of stars (upper panel) and dark matter. The colour scales are logarithmic with two contours per dex.

Figure 14. Density as a function of distance from the plane at \( R_0 \). From lower left to upper right, the blue curves show the contributions of the young, middle-aged, old and thick discs. The dotted red curve shows the contribution of the stellar halo and the full black curve shows the sum of stellar contributions. The density of dark matter is shown by the dashed black line. The red data points show the measurements of Gilmore & Reid (1983) while the open squares show the analytic fit to data of Jurić et al. (2008).

In Fig. 15 the dotted blue curve shows the contributions to \( V_c \) from stars and gas. Baryons make the largest contribu-
tion to the circular-speed curve within 6 kpc and the bulge contributes more than the disc inside 2.3 kpc. In fact, in the inner kiloparsec the gravitational force from the disc is directed outwards, so it actually reduces $V_c$—see below.

Table 4 gives the local densities of baryons and dark matter according to the model shown here, and also the surface density $\Sigma(R_\odot) = 61.2 \, M_\odot \, pc^{-2}$ of mass that lies within 1.1 kpc of the plane at the solar radius. This comprises $13.6 \, M_\odot \, pc^{-2}$ of gas, $21.6 \, M_\odot \, pc^{-2}$ of stars and $26.0 \, M_\odot \, pc^{-2}$ of dark matter. For comparison, Kuiper & Gilmore (1991), using a sample of K dwarfs, inferred $\Sigma(R_\odot, 1.1) = 71 \pm 6 \, M_\odot \, pc^{-2}$, while Holmberg & Flynn (2004), using K giants as tracers, obtained $\Sigma(R_\odot, 1.1) = 74 \pm 6 \, M_\odot \, pc^{-2}$. More recently, by counting stars in Gaia EDR3 Everall et al. (2022a) infer $(23 \pm 2.4) \, M_\odot \, pc^{-2}$ of stars, and by applying the Jeans equations to data from Gaia DR3 Nitschai et al. (2021) find $\Sigma(R_\odot, 1.1) = (55 \pm 1.7) \, M_\odot \, pc^{-2}$.

Table 4 reports the model's local density of stars to be $0.0317 \, M_\odot \, pc^{-3}$, which lies below but is consistent with the value $(0.0366 \pm 0.005) \, M_\odot \, pc^{-3}$ Everall et al. (2022a) inferred from Gaia EDR3. The local density of the stellar halo is $4.7 \times 10^{-5} \, M_\odot \, pc^{-3}$. Table 4 reports the local density of dark matter to be $\rho_{DM}(R_\odot) = 0.0121 \, M_\odot \, pc^{-3} = 0.46 GeV cm^{-3}$. For comparison, from their study of red giants, Eilers et al. (2019) inferred the local density of dark matter to be $0.0079 \pm 0.0008 \, M_\odot \, pc^{-3}$, while Nitschke et al. (2021) obtained $0.0089 \pm 0.0006 \, M_\odot \, pc^{-3}$ and Cantau et al. (2020) found $0.0085 \pm 0.0005 \, M_\odot \, pc^{-3}$. Thus our model is more dark-matter dominated than several recent studies prefer. Evidently we still have some way to go before a consistent picture emerges of the local structure of our Galaxy. A comprehensive review of estimates of the dark halo can be found in de Salas & Widmark (2021).

Given that the circular speed is falling at $R_\odot$, there is scope for confusion in this area because the surface density $\Sigma(z)$ within $z$ of the plane and the vertical acceleration $K_z(z)$ satisfy

$$\Sigma(z) = \frac{1}{2\pi G} \left( K_z + \frac{z \, dV_c^2}{R \, dz} \right). \hspace{1cm} (17)$$

Often values of $K_z/(2\pi G)$ are quoted instead of $\Sigma$. In the present model $K_z/(2\pi G) = 64.1 \, M_\odot \, pc^{-2}$.

The full black curve in Fig. 16 shows the sum of the mid-plane densities of all the stellar components. It is very nearly exponential although comprised of contributions that have very different scale-lengths. The steep black, dotted line shows the nearly exponential density of the bulge. At the Sun the dark halo contributes a third of the density contributed by stars.

The lower panel of Fig 16 shows the mean-streaming speeds $\langle V_\phi \rangle$ of the stellar discs at the mid-plane. For the thin-disc components $\langle V_\phi \rangle$ is only slightly smaller than $V_c$, while the value for the thick disc is generally reduced by an asymmetric drift in excess of $>10 \, km \, s^{-1}$.

Fig. 17 shows what the Galaxy would look like when seen face-on. The solid blue curve in the left panel shows the contributions of the four disc components while the dotted black curve shows the density obtained on adding the contributions of the bulge and stellar halo. The profiles are not far from exponential.

The dashed red line shows the slope characteristic of a disc with a scale length $R_d = 2.6 \, kpc$, slightly larger than the value $R_d = 2.4 \, kpc$ reported by Robin et al. (2003) from mid-IR photometry. There is a strong correlation between the disc’s scale length and degree of dark-matter domination.
Figure 17. Quantities at face-on projection. Left: the disc’s surface density (solid blue curve) and the surface density including all stellar components (black dotted curve). The dashed red line show the slope of a disc with scale length 2.6 kpc. Centre: rms heights of the disc components (full line for thin disc, dotted for middle disc, short-dashed for old disc and long-dashed for thick disc). Right: projected line-of-sight velocity dispersions of the disc components.

Figure 18. From left to right the radial dependence of the velocity dispersions $\sigma_R$, $\sigma_z$ and $\sigma_\phi$. The black long-dashed curve is for the dark halo, the red dotted curve is for the stellar halo. The dispersions of the discs are given by the four blue curves at the bottom, with the young disc having smallest dispersions and the thick disc the largest. The black dashed curves that reach only to $R = 6$ kpc show the dispersion in the bulge.

Table 5. Characteristics of the dark halo. Distances in kpc, masses in $10^{10} M_\odot$.

|        | $r_s$ | $r_{200}$ | $M_{200}$ |
|--------|-------|-----------|-----------|
| Dark halo | 20.0  | 192       | 80.5      |

$R = 20$ kpc. The stellar halo (red dotted curves) is radially biased, with $\sigma_R$ similar to the dispersions of the dark halo.

The black dotted curves in Fig. 18, which show the velocity dispersions of the bulge, show $\sigma_R$ and $\sigma_z$ tracking the dispersions of the thick disc, but $\sigma_\phi$ falling much more steeply.

In all disc components $\sigma_R > \sigma_\phi > \sigma_z$, everywhere although $\sigma_R$ invariably falls faster with increasing $R$ than does $\sigma_z$.

4.5 The halo without baryons

Our understanding of the statistical properties of dark haloes depend to a large extent on cosmological simulations that exclude baryons. Comparison of the Galaxy’s dark halo with expectations raised by such simulations requires knowledge of the structure of the dark halo before the baryons fell in. The classic assumption is that the accumulation of baryons was an adiabatic process (Blumenthal et al. 1986; Sellwood & McGaugh 2005). In this case, the primordial structure of the dark halo can be recovered by constructing the self-consistent model defined by the dark halo’s DF...
The structure of the dark halo before the infall of baryons. Upper panel: the full black curve shows the density profile generated in isolation by the DF of the Milky Way’s dark halo. The dashed blue curve shows the profile generated by this DF after removal of the Cole–Binney core. The square marks the radius, \( r_s = 20.0 \) kpc, at which \( \ln \rho / \ln r = -2 \) and the red curve shows the density profile of the NFW halo with the same scale radius. The black dashed curve shows the density of the dark halo as it actually is in the Galaxy. The sloping dotted lines show the asymptotic slopes, \(-1\) and \(-3\) of an NFW profile. Lower panel: the circular speed and the principal velocity dispersions of the halo before removal of the Cole–Binney core.

There is now significant evidence that baryonic infall was significantly non-adiabatic (Chan et al. 2015; Pontzen & Governato 2014; Pascale et al. 2018). Indeed it is natural that in the region in which baryons dominate the gravitational field, fluctuations in the gravitational field will have upscattered dark-matter particles, with the consequence that their phase-space density is now lower than it was before the baryons fell in. The Cole–Binney core of our halo is designed to model the consequences of this physics, so to obtain the best estimate of what the dark halo looked like before the baryons fell in, we should remove this core from the DF. The dashed blue curve in Fig. 19 shows the density profile of this primordial dark halo. It closely hugs the red curve, which plots the density of the NFW model that has the same scale radius \( r_s = 20 \) kpc (the radius at which \( \ln \rho / \ln r = -2 \)). This closeness of fit is no accident: the search for a dark halo DF was confined to ones that in isolation generate an NFW profile after removal of the Cole–Binney core. This curve yields \( r_{200} = 192 \) kpc for the radius at which the mean density is 200 times the mean cosmic density and \( M_{200} = 0.805 \times 10^{12} M_\odot \) for the mass interior to that radius. For comparison, Eilers et al. (2019) derive \( M_{200} = (0.725 \pm 0.026) \times 10^{12} M_\odot \), while two estimates obtained by Ablimit et al. (2020) are \( |r_s| = (14.5 \pm 0.5) \) kpc, \( M_{200} = (0.66 \pm 0.07) \times 10^{12} M_\odot \), \( r_{200} = (179 \pm 5) \) kpc and \( |r_s| = (14.7 \pm 0.4) \) kpc, \( M_{200} = (0.82 \pm 0.05) \times 10^{12} M_\odot \), \( r_{200} = (192 \pm 4) \) kpc. When comparing these estimates one must bear in mind that the RVS data can constrain the density only at \( R \lesssim 12 \) kpc, so quoted values of even \( r_s \) involve significant extrapolation and values of \( r_{200} \) and \( M_{200} \) are dangerously exposed to assumptions about the halo’s density profile – for example, the mass of our halo could be made larger either by increasing its truncation action \( J_{\text{cut}} \) or making its outer slope shallower. The globular cluster system provides some sensitivity to mass that lies beyond \( r = 12 \) kpc, and from a study of this system Wang et al. (2021) inferred \( 0.54 \lesssim M_{200}/10^{12} M_\odot \lesssim 0.78 \).

Using both globular clusters and dwarf spheroidal galaxies, and taking into account perturbation by the LMC, Correa Magnus & Vasiliev (2022) finds \( M_{200} = (0.7 \pm 1.6) \times 10^{12} M_\odot \).

Whereas ours estimates are based on a reconstruction of the dark halo prior to baryon infall, those of earlier work relate to models of the halo as it now is.

The dashed curve in the upper panel of Fig. 19 shows the dark halo’s current density profile; despite the upscattering of dark-matter particles, the gravitational pull of the baryons has more than tripled its density at \( R \approx 1 \) kpc from what it was even before the upscattering. The density enhancement remains significant out to \( r_s \). Comparing the current of the dashed and full black curves in the upper panel of Fig. 19 we see that the widespread assumption that the dark haloes of luminous galaxies have NFW profiles is indefensible. Dark haloes will have such profiles only if upscattering of dark-matter particles precisely cancels adiabatic compression, and there are no grounds for believing this to be true. For more on this topic, see Cautun et al. (2020).

The lower panel of Fig. 19 plots velocities associated with this reconstruction of the primordial dark halo. Its circular speed, plotted in red, peaks at \( 170 \text{ km s}^{-1} \) and its velocity distribution is nearly isotropic. Removal of the core changes the dispersions negligibly at \( R > 5 \) kpc.

### 4.6 Chemodynamics

Fig. 4 of Hayden et al. (2015) is one of the most significant products of the APOGEE survey (Majewski et al. 2017). It shows the distribution of stars in the \([\alpha/\text{Fe}], [\text{Fe}/\text{H}]\) plane in each of 18 bins in the \((R,z)\) plane. The robustness of this result was recently confirmed by Eilers et al. (2022). The upper block of panels in Fig. 20 is an analogous plot using the same data, kindly supplied by M. Hayden. Two stellar populations are clearly visible. There is a high-\(\alpha\) population that always occupies the same location in the chemical plane but is much more prominent at large \(|z|\) and small \(R\); it has essentially disappeared at \(R > 10\) kpc and is very faint at \(|z| < 3\) kpc. The other, normal-\(\alpha\), population dominates near the plane and at large \(R\). It moves to lower \([\text{Fe}/\text{H}]\) with increasing \(R\).

If we conjecture how the stellar populations of our dynamical model are distributed in the chemical plane, we can create an analogous figure. The lower block of panels in Fig. 20 shows the result of doing this as follows.

At the centre of each spatial bin we determined the fractional contribution to the stellar density by stars be-
longing to each of the six populations, young disc, middle
disc, old disc, thick disc, stellar halo and bulge. Then mock
stars were drawn from the velocity distributions of the com-
ponents in the corresponding proportions, the total number
of stars being the same as the number of real stars reported
by Hayden et al. (2015) for that spatial bin. The mock stars
were assigned values of $[\text{Fe/H}]$ and $[\alpha/\text{Fe}]$ by sampling its
component’s chemical pdf. Each pdf was a two-dimensional
Gaussian with principal axes

$$
\begin{pmatrix}
  x \\
  y
\end{pmatrix} = 
\begin{pmatrix}
  \cos \theta & -\sin \theta \\
  \sin \theta & \cos \theta
\end{pmatrix}
\begin{pmatrix}
  [\text{Fe/H}] \\
  [\alpha/\text{Fe}]
\end{pmatrix}
$$

(18)

that are rotated with respect to the $[\text{Fe/H}]$ and $[\alpha/\text{Fe}]$ axes. The angle $\theta$ being small, $x$ is close to $[\text{Fe/H}]$ while $y$ is close
to $[\alpha/\text{Fe}]$.

In the normal-$\alpha$ components, the distribution in $y$ was
independent of $J_\phi$ but the mean value of $x$ was taken to be

$$
\tau = \tau_\odot + \frac{\alpha \tau}{dJ_\phi} (J_\phi - J_\phi) ,
$$

(19)

where $J_\phi$ is the circular angular momentum at the Sun and
$\tau_\odot$ is approximately the mean metallicity of the component
at $R_0$. This dependence of $\tau$ on $J_\phi$ generates a metallicity
gradient $d[\text{Fe/H}]/dR \simeq V_c d\tau/dJ_\phi$ in the disc.

The top row of Fig. 20 shows that the high-$\alpha$ popula-
tion is scarcely visible at $R \gtrsim 10$ kpc. The obvious way to
engineer this fading of the high-$\alpha$ is to make the probabili-

Figure 20. Upper block of panels: a plot analogous to Fig. 4 in Hayden et al. (2015) using the same APOGEE data. Lower block: a
similar plot computed from our dynamical model using the chemical compositions of the components given by equations (18) to (21).
The mean values of $R$ and $z$ for the stars/pseudo-stars that fall in each spatial bin are given at upper right of each panel. Contours are
drawn that enclose 90 percent and 50 percent of stars.
Table 6. Parameters of the chemical pdfs used to construct Fig. 20. The units of \( x, y, \sigma_x, \sigma_y \) are dex, those of \( V_c \) are \( J_\phi \) are \( \text{kpc km s}^{-1} \), those of \( \theta \) are degrees and those of \( J_{z0} \) and \( \Delta_z \) are \( \text{kpc} \). 

| Component     | \( \pi_\circ \) | \( V_c d\pi/dJ_\phi \) | \( \bar{x} \) | \( \sigma_x \) | \( \sigma_y \) | \( \theta \) | \( J_{z0} \) | \( \Delta_z \) |
|---------------|---------------|-----------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Young disc    | 0.2           | -0.07           | 0           | 0.1         | 0.03        | 2           | -           | -           |
| Middle disc   | 0             | -0.07           | 0.03        | 0.1         | 0.03        | 2           | -           | -           |
| Old disc      | -0.2          | -0.07           | 0.06        | 0.1         | 0.05        | 2           | -           | -           |
| Thick disc    | -0.5          | 0               | 0.2         | 0.2         | 0.03        | 7           | 50          | 15          |
| Stellar disc  | -1            | 0               | 0.2         | 0.3         | 0.1         | 3           | -           | -           |
| Bulge         | -0.5          | 0               | 0.2         | 0.2         | 0.02        | 3           | -           | -           |

Figure 21. Upper panel: the probability that a thick-disc star has high-\( \alpha \) chemistry. Lower panel: the \( J_z \)-dependence of the DF of the high-\( \alpha \) population.

\[ P_{hi} = \frac{1}{2} \left[ 1 - \tanh \left( \frac{J_z - J_{z0}}{\Delta_z} \right) \right]. \tag{20} \]

produced results that were unsatisfactory in two respects: i) an implausibly rapid transition was required, specifically \( J_{z0}/\Delta_z \gg 100 \), and ii) it left the high-\( \alpha \) population too prominent at \( z \simeq 0.7 \) and 0.2 kpc. The more satisfactory results shown in the lower block of panels of Fig. 20 were obtained by taking the probability of belonging to the high-\( \alpha \) population to be

\[ P_{hi}(J_z) = \frac{1}{2} \left[ 1 + \tanh \left( \frac{J_z - J_{z0}}{\Delta_z} \right) \right]. \tag{21} \]

Now \( P_{hi} \approx 1 \) for \( J_z \gg J_{z0} \) and \( P_{hi} \approx 0 \) for \( J_z \ll J_{z0} \). If a star was assigned to the high-\( \alpha \) portion of the thick disc, its chemistry was drawn from a Gaussian in \((x, y)\) independently of \( J_\phi \), while if it was assigned to the normal-\( \alpha \) disc, its chemistry was determined as if it were a member of the old disc. The upper panel of Fig. 21 is a plot of \( P_{hi}(J_z) \) for the chosen values of \( J_{z0} \) and \( \Delta_z \).

Table 6 lists for each component the numbers that define the two-dimensional Gaussians pdfs, and the parameters \( J_{z0} \) and \( \Delta_z \) that determine the distribution of high-\( \alpha \) stars within the thick disc. Comparison of the upper and lower blocks of panels in Fig. 20 shows that the model captures the essential features of the observations, including the dominance of the high-\( \alpha \) population at small \( R \) and large \( z \), and the drift of the normal-\( \alpha \) population towards lower [Fe/H] with increasing \( R \).}

5 DISCUSSION

5.1 Extracting information from data

Much previous galaxy modelling has been based on the Jeans equations (Binney et al. 1990; Robin et al. 2003; Garbani et al. 2012; Read & Steger 2017; Nitschka et al. 2020, 2021; Sivertsson et al. 2022). Gaia produces such large samples that the full distribution of velocities in a small spatial region can be determined with exquisite precision. These distributions are highly non-Gaussian, so a technique which characterises them by just two or three moments is throwing away valuable information.

Extraction of the maximum information from a stellar survey is probably achieved by evaluating the likelihood of the survey data given candidate models. Unfortunately, this technique is unfeasible when there is significant obscuration by dust and the density of dust is not well known (e.g. Li & Binney 2022b). Moreover, evaluating likelihoods for individual stars is very costly for samples as large as those yielded by Gaia. In light of these considerations there is a compelling case for binning stars spatially and evaluating the fit to models of the resulting velocity distributions. The comparison would be better made by evaluating likelihoods in velocity space than by binning stars by their single velocity components, but the cost of likelihood evaluation seems currently prohibitive. Humans can more easily assess fit quality from one-dimensional velocity distributions than from three-dimensional ones, but for an MCMC search of model space one should probably bin stars into cell in velocity space that are not the two-dimensional slices right across the space employed here.

5.2 The dark-matter baryon balance

Hand-fitting models to the data reveals a strong connection between the radial structure of the dark halo and the extent to which the disc is self-gravitating. On the one hand, the large distances from the plane reached by thick-disc stars places an upper limit on the disc’s mass. On the other hand, the clear outwards decline in the circular speed at \( R \gtrsim 6 \) kpc requires both a significant contribution to the gravitational field from stars, and that the circular speed contributed by the dark halo does not rise strongly in this radial range. The latter requirement places an upper limit on the core parameter \( J_c \), and a lower limit on the index \( \alpha \) of the dark halo’s DF. It would be very interesting to use the MCMC algorithm to explore how big a region of model space acceptable models occupy.
5.3 We need better halo DFs

Action-based DFs that envisage a non-zero density of stars in the neighbourhood of $J_\phi = 0$ need to conform to certain constraints (Pfiff et al. 2015; Binney 2022). The disc DFs introduced here have vanishing star densities at $J_\phi = 0$ so do not need to engage with these constraints, but the spheroidal DFs of the dark halo, the bulge and the stellar halo should conform to these constraints, and they do so to only a limited extent. To avoid excessive violation of the constraints, and the introduction of unphysical features, the parameters of the spheroidal DFs have been chosen to avoid plausible levels of radial bias. The absence of radial bias in the dark halo and the bulge is not significant observationally because the data contain few bulge stars and no dark-matter particles. But the stellar halo needs to have greater radial bias and work to enable this should be a high priority. Moreover, restricting the velocity anisotropy of the dark halo’s DF limits the dark halo’s flattening, and a flattened halo will imply a less massive disc. Hence, better functional forms for spheroidal DFs may significantly change our understanding of the disc-halo balance.

5.4 Structure of the high-α population

The quality of the fit to the Hayden et al. (2015) data that is provided by the restriction of high-α stars to orbits with large $J_z$ is so good that it seems certain that the restriction is real, but its historical origin is far from evident. The a priori expectation is that high-α stars would be restricted to low $|J_\phi|$ because these would be the first stars to form in the natural inside-out formation picture. However, fruitless attempts to model the data in this way left us convinced that it is to orbits with large $J_z$ that high-α stars are confined.

In light of this result the high-α population would form a very strange component if considered in isolation, for it comprises stars that have apocentres in $z$ that fall in quite a narrow range – the lower panel of Fig. 21 illustrates this fact by showing how the DF of this population depends on $J_z$. It seems that at low $|z|$ the density of high-α stars increases with $z$, peaking at a star with $J_z \simeq 50 \text{ kpc km s}^{-1}$ turns around.

Could the $J_z$-distribution of the high-α stars have arisen when a primordial, high-α, thin disc was shattered in a merger (Belokurov et al. 2020)? The problem with this idea is that we would expect the merger to spread stars in $J_z$ rather than shift them all up in $J_z$, leaving orbits around $J_z = 0$ vacant until occupied by the subsequent formation of low-α stars. A possibly more promising scenario that the high-α stars were stripped from a satellite that was finally disrupted on a nearly circular orbit, and their characteristic value of $J_z \sim 50 \text{ kpc km s}^{-1}$ was the vertical action of the satellite when it dispersed. Another possibility is that we are seeing the result of ‘levitation’ (Sridhar & Touma 1996): stars trapped in a resonance such as $\Omega_2 = \Omega_z$ moving up in $J_z$ along with the resonance as the disc gains mass and the values of $\Omega_z/\Omega_r$ in the plane increase.

5.5 What to do with all Gaia stars

Gaia is a survey instrument that provides photometry of unprecedented precision. Hence it is the ultimate tool with which to determine how the density of stars varies with location. Yet we have completely neglected this possibility and instead used a forty-year-old determination of the density of stars above the Sun (Gilmore & Reid 1983). The problem is that the Gaia catalogues have complex selection functions. Even after these selection functions have been determined (Everall & Bouwens 2022), to predict how many stars will be catalogued at a given sky position one needs accurate knowledge of the extinction as a function of distance along the relevant line of sight. This knowledge is currently woefully lacking for most sight lines, but fortunately along lines of sight at $|b| \gtrsim 80$ deg extinction is sufficiently low that a poor understanding of its distribution is not a major problem, so it is possible to deduce the vertical profile $\rho_\star(z)$ of the disc near the Sun from Gaia data, and Everall et al. (2022b,a) have recently done this. Their work appeared after the present work was largely completed and a decision was made not to replace the Gilmore & Reid (1983) data with the new determination of $\rho_\star(z)$ because it would be fundamentally sounder to fit the self-consistent models directly to Gaia’s star counts rather than to a density profile that has been fitted to the star counts – Everall et al. (2022b,a) needed to assume that the luminosity function is independent of $z$, though it will not be because younger populations are more prominent at low $|z|$, and older and more metal-poor populations dominate at high $|z|$.

Since each population of the current models comes with an age range and therefore a luminosity function, predicting the star counts along a dust-free line of sight is straightforward. Comparison of these star counts, modified by the relevant Gaia selection function, with the actual star counts will then yield a constraint on the models that can replace that provided here by Fig. 14.

Probably the best way to use comparisons between predicted and actual Gaia star counts along general lines of sight is for the construction of a global three-dimensional dust map. Current attempts to construct dust maps are restricted to the tiny fraction of stars for which it is possible to obtain an extinction (Sale & Magorrian 2014; Green 2018; Green et al. 2019; Lallement et al. 2022). Such stars, mainly intrinsically blue stars, are for some reason susceptible to having their absolute magnitudes predicted. A model of the type presented here, combined with a trial dust model, yields for each field of view star counts as a function of parallax that can be compared with the actual star counts down to the faintest magnitudes. Hence every star in the Gaia catalogue can be used to constrain the dust model.

6 CONCLUSIONS

Fully self-consistent dynamical galaxy models constitute valuable tools for the interpretation of observational data for both our Galaxy and external galaxies because they pull constraints from a variety of observational probes into a coherent physical framework. They minimise the number of parameters that must be determined from observations by exploiting to the full the constraints imposed by the laws of dynamics.

Such models are inherently steady-state models, and currently AGAMA can only construct axisymmetric models. However, from such a model one may easily draw an N-body
sample, and by perturbing this either study time-dependent phenomena (e.g. Binney & Schönrich 2018; Al Kazwini et al. 2022) or construct barred models by the made-to-measure technique (Syer & Tremaine 1996; de Lorenzi et al. 2007).

Hitherto, self-consistent models of our Galaxy (Piffl et al. 2015; Binney & Piffl 2015; Cole & Binney 2017) have represented the discs with the quasi-exponential DF. Unfortunately, the radial and vertical epicycle frequencies $\kappa(J_2)$ and $\nu(J_2)$, and the circular radius $R_c(J_2)$ play significant roles in this DF, so the DF is only fully specified when a potential is given. This fact makes the quasi-isothermal DF ill-suited to self-consistent galaxy modelling because the galaxy’s potential should emerge from the DF, not precede it.

Therefore in Section 2 we introduced a new family of DFs for discs that are fully specified by the disc’s mass and the values of seven parameters. We have elucidated the physical significance of the parameters: three characteristic actions set the scale length of the disc and the in-plane and vertical velocity dispersions. Two further parameters set the radial gradients of the dispersions and the final two control the central structure of the disc.

In Section 4 we built a model of our Galaxy that comprises four of these stellar discs, a gas disc, and three spheroidal components. The model reproduces to good accuracy the velocity distributions of stars in the Gaia RVS sample at 35 locations distributed through the rectangle $R_0 \pm 3\,\text{kpc}$ and $|z| \lesssim 3\,\text{kpc}$. It also fits old observations of the density of stars as a function of $|z|$ in the column above the Sun, which effectively set the balance between baryon and dark mass.

The model’s circular-speed curve is tightly constrained over the $\sim 6\,\text{kpc}$ radial range covered by the Gaia RVS sample and it agrees well with recent determinations using specific tracer populations. Its stellar disc has mass $3.9 \times 10^{10} M_\odot$, scale length 2.6 kpc and local surface density $22 M_\odot \, \text{pc}^{-2}$. Its dark halo has mass $M_{200} = 8.05 \times 10^{10} M_\odot$ scale radius $r_s = 20\,\text{kpc}$ and local density $0.012 M_\odot \, \text{pc}^{-3}$. Its original central cusp is assumed to have been eroded by interactions with baryons. We reconstructed the dark halo prior to baryon infall. Prior to the addition of baryons it had peak circular speed $170 \,\text{km} \, \text{s}^{-1}$ and virial radius $r_{200} = 192\,\text{kpc}$.

The model’s stellar disc is a superposition of three relatively cool “thin” discs of increasing velocity dispersions and therefore thickness, and a much hotter and thicker disc. The disc’s scale lengths decrease as their dispersions and thicknesses increase. The hotter a disc is, the faster its velocity dispersions fall off with increasing radius.

After assigning a chemical composition to each of the model’s six stellar components, we could predict from the model how the chemistry of stars varies with location in the $Rz$ plane. We identified simple compositions of the components that yielded good agreement between the model’s predictions and the chemistry measured by the APOGEE survey (Hayden et al. 2015). The key feature of these compositions is the restriction of high-$\alpha$ stars to orbits with high $J_z$. This restriction ensures that normal-$\alpha$ stars dominate both at low $|z|$ and small $R$, and at high $|z|$ and large $R$.

The Galaxy model presented here is designed mainly to illustrate the opportunities that are opened up by the new disc DFs. There is much scope to extend the work by a wider exploration of model space and the addition of more and better observational constraints.

Here we have presented a single model that provides good fits to most of the data. Attempts to find a best-fitting model were unsuccessful on account of the high dimensionality of the model parameter space and the cost of computing observables from a model. Nevertheless, soon the MCMC algorithm should be used to explore model space starting from this model and thus to obtain insight into the range of model parameters that generate acceptable models.

ACKNOWLEDGEMENTS

This work was supported by the UK Science and Technology Facilities Council under grant number ST/N000919/1. JB also acknowledges support from the Leverhulme Trust through an Emeritus Fellowship. We thank Michael Hayden for providing the APOGEE data re-plotted in Fig. 20.

This work presents results from the European Space Agency (ESA) space mission Gaia. Gaia data are being processed by the Gaia Data Processing and Analysis Consortium (DPAC). Funding for the DPAC is provided by national institutions, in particular the institutions participating in the Gaia MultiLateral Agreement (MLA). The Gaia mission website is https://www.cosmos.esa.int/gaia. The Gaia archive website is https://archives.esac.esa.int/gaia.

DATA AVAILABILITY

The code that generates Galaxy models can be downloaded from the AGAMA website https://github.com/GalacticDynamics-Oxford/Agama

REFERENCES

Abazajian K. et al., 2003, AJ, 126, 2081
Ablimit I., Zhao G., Flynn C., Bird S. A., 2020, ApJL, 895, L12
Al Kazwini H. et al., 2022, A&A, 658, A50
Antoja T. et al., 2018, Nat, 561, 360
Belokurov V., Erkal D., Evans N. W., Koposov S. E., Deason A. J., 2018, MNRAS, 478, 611
Belokurov V., Sanders J. L., Fattahi A., Smith M. C., Deason A. J., Evans N. W., Grand R. J. J., 2020, MNRAS, 494, 3880
Binney J., 2010, MNRAS, 401, 2318
Binney J., 2012, MNRAS, 426, 1328
Binney J., 2014, MNRAS, 440, 787
Binney J., 2022, in preparation, xxx
Binney J., McMillan P., 2011, MNRAS, 413, 1889
Binney J., Piffl T., 2015, MNRAS, 454, 3653
Binney J., Schönrich R., 2018, MNRAS, 481, 1501
Binney J. J., Davies R. L., Illingworth G. D., 1990, ApJ, 361, 78
Blumenthal G. R., Faber S. M., Flores R., Primack J. R., 1986, ApJ, 301, 27
Blumenthal G. R., Faber S. M., Flores R., Primack J. R., 1986, ApJ, 301, 27
Boubert D., Everall A., 2020, MNRAS, 497, 4246
Bovy J., Rix H.-W., 2013, ApJ, 779, 787
Carollo D. et al., 2007, Nat, 450, 1020
Cautun M. et al., 2020, MNRAS, 494, 4291
