

Jeans modelling of the Milky Way’s nuclear stellar disc

Mattia C. Sormani¹,¹★ John Magorrian²,² Francisco Nogueras-Lara³,³ Nadine Neumayer,³
Ralph Schönrich⁴,⁴ Ralf S. Klessen^{1,5} and Alessandra Mastrobuono-Battisti⁶

¹*Institut für Theoretische Astrophysik, Zentrum für Astronomie, Universität Heidelberg, Albert-Ueberle-Str 2, D-69120 Heidelberg, Germany*

²*Rudolf Peierls Centre for Theoretical Physics, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, UK*

³*Max Planck Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany*

⁴*Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK*

⁵*Interdisziplinäres Zentrum für Wissenschaftliches Rechnen, Universität Heidelberg, Im Neuenheimer Feld 205, D-69120 Heidelberg, Germany*

⁶*Department of Astronomy and Theoretical Physics, Lund Observatory, Box 43, SE-22100 Lund, Sweden*

Accepted 2020 September 9. Received 2020 September 7; in original form 2020 July 13

ABSTRACT

The nuclear stellar disc (NSD) is a flattened stellar structure that dominates the gravitational potential of the Milky Way at Galactocentric radii $30 \lesssim R \lesssim 300$ pc. In this paper, we construct axisymmetric Jeans dynamical models of the NSD based on previous photometric studies and we fit them to line-of-sight kinematic data of the Apache Point Observatory Galactic Evolution Experiment (APOGEE) and silicon monoxide (SiO) maser stars. We find that (i) the NSD mass is lower but consistent with the mass independently determined from photometry by Launhardt et al. Our fiducial model has a mass contained within spherical radius $r = 100$ pc of $M(r < 100 \text{ pc}) = 3.9 \pm 1 \times 10^8 M_\odot$ and a total mass of $M_{\text{NSD}} = 6.9 \pm 2 \times 10^8 M_\odot$. (ii) The NSD might be the first example of a vertically biased disc, i.e. with ratio between the vertical and radial velocity dispersion $\sigma_z/\sigma_R > 1$. Observations and theoretical models of the star-forming molecular gas in the central molecular zone suggest that large vertical oscillations may be already imprinted at stellar birth. However, the finding $\sigma_z/\sigma_R > 1$ depends on a drop in the velocity dispersion in the innermost few tens of parsecs, on our assumption that the NSD is axisymmetric, and that the available (extinction corrected) stellar samples broadly trace the underlying light and mass distributions, all of which need to be established by future observations and/or modelling. (iii) We provide the most accurate rotation curve to date for the innermost 500 pc of our Galaxy.

Key words: Galaxy: centre – Galaxy: kinematics and dynamics – Galaxy: structure.

1 INTRODUCTION

The nuclear stellar disc (NSD) is a flattened stellar structure that dominates the gravitational potential of the Milky Way at Galactocentric radii $30 \lesssim R \lesssim 300$ pc (see e.g. fig. 14 in Launhardt, Zylka & Mezger 2002). Current observational constraints are consistent with the NSD being an axisymmetric structure (Gerhard & Martinez-Valpuesta 2012), although it cannot be ruled out that it actually consists of a secondary nuclear bar (Alard 2001; Rodríguez-Fernández & Combes 2008). The radius and exponential scale height determined from near-infrared photometry and star counts are $R \simeq 100\text{--}200$ pc and $H \simeq 45$ pc, respectively (Catchpole, Whitelock & Glass 1990; Launhardt et al. 2002; Nishiyama et al. 2013; Gallego-Cano et al. 2020).

The NSD is cospatial with the central molecular zone (CMZ), a ring-like accumulation of molecular gas at $R \lesssim 200$ pc, which is the Milky Way’s counterpart of the star-forming nuclear rings that are commonly found at the centre of barred galaxies (Molinari et al. 2011; Henshaw et al. 2016; Tress et al. 2020). This cospatiality is presumably not a coincidence, and suggests that the NSD is made of stars born in the dense CMZ gas (Baba & Kawata 2020; Sormani

et al. 2020). This picture is consistent with kinematic observations that show that the NSD is rotating with velocities similar to those of the molecular gas in the CMZ (Schönrich, Aumer & Sale 2015). The rotation of the NSD has been detected in the Apache Point Observatory Galactic Evolution Experiment (APOGEE) data by Schönrich et al. (2015), in OH/IR and silicon monoxide (SiO) maser stars by Lindqvist, Habing & Winnberg (1992) and Habing et al. (2006), in the Very Large Telescope (VLT)/Infrared Spectrometer and Array Camera (ISAAC) near-infrared integral-field spectroscopy by Feldmeier et al. (2014), and in classical Cepheids by Matsunaga et al. (2015).

Since the CMZ gas currently flows in the gravitational potential created by the NSD, having an accurate representation of the NSD mass and density distribution is crucial to understand gas flows in the CMZ. Hydrodynamical simulations confirm this by showing that macroscopic properties such as the size of the CMZ strongly depends on the mass and density profile of the NSD (e.g. Sormani et al. 2018; Li, Shen & Schive 2020; Tress et al. 2020). However, Bland-Hawthorn & Gerhard (2016) note that the kinematic data from Schönrich et al. (2015) suggest a mass that is on the lower side of that determined from near-infrared photometry by Launhardt et al. (2002): the former report a rotation velocity of $v \simeq 120 \text{ km s}^{-1}$ at $R \simeq 100$ pc, which naively suggests (ignoring asymmetric drift)

★ E-mail: mattia.sormani@alumni.sns.it

a mass of $M_{\text{NSD}} \simeq Rv^2/G \simeq 3 \times 10^8 M_{\odot}$, while the latter report a mass of $M_{\text{NSD}} = 6 \pm 2 \times 10^8 M_{\odot}$ at the same radius. It is thus important to constrain the NSD mass more precisely.

The mass and structure of the NSD can be constrained by constructing stellar dynamical models of the NSD and comparing them with the available kinematic/star counts data. The only attempt available in the literature is a very simple spherical Jeans modelling from Lindqvist et al. (1992) based on a sample of 148 OH/IR maser stars. However, this model neglects that the stellar density distribution of the NSD is strongly flattened (Launhardt et al. 2002; Nishiyama et al. 2013; Gallego-Cano et al. 2020) and is based on a limited number of stars.

Dynamical modelling of the NSD is also interesting from a general theoretical perspective. NSDs are common in the centre of spiral galaxies (Pizzella et al. 2002; Cole et al. 2014; Gadotti et al. 2019, 2020). The radii of nuclear rings in the sample of Gadotti et al. (2019, 2020) range from $R \sim 100$ to ~ 1000 pc, so the size of the Milky Way's NSD is consistent with but on the lower side of their distribution (see fig. 5 and table 2 in Gadotti et al. 2020). Since NSDs have different formation and evolution history than more well-studied disc systems such as galactic discs, they may be expected to have qualitatively different structural and kinematic properties.

In this paper, we aim to construct Jeans-type dynamical models of the NSD that are consistent with previous photometric/star counts studies and to compare them with line-of-sight kinematic data. This will provide constraints on the mass and structure of the NSD.

The paper is structured as follows. In Section 2, we describe the observational data. In Section 3, we describe the Jeans modelling methodology. In Section 4, we present our results, and in Section 5 we discuss them. We sum up in Section 6.

2 OBSERVATIONAL DATA

2.1 APOGEE data

We use data from the Sloan Digital Sky Survey IV (SDSS-IV)/APOGEE survey (Majewski et al. 2017) Data Release 16 (DR16; Ahumada et al. 2020), which is publicly available at <https://www.sdss.org/dr16/irspec/>. APOGEE is the first large-scale spectroscopic survey of the Milky Way in the near-infrared (H band, $1.51\text{--}1.70\ \mu\text{m}$). Most of the stars that we will use for the modelling in this paper are part of the ‘GALCEN’ field, which is a special additional target not part of the main survey targets (see section 8 in Zasowski et al. 2013). Since observations of the Galactic Centre are hampered by the extreme crowding and the high extinction and differential reddening (Nishiyama et al. 2008; Schödel et al. 2010; Nogueras-Lara et al. 2018a, 2019a, 2020a), the majority of stars that we can observe using APOGEE are bright giants (Bovy et al. 2014, 2016).

In order to diminish foreground contamination, we apply a series of cuts to the data. Our ‘standard’ filter is constructed as follows. First, we exclude all stars outside the range $|l| < 1.5$ and $|b| < 0.25$ (see red dashed box in the top panel of Fig. 1). Assuming a Sun–Galactic Centre distance of 8.2 kpc (e.g. Gravity Collaboration et al. 2019), this corresponds to projected radial and vertical distances of $|R| < 215$ pc and $|z| < 36$ pc, roughly the NSD radius and scale height (see Section 1). In this region, the surface density of the NSD is higher than that of the Galactic disc and therefore the percentage of contaminating stars are expected to be relatively low (see e.g. table 5 in Catchpole et al. 1990). A total of 405 APOGEE

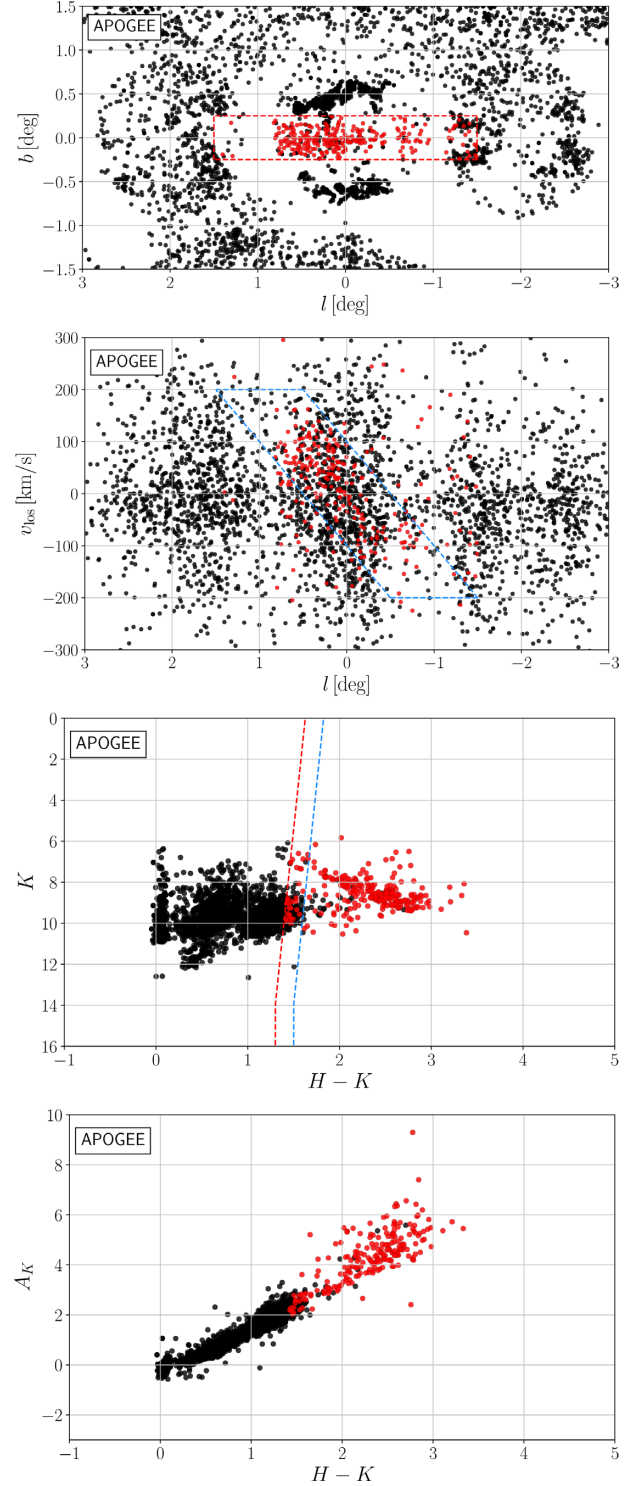


Figure 1. All APOGEE stars contained in the region $|b| < 0.25$, $|l| < 1.5$. Each point represents an individual star. Red stars are those that satisfy all the selection criteria defined in Section 2.1, while black stars are those excluded by the various cuts. The red dashed lines indicate these cuts: stars outside of the box in the top panel, or to the left of the red-dashed line in the third panel, are excluded. The blue dashed lines indicate additional cuts that we use to check the robustness of our results against variations in the selection criteria (see Section 4). v_{los} is the line-of-sight velocity, H and K are the Two Micron All Sky Survey (2MASS) H - and K -band magnitudes, and A_K is the K -band extinction from the *Wide-field Infrared Survey Explorer* (WISE) survey (Wright et al. 2010).

stars are contained in this region. We then apply a parallax cut by excluding stars that according to the APOGEE data file have $p - 3\delta_p > 1/d_{\min}$, where p is the *Gaia* Data Release 2 (DR2) parallax, δ_p is the *Gaia* DR2 parallax uncertainty, and $d_{\min} = 7$ kpc (so we remove stars that are closer than this distance). Only a small subset of stars has *Gaia* parallax defined, so this cut only removes two stars from the 405, leaving 403. Then we apply a proper motion cut by excluding stars that have $\mu_\alpha - 3\delta\mu_\alpha > \mu_{\max}$ or $\mu_\delta - 3\delta\mu_\delta > \mu_{\max}$, where μ_α and μ_δ are the *Gaia* DR2 proper motions in RA and Dec. directions, $\delta\mu_\alpha$ and $\delta\mu_\delta$ are the associated uncertainties, and $\mu_{\max} = 400/(4.74 \times 7000) \times 1000 \text{ mas yr}^{-1}$ corresponds to a proper motion velocity of 400 km s^{-1} at 7 kpc (i.e. we exclude stars that at distance of $d > 7$ kpc move faster than 400 km s^{-1}). After applying this cut, we are left with 366 stars. Finally, we apply a colour–magnitude cut. We follow the methodology explained in Nogueras-Lara et al. (2020b) and consider only stars with $H - K > \max(-0.0233K + 1.63, 1.3)$, see red dashed line in the third panel of Fig. 1. Because of the high extinction that characterizes the Galactic Centre ($A_K \sim 2.5$ mag; e.g. Nishiyama et al. 2008; Nogueras-Lara et al. 2018a, 2019a, 2020b), this colour cut effectively excludes the foreground stellar population belonging to the Galactic disc, whose absolute extinction is significantly lower, and also the majority of stars from the inner bulge ($A_K \sim 1.2$ mag, corresponding to $(H - K) \sim 1$; Nogueras-Lara et al. 2018b). The final set of stars, which consists of 273 stars, is shown in red in Fig. 1.

In order to compare the data with Jeans models, we bin the final set of stars using the VORBIN package from <https://pypi.org/user/micappe/>. This is an implementation of the two-dimensional adaptive spatial binning method of Cappellari & Copin (2003), which uses a Voronoi tessellation to bin data with given minimum signal-to-noise ratio. Here, we only use this as a convenient method to define a Voronoi tessellation that has approximately the same number of stars in each bin. The signal-to-noise ratio parameter essentially controls the average number of stars in each bin: a higher (lower) value results in less (more) bins with a higher (lower) number of stars in each of them. We assign constant signal = 1, noise = 1 and use a target signal-to-noise ratio of 3.2, which gives an average of $\simeq 10$ stars per bin. The result is shown in Fig. 2.

2.2 SiO maser data

We use the 86 GHz SiO maser survey of the inner Galaxy from Messineo et al. (2002, 2004, 2005). The SiO maser stars targeted in this survey are stars in the asymptotic giant branch (AGB) phase with estimated ages in the range 0.2–2 Gyr (e.g. Habing et al. 2006, and references therein).

Similarly to what we have done in Section 2.1 for the APOGEE data, in order to reduce foreground contamination we apply a series of cuts to the maser data. There are initially 67 maser stars in the region $|l| < 1.5$ and $|b| < 0.25$, four of which are flagged as foreground contamination by Messineo et al. (2005). After excluding these four stars, we apply the same colour–magnitude cut defined in Section 2.1 using H and K determined from the Two Micron All Sky Survey (2MASS; Skrutskie et al. 2006); see red dashed line in the third panel of Fig. 3. This excludes only one more star. The final set therefore consists of 62 stars, which are shown in red in Fig. 3.

As for the APOGEE data, we bin the final set of stars using the VORBIN package. Again we assign constant signal = 1, noise = 1 and use a target signal-to-noise ratio of 3.2, which gives an average of $\simeq 10$ stars per bin. The result is shown in Fig. 4.

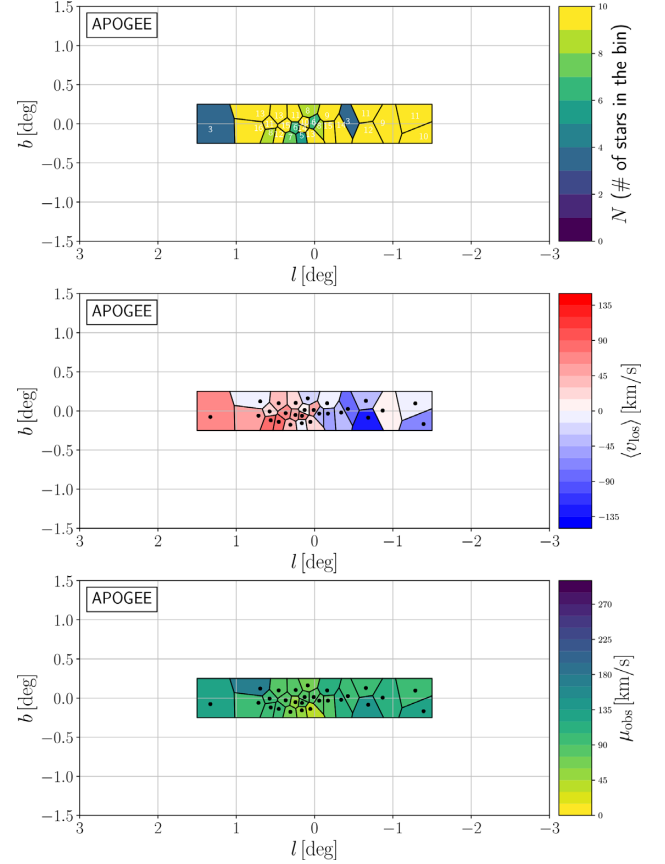


Figure 2. Voronoi binning of the APOGEE stars that satisfy all the selection criteria defined in Section 2.1 (the red stars in Fig. 1). Top: the number of stars in the bin. Middle: average line-of-sight velocity (equation 17). Bottom: root-mean-square velocity (equation 18). The latter is the quantity that we fit in our Jeans modelling.

3 JEANS MODELLING

We model the line-of-sight stellar kinematics using an anisotropic axisymmetric Jeans formalism (Cappellari 2008). Section 3.1 reviews the basic equations of this formalism. Section 3.2 describes how we compute the observables from the model, and Section 3.3 describes our fitting procedure. Section 3.4 describes the mass distribution and gravitational potential models that we employ.

3.1 Review of Jeans equations

The dynamics of a collisionless stellar system is described by the collisionless Boltzmann equation, which in cylindrical coordinates (R, ϕ, z) for an axisymmetric system reads (see equation 4–17 of Binney & Tremaine 1987)

$$\begin{aligned} \frac{\partial f}{\partial t} + v_R \frac{\partial f}{\partial R} + \frac{v_\phi}{R} \frac{\partial f}{\partial \phi} + v_z \frac{\partial f}{\partial z} + \left(\frac{v_\phi^2}{R} - \frac{\partial \Phi}{\partial R} \right) \frac{\partial f}{\partial v_R} \\ - \frac{1}{R} \left(v_R v_\phi + \frac{\partial \Phi}{\partial \phi} \right) \frac{\partial f}{\partial v_\phi} - \frac{\partial \Phi}{\partial z} \frac{\partial f}{\partial v_z} = 0, \end{aligned} \quad (1)$$

where $f(\mathbf{x}, \mathbf{v}, t)$ is the distribution function (DF), $f(\mathbf{x}, \mathbf{v}, t) d^3\mathbf{x} d^3\mathbf{v}$ is the number of stars in the small volume $d^3\mathbf{x} = R dR d\phi dz$ centred on \mathbf{x} and with velocities in the small range $d^3\mathbf{v} = dv_R dv_\phi dv_z$ centred on \mathbf{v} , and $\Phi(\mathbf{x}, t)$ is the gravitational potential. Note that, for equation (1) to be valid, it is not necessary that Φ is the potential

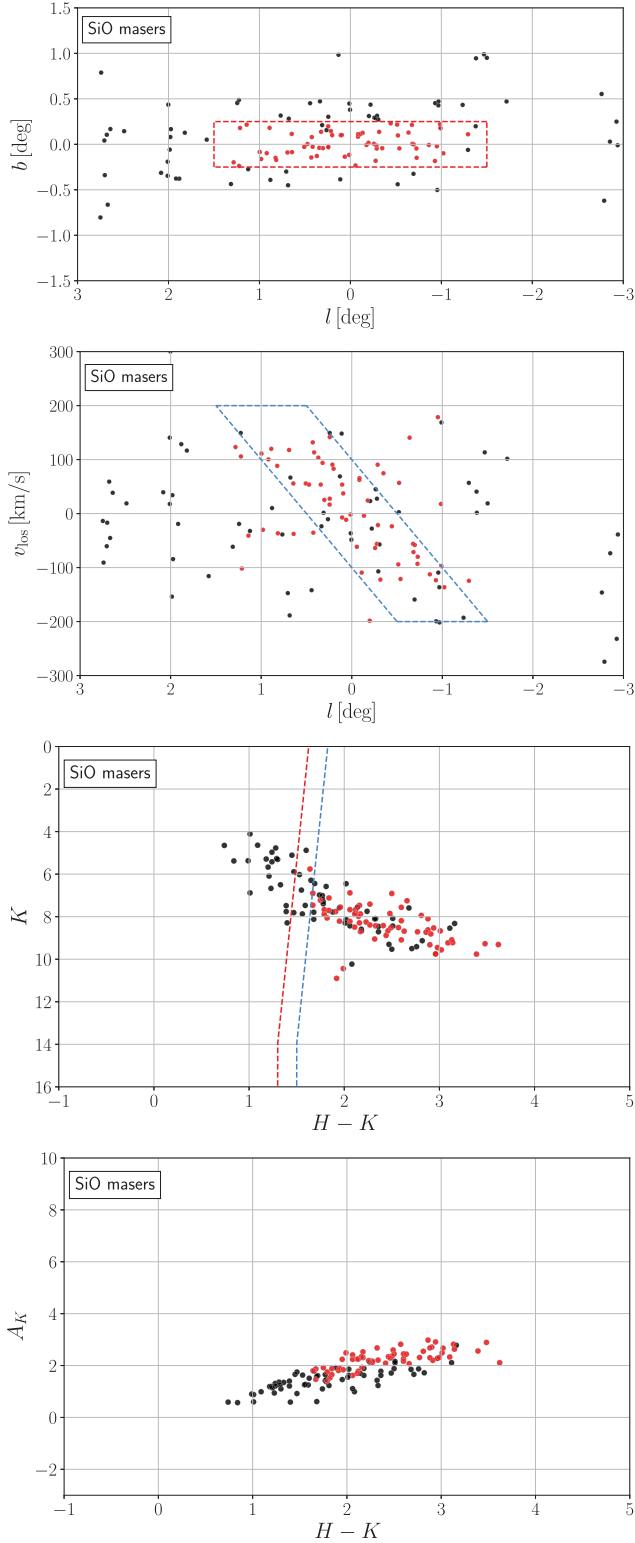


Figure 3. Same as Fig. 1, but for the SiO maser stars from Messineo et al. (2002, 2004, 2005) and using the selection criteria defined in Section 2.2.

self-consistently generated by the (tracer) density distribution calculated from f (see equation 2 below). For example, f might describe a subpopulation of stars that only partially contributes to the overall gravitational potential Φ .

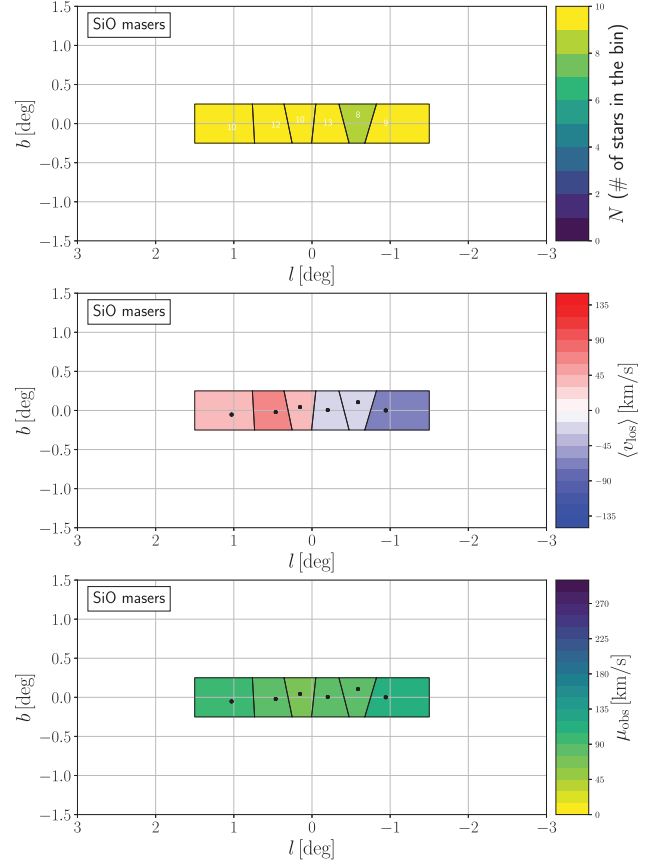


Figure 4. Same as Fig. 2, but for SiO maser stars shown in red in Fig. 3.

The spatial density of tracer stars $\rho(\mathbf{x})$, the mean velocities $\bar{v}_i(\mathbf{x})$, and the velocity ellipsoid $\sigma_{ij}(\mathbf{x})$ are defined as

$$\rho = \int f \, d^3\mathbf{v}, \quad (2)$$

$$\bar{v}_i = \frac{1}{\rho} \int f v_i \, d^3\mathbf{v}, \quad (3)$$

$$\bar{v}_i \bar{v}_j = \frac{1}{\rho} \int f v_i v_j \, d^3\mathbf{v}, \quad (4)$$

$$\sigma_{ij}^2 = (\overline{v_i - \bar{v}_i})(\overline{v_j - \bar{v}_j}) = \bar{v}_i \bar{v}_j - \bar{v}_i \bar{v}_j, \quad (5)$$

where $i, j = R, \phi, \text{ or } z$. Multiplying equation (1) by v_R , v_ϕ , or v_z respectively, assuming axisymmetry ($\partial_\phi = 0$), and integrating over all velocities, we obtain the following Jeans equations¹ (see equations 4-29a-4-29c in Binney & Tremaine 1987):

$$\frac{\partial(\rho \bar{v}_R)}{\partial t} + \frac{\partial(\rho \bar{v}_R^2)}{\partial R} + \frac{\partial(\rho \bar{v}_R \bar{v}_z)}{\partial z} + \rho \left(\frac{\bar{v}_R^2 - \bar{v}_\phi^2}{R} + \frac{\partial \Phi}{\partial R} \right) = 0, \quad (6)$$

$$\frac{\partial(\rho \bar{v}_\phi)}{\partial t} + \frac{\partial(\rho \bar{v}_R \bar{v}_\phi)}{\partial R} + \frac{\partial(\rho \bar{v}_\phi \bar{v}_z)}{\partial z} + \frac{2\rho}{R} \bar{v}_\phi \bar{v}_R = 0, \quad (7)$$

$$\frac{\partial(\rho \bar{v}_z)}{\partial t} + \frac{\partial(\rho \bar{v}_R \bar{v}_z)}{\partial R} + \frac{\partial(\rho \bar{v}_z^2)}{\partial z} + \frac{\rho}{R} \bar{v}_R \bar{v}_z + \rho \frac{\partial \Phi}{\partial z} = 0. \quad (8)$$

¹The steps involve integrating some terms by parts and assuming that $f \rightarrow 0$ for $|\mathbf{v}| \rightarrow \infty$.

The typical situation in Jeans modelling is one in which given ρ and Φ , and under the assumption of steady state ($\partial_t = 0$), we want to use equations (6)–(8) to generate predictions for the six velocity moments ($\overline{v_R^2}$, $\overline{v_\phi^2}$, $\overline{v_z^2}$, $\overline{v_R v_\phi}$, $\overline{v_R v_z}$, and $\overline{v_\phi v_z}$) that can be compared with kinematic observations. However, for a steady state system ($\partial_t = 0$), equations (6)–(8) provide only three constraints among these six moments. Therefore, in order to proceed, one has to make some assumptions that reduce the number of unknowns to match the number of equations. Following Cappellari (2008) we assume that

- (i) $\overline{v_R} = \overline{v_z} = 0$, i.e. any mean streaming motion within the disc is purely tangential;
- (ii) $\overline{v_R v_\phi} = \overline{v_z v_\phi} = 0$;
- (iii) $\sigma_{Rz} = \overline{v_R v_z} = 0$, i.e. the principal axes of the velocity ellipsoid σ_{ij} (which can always be diagonalized since it is a symmetric tensor) are parallel to the R and z axes;
- (iv) $\overline{v_R^2} = b \overline{v_z^2}$, where the anisotropy b is a constant.

Assumptions (i) and (ii) are in the spirit of our assumption that the disc is axisymmetric. Assumption (iii) is stronger than (ii), and does not have such a natural justification. It is reasonable to assume that $\overline{v_R v_z} = 0$ in the $z = 0$ plane, since this follows if we assume reflection symmetry with respect to the plane $z = 0$. In the solar neighbourhood, as one moves away from the $z = 0$ plane the velocity ellipsoid ‘tilts’ in the sense that it is more closely aligned with the r and θ axes of a spherical polar coordinate system (Siebert et al. 2008; Binney et al. 2014; Everall et al. 2019). Nevertheless, even if the velocity ellipsoid does tilt like this, then (using the standard rules for the transformation of tensors under rotations) we would have $\overline{v_R v_z} = (\sigma_r^2 - \sigma_\theta^2) \sin \theta \cos \theta$, which is much smaller than the other terms in equations (6) and (8) when one is close to the plane ($\theta = \pi/2$). We have tested that assuming that the principal axes are aligned on spherical rather than cylindrical coordinates does not affect the conclusions of the paper (see Section 5.2 for more details). Assumption (iv) is harder to justify a priori and is mainly motivated by simplicity. We will see in Section 4 that it gives an adequate representation of the available data.

Under these assumptions, equation (7) is identically zero, while equations (6) and (8) become

$$\frac{\partial (\rho b \overline{v_z^2})}{\partial R} + \rho \left(\frac{b \overline{v_z^2} - \overline{v_\phi^2}}{R} + \frac{\partial \Phi}{\partial R} \right) = 0, \quad (9)$$

$$\frac{\partial (\rho \overline{v_z^2})}{\partial z} + \rho \frac{\partial \Phi}{\partial z} = 0. \quad (10)$$

These two equations can be solved in the two unknowns $\overline{v_z^2}$ and $\overline{v_\phi^2}$. Integrating equation (10) using the boundary condition $\rho \overline{v_z^2} \rightarrow 0$ as $z \rightarrow \infty$ and then substituting in equation (9) we obtain

$$\rho \overline{v_z^2}(R, z) = \int_z^\infty \rho \frac{\partial \Phi}{\partial z} dz, \quad (11)$$

$$\rho \overline{v_\phi^2}(R, z) = b \left[R \frac{\partial (\rho \overline{v_z^2})}{\partial R} + \rho \overline{v_z^2} \right] + R \rho \frac{\partial \Phi}{\partial R}. \quad (12)$$

These equations allow one to generate predictions for $\overline{v_z^2}(R, z)$ and $\overline{v_\phi^2}(R, z)$ given $\rho(R, z)$, $\Phi(R, z)$, and the parameter b . In Section 3.2, we show that it is straightforward to project these quantities along lines of sight and to compare the results against the (density-weighted) projected second moment constructed from the observed stellar samples. We stress however that equations (11) and (12) rely on the simplifying and somewhat arbitrary assumptions (iii) and (iv)

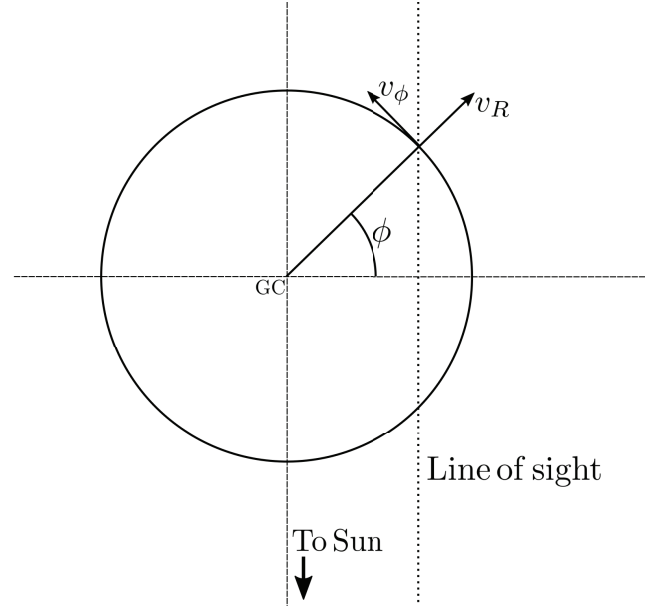


Figure 5. Geometry of the line-of-sight integration (see Section 3.2).

above. One of the biggest shortcomings of Jeans modelling is that even if a tracer density model ρ and a gravitational potential Φ are found such that the moments calculated using equations (11) and (12) project to give a good representation of the data, this does not guarantee that this model is physical: it may not exist a well-defined DF ($f > 0$) in the potential Φ that corresponds to the density ρ and that satisfies all the assumptions made in this section (steady state, axisymmetry, i–iv; see e.g. section 4.4.1 in Binney & Tremaine 2008).

3.2 Calculation of observables

Equations (11) and (12) allow one to calculate predictions for $\overline{v_z^2}(R, z)$ and $\overline{v_\phi^2}(R, z)$. However, we do not have direct observations of these two quantities for the NSD. In order to calculate the observables that can be compared to our data, we need to integrate them along the line of sight.

We assume that the NSD is exactly edge-on and that all lines of sight can be considered parallel at the distance of the Galactic Centre. Under these assumptions the line-of-sight velocity is given by (see Fig. 5)

$$v_{\text{los}}(R, z) = v_\phi(R, z) \cos \phi + v_R(R, z) \sin \phi. \quad (13)$$

Taking the square of this equation and then averaging² we find (see e.g. appendix A of Evans & de Zeeuw 1994)

$$\overline{v_{\text{los}}^2}(R, z) = \overline{v_\phi^2}(R, z) \cos^2 \phi + \overline{v_R^2}(R, z) \sin^2 \phi, \quad (14)$$

where we have used that $\overline{v_\phi v_R} = 0$ (see Section 3.1). The second moment of the line-of-sight velocity is obtained by integrating equation (14) along the line of sight weighting by density:³

$$\Sigma \mu_{\text{los}}^2 = \int_{-\infty}^{\infty} \rho (\overline{v_\phi^2} \cos^2 \phi + \overline{v_R^2} \sin^2 \phi) ds, \quad (15)$$

²The average of a generic quantity $G(x, v)$ are defined here as $\bar{G}(R, z) = \int [f G d^3 v] / \rho$, where $\rho = \int f d^3 v$.

³Note that we use the same Greek letter Σ to denote both surface density and the summation symbol. The two can be distinguished since the latter is always accompanied by an index of summation (i or j), while the former never is.

where s indicates the distance along the line of sight and the surface density is defined as

$$\Sigma = \int_{-\infty}^{\infty} \rho \, ds. \quad (16)$$

3.3 Fitting procedure

We calculate the likelihood of a model as follows.

(i) For each bin j (see Figs 2 and 4), we calculate the mean line-of-sight velocity and the mean square line-of-sight velocity from the observed sample as

$$\langle v_{\text{los,obs}} \rangle_j = \frac{1}{N_j} \sum_{i=0}^{N_j} v_{\text{los,obs},i}, \quad (17)$$

$$\mu_{\text{obs},j}^2 \equiv \langle v_{\text{los,obs}}^2 \rangle_j = \frac{1}{N_j} \sum_{i=0}^{N_j} v_{\text{los,obs},i}^2, \quad (18)$$

where the sum is extended over all stars contained in the bin, the index j labels the bins, the index i labels individual stars, $v_{\text{los,obs},i}$ is the observed line-of-sight velocity of the star i and N_j is the number of stars in the bin. We use the notation $\langle \cdot \rangle_j = \sum_i \cdot / N_j$ to denote averages over the bin j , while we reserve the overline symbol $\bar{\cdot}$ to denote averages over the DF (e.g. equation 3). The values of N_j , $\langle v_{\text{los,obs}} \rangle_j$ and $\mu_{\text{obs},j}^2$ calculated in this way are shown in Figs 2 and 4 for the APOGEE stars and SiO maser stars, respectively.

(ii) For each bin j , we calculate the second moment of the line-of-sight velocity $\mu_{\text{los},i}^2$ at the on-sky position of each individual star i within the bin by performing the integrals in equations (15) and (16). Then we average these over the bin,

$$\mu_{\text{model},j}^2 \equiv \langle \mu_{\text{los}}^2 \rangle_j = \frac{1}{N_j} \sum_{i=0}^{N_j} \mu_{\text{los},i}^2. \quad (19)$$

Given the small number of stars in each bin, the quantity (19) is a reasonable proxy of the observed quantity (18).

(iii) We assume that the estimates (18) are normally distributed about their true values. Then the likelihood of a model is

$$P = \exp(-\chi^2/2), \quad (20)$$

where

$$\chi^2 = \sum_j \left[\frac{\mu_{\text{obs},j} - \mu_{\text{model},j}}{\Delta \mu_j} \right]^2, \quad (21)$$

where the sum is extended over all bins and $\Delta \mu_j$ is the error on $\mu_{\text{obs},j}$, which we estimate as

$$\Delta \mu_j = \frac{\mu_{\text{obs},j}}{\sqrt{N_j}}. \quad (22)$$

3.4 Gravitational potential and density distribution

The Jeans equations (11) and (12) require assuming a gravitational potential $\Phi(R, z)$ and a density distribution $\rho(R, z)$ in order to generate predictions for $\bar{v}_z^2(R, z)$ and $\bar{v}_\phi^2(R, z)$. Note that, as mentioned in Section 3.1, for equations (11) and (12) to be valid it is not necessary that $\Phi(R, z)$ is the potential self-consistently generated by $\rho(R, z)$. We will consider both models in which $\Phi(R, z)$ is generated by $\rho(R, z)$ and models in which it is not.

3.4.1 Gravitational potential

At the range of Galactocentric radii considered in this paper only two components contribute significantly to the potential: the NSD, which dominates the potential at $30 \lesssim R \lesssim 300$ pc, and the nuclear stellar cluster (NSC), which dominates the potential at $1 \lesssim R \lesssim 30$ pc (e.g. Launhardt et al. 2002; Schödel et al. 2014; Gallego-Cano et al. 2020). Therefore we take a gravitational potential of the following form:

$$\Phi(R, z) = \alpha \Phi_{\text{NSD}}(R, z) + \beta \Phi_{\text{NSC}}(R, z), \quad (23)$$

where parameter α is the mass scaling of the NSD and will be left as a free parameter in our fitting procedure below. The value $\alpha = 1$ will correspond to the normalizations of ρ_{NSD} as given below in this section. The parameter β is the mass scaling of the NSC, which we keep fixed in all our fitting procedures. We will consider models with ($\beta = 1$) and without ($\beta = 0$) the NSC.

To calculate the potential $\Phi_{\text{NSD}}(R, z)$ we consider three different NSD models, the diversity of which reflects the current large uncertainties in the mass distribution of the NSD. The first is the best-fitting model from Launhardt et al. (2002) (see their section 5.2; see also equation 1 in Li et al. 2020):

$$\rho_{\text{NSD}}(R, z) = \rho_1 \exp \left\{ -\log(2) \left[\left(\frac{R}{R_1} \right)^{n_R} + \left(\frac{|z|}{z_0} \right)^{n_z} \right] \right\} \\ + \rho_2 \exp \left\{ -\log(2) \left[\left(\frac{R}{R_2} \right)^{n_R} + \left(\frac{|z|}{z_0} \right)^{n_z} \right] \right\}, \quad (24)$$

where $n_R = 5$, $n_z = 1.4$, $R_1 = 120$ pc, $R_2 = 220$ pc, $z_0 = 45$ pc, $\rho_1/\rho_2 = 3.9$, $\rho_1 = 15.2 \times 10^{10} \text{ M}_\odot \text{ kpc}^{-3}$, and $\log(2) \simeq 0.693$. Launhardt et al. (2002) showed that this model fits well the *Cosmic Background Explorer* (COBE) 4.9 μm emission. The vertical scale height z_0 has been independently confirmed from star counts by Nishiyama et al. (2013).

The second NSD model we consider is the best-fitting axisymmetric model of Chatzopoulos et al. (2015), which has a density distribution given by (see their equation 17)

$$\rho_{\text{NSD}}(R, z) = \frac{(3 - \gamma)M}{4\pi q} \frac{a_0}{a^\gamma (a + a_0)^{4-\gamma}}, \quad (25)$$

where

$$a(R, z) = \sqrt{R^2 + \frac{z^2}{q^2}}, \quad (26)$$

and $\gamma = 0.07$, $q = 0.28$, $a_0 = 182$ pc, and $M = 6.2 \times 10^9 \text{ M}_\odot$. Note that this is only the second component from equation (17) of Chatzopoulos et al. (2015), while the first component represents the NSC (see below).

The third NSD model we consider is obtained by deprojecting model 2 of Gallego-Cano et al. (2020) (see their equation 3 and their table 4). These authors have fitted a Sérsic profile to the *Spitzer*/Infrared Array Camera (IRAC) 4.5 μm stellar surface density maps of the central $300 \times 250 \text{ pc}^2$. Their models give a projected density $\Sigma(R, z)$ that can be deprojected to obtain the 3D density distribution $\rho(R, z)$. For an edge-on system, this deprojection is unique and can be done using the Abel transform (see e.g. appendix A in Mamon & Boué 2010). The following analytical density distribution gives an excellent approximation to the unique deprojected density distribution:

$$\rho_{\text{NSD}}(R, z) = \rho_1 \exp \left[-\left(\frac{a}{R_1} \right)^{n_1} \right] + \rho_2 \exp \left[-\left(\frac{a}{R_2} \right)^{n_2} \right], \quad (27)$$

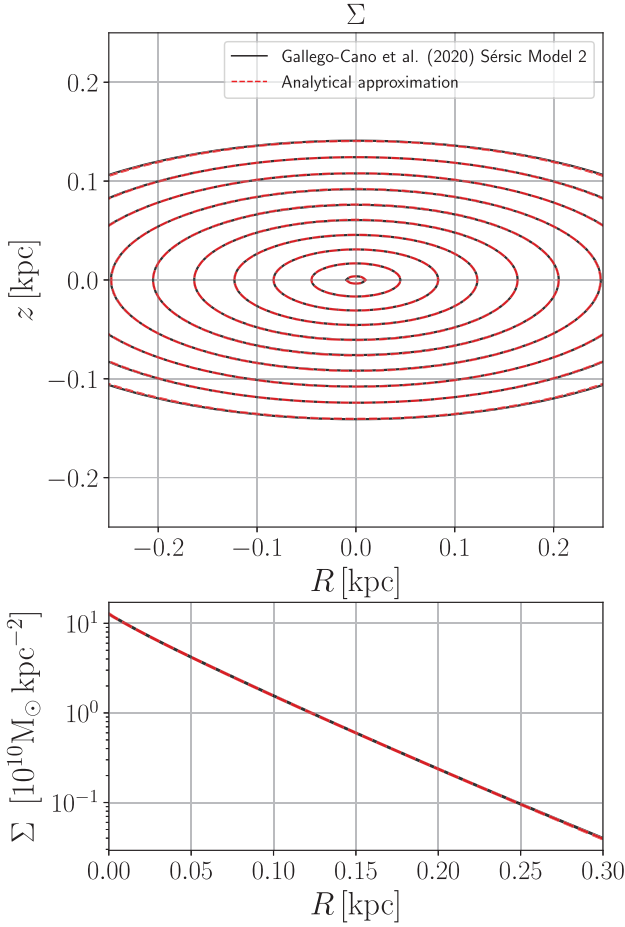


Figure 6. Comparison between Sérsic model 2 of Gallego-Cano et al. (2020) (full black lines) and the analytical approximation given by equation (27) (dashed red lines). Top panel: contours of surface density $\Sigma(R, z)$. Bottom panel: surface density radial profile in the plane $z = 0$. The normalization of both models is arbitrarily chosen so that $\Sigma = 2 \times 10^{10} \text{ M}_\odot \text{ kpc}^{-3}$ at the centre.

where $a(R, z)$ is defined as in equation (26), $q = 0.37$, $n_1 = 0.72$, $n_2 = 0.79$, $R_1 = 5.06 \text{ pc}$, $R_2 = 24.6 \text{ pc}$, $\rho_1/\rho_2 = 1.311$, and $\rho_2 = 170 \times 10^{10} \text{ M}_\odot \text{ kpc}^{-3}$. Since Gallego-Cano et al. (2020) normalize their model using observed intensity and not surface density, we choose the arbitrary normalization ρ_2 by requiring that the surface density is $\Sigma = 2 \times 10^{10} \text{ M}_\odot \text{ kpc}^{-3}$ at the centre. The normalization with respect to this value, quantified by the parameter α , will be determined by the fitting procedure in Section 4. Fig. 6 shows that there is excellent agreement between the surface density of model 2 of Gallego-Cano et al. (2020) and that obtained with equation (27).

Fig. 7 compares the three NSD models described above. It can be seen that they have rather different density contours. Since the Gallego-Cano et al. (2020) NSD is obtained using data at much higher resolution than those of Launhardt et al. (2002) and Chatzopoulos et al. (2015), we consider it is the most accurate of the three. We will see in Section 4 that the main results of this paper are not affected by the choice of the NSD model.

To calculate the potential $\Phi_{\text{NSC}}(R, z)$ generated by the NSC, we adopt the mass density of the best-fitting axisymmetric model from Chatzopoulos et al. (2015) (see their equation 17):

$$\rho_{\text{NSC}}(R, z) = \frac{(3 - \gamma)M_{\text{NSC}}}{4\pi q} \frac{a_0}{a^\gamma (a + a_0)^{4-\gamma}}, \quad (28)$$

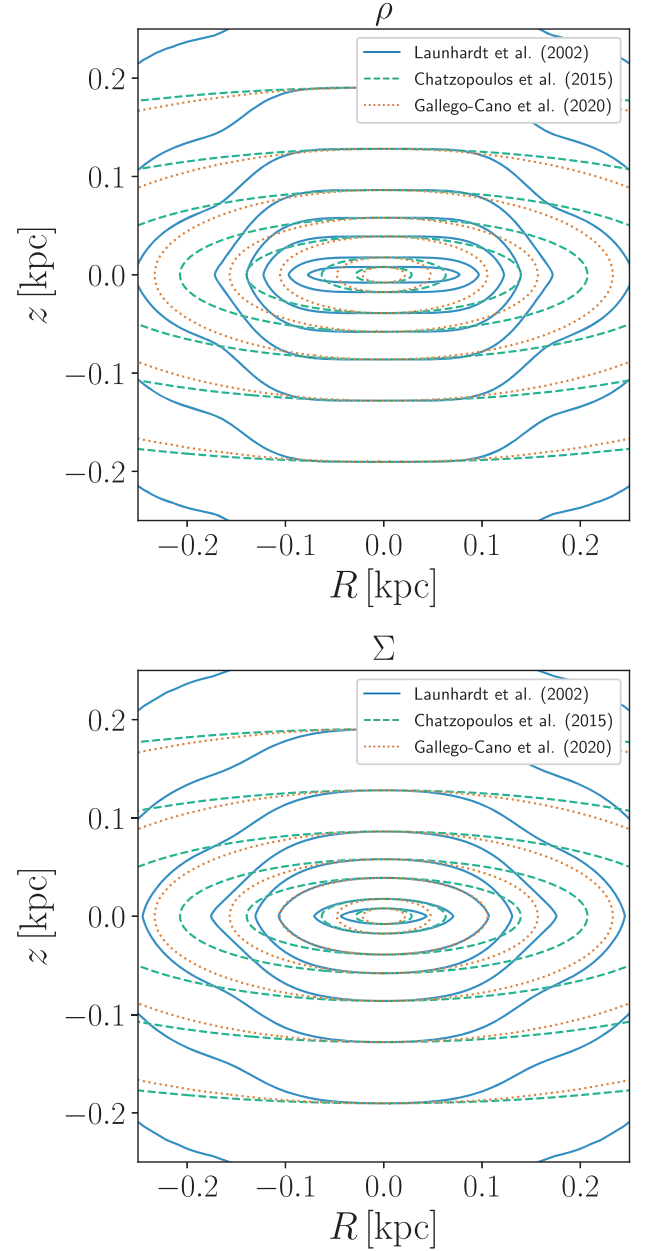


Figure 7. Comparison of the three NSD models considered in this paper (see Section 3.4.1). Top: density $\rho(R, z)$. Bottom: surface density $\Sigma(R, z)$.

where

$$a(R, z) = \sqrt{R^2 + \frac{z^2}{q^2}}, \quad (29)$$

and $\gamma = 0.71$, $q = 0.73$, $a_0 = 5.9 \text{ pc}$, and $M_{\text{NSC}} = 6.1 \times 10^7 \text{ M}_\odot$. This corresponds to the first component from equation (17) of Chatzopoulos et al. (2015), while the second component corresponds to the NSD as mentioned above.

3.4.2 Tracer density distribution

For the density distribution $\rho(R, z)$ we consider two cases.

Table 1. Best-fitting parameters for all the models considered in this paper. Each row corresponds to a panel in Figs 8, A1, and A2. Columns are defined as follows. Filter: selection criteria used to filter the data. ‘Standard’ denotes the criteria as described in Sections 2.1 and 2.2 (see red dashed lines in Figs 1 and 3). ‘Restrictive’ is the same as standard, but using a more restrictive colour–magnitude cut that is shifted by 0.2 mag (see blue lines in the third panels of Figs 1 and 3). ‘(*l*, *v*) cut’ is the same as standard, but with an additional cut that excludes all stars outside the parallelogram shown in blue dashed lines in the second panels of Figs 1 and 3. $\Phi(R, z)$: employed gravitational potential. ‘NSD+NSC’ and ‘NSD only’ mean that the potential is calculated using equation (23) with fixed $\beta = 1$ and $\beta = 0$, respectively. $\rho(R, z)$: employed tracer density distribution, which can be either the same as the one that generates the potential $\Phi(R, z)$ (equation 30) or the ring distribution (equation 31). NSD: mass model used to calculate Φ_{NSD} (see Section 3.4.1). α : best-fitting mass scaling of the NSD relative to the normalization as given in Section 3.4.1. b : best-fitting anisotropy parameter. χ^2 : value for the best-fitting model. $M(r < 100 \text{ pc})$: mass contained within spherical radius $r = 100 \text{ pc}$ in units of $10^8 M_\odot$.

Model	Fitted to	Filter	$\Phi(R, z)$	$\rho(R, z)$	NSD	α	b	χ^2	$M(r < 100 \text{ pc})$
1	APOGEE	Standard	NSD+NSC	NSD+NSC	Launhardt et al. (2002)	0.8	0.475	11.74	4.3
2	APOGEE	Standard	NSD+NSC	NSD+NSC	Chatzopoulos et al. (2015)	0.875	0.45	10.79	5.0
3 (fiducial)	APOGEE	Standard	NSD+NSC	NSD+NSC	Gallego-Cano et al. (2020)	0.9	0.4	10.73	3.9
4	SiO masers	Standard	NSD+NSC	NSD+NSC	Launhardt et al. (2002)	0.675	0.8	0.82	3.7
5	SiO masers	Standard	NSD+NSC	NSD+NSC	Chatzopoulos et al. (2015)	0.675	0.925	0.74	4.0
6	SiO masers	Standard	NSD+NSC	NSD+NSC	Gallego-Cano et al. (2020)	0.85	0.725	0.80	3.7
7	APOGEE	Restrictive	NSD+NSC	NSD+NSC	Launhardt et al. (2002)	0.85	0.425	11.91	4.5
8	APOGEE	Restrictive	NSD+NSC	NSD+NSC	Chatzopoulos et al. (2015)	0.925	0.375	11.84	5.2
9	APOGEE	Restrictive	NSD+NSC	NSD+NSC	Gallego-Cano et al. (2020)	0.875	0.4	12.79	3.8
10	APOGEE	(<i>l</i> , <i>v</i>) cut	NSD+NSC	NSD+NSC	Launhardt et al. (2002)	0.75	0.225	4.19	4.0
11	APOGEE	(<i>l</i> , <i>v</i>) cut	NSD+NSC	NSD+NSC	Chatzopoulos et al. (2015)	0.85	0.175	2.42	4.9
12	APOGEE	(<i>l</i> , <i>v</i>) cut	NSD+NSC	NSD+NSC	Gallego-Cano et al. (2020)	0.675	0.125	1.83	3.1
13	APOGEE	Standard	NSD only	NSD only	Launhardt et al. (2002)	0.925	0.6	13.22	4.3
14	APOGEE	Standard	NSD only	NSD only	Chatzopoulos et al. (2015)	0.975	0.625	12.71	4.9
15	APOGEE	Standard	NSD only	NSD only	Gallego-Cano et al. (2020)	1.175	0.65	11.80	4.4
16	APOGEE	Standard	NSD+NSC	Ring	Launhardt et al. (2002)	0.95	0.825	13.47	5.0
17	APOGEE	Standard	NSD+NSC	Ring	Chatzopoulos et al. (2015)	0.875	0.625	13.37	5.0
18	APOGEE	Standard	NSD+NSC	Ring	Gallego-Cano et al. (2020)	1.45	0.625	13.04	6.0

(i) The stellar populations traced by our data (Section 2) are distributed in the same way as the stars that make up most of the mass of the NSD/NSC. In this case, we take for $\rho(R, z)$ in equations (11) and (12) the density distribution that generates the potential $\Phi(R, z)$ given in equation (23), i.e.

$$\rho(R, z) = \alpha \rho_{\text{NSD}}(R, z) + \beta \rho_{\text{NSC}}(R, z). \quad (30)$$

(ii) The stellar populations traced by our data (Section 2) are distributed differently than stars that make up most of the mass of the NSD/NSC. Indeed, the selection function of APOGEE favours younger populations of order $\sim 1 \text{ Gyr}$ of age (e.g. fig. 1 of Aumer & Schönrich 2015), while the SiO maser stars have estimated ages of 0.2–2 Gyr (Habing et al. 2006), which may be distributed differently than the older stars that are believed to make up most of the NSD mass (Nogueras-Lara et al. 2020b). Since the gas in the CMZ is believed to have a ring-like morphology (e.g. Molinari et al. 2011; Kruijssen, Dale & Longmore 2015; Sormani et al. 2018; Tress et al. 2020), and that the distribution of relatively young stars traced by our data might still reflect the gas distribution, we consider a ring-like stellar density distribution given by

$$\rho(R, z) = \rho_0 \exp \left[-\eta \left(\frac{R_0}{R} + \frac{R}{R_0} \right) - \frac{|z|}{z_0} \right], \quad (31)$$

where $R_0 = 100 \text{ pc}$ is the radius at which the density is maximum, $\eta = 1$ is a parameter that controls the width of the ring, and for the vertical scale height we take $z_0 = 45 \text{ pc}$ as in equation (24).

4 RESULTS

All the models considered in this paper are listed in Table 1. Fig. 8 shows the probability distributions for models 1–6. The left-hand, middle, and right-hand columns differ for the NSD models employed,

and correspond to those of Launhardt et al. (2002), Chatzopoulos et al. (2015), and Gallego-Cano et al. (2020), respectively. The top row corresponds to models that are fitted to APOGEE data, while the bottom row corresponds to models that are fitted to SiO maser data.

The top row in Fig. 8 shows that although the three NSD models have rather different density distributions (see Fig. 7), they all give similar values for $M(r < 100 \text{ pc})$ and for the anisotropy parameter b when fitted to APOGEE data (Table 1). The best-fitting value for our (fiducial) model 3 is $\alpha = 0.9 \pm 0.2$ that corresponds to a mass $M(r < 100 \text{ pc}) = (3.9 \pm 1) \times 10^8 M_\odot$ and a total NSD mass of $M_{\text{NSD}} = (6.9 \pm 2) \times 10^8 M_\odot$. The anisotropy parameter is consistently $b \sim 0.5$ for all models. Fitting the same model to SiO masers (bottom left-hand panel) gives results that are consistent with the fit to APOGEE data, but with significantly larger uncertainty (as is expected given the smaller number of stars in the SiO masers data).

To assess the impact of our data selection criteria, we follow a strategy similar to that of Nogueras-Lara et al. (2020b) and repeat the fits using different cuts. Models 7–9 are identical to models 1–3 except that we use a more restrictive colour–magnitude cut that is shifted by 0.2 mag with respect to the standard cut (see blue dashed line in Fig. 1). This excludes 30 additional APOGEE stars from the sample, leaving 243. Table 1 and Fig. A1 show that this does not affect the results significantly.

The second panel in Fig. 1 displays several stars at negative (positive) longitude that have large positive (negative) line-of-sight velocities and therefore naively appear to be counter-rotating. Such stars are most likely stars on elongated x_1 -like orbits that belong to the Galactic bar (Aumer & Schönrich 2015; Molloy et al. 2015), and indeed occupy the same area in the (*l*, *v*) plane as the so-called ‘forbidden velocity’ gas, which has a similar interpretation (Binney et al. 1991; Fux 1999; Sormani, Binney & Magorrian 2015b). Also visible in the second panel of Fig. 1 are stars with very high

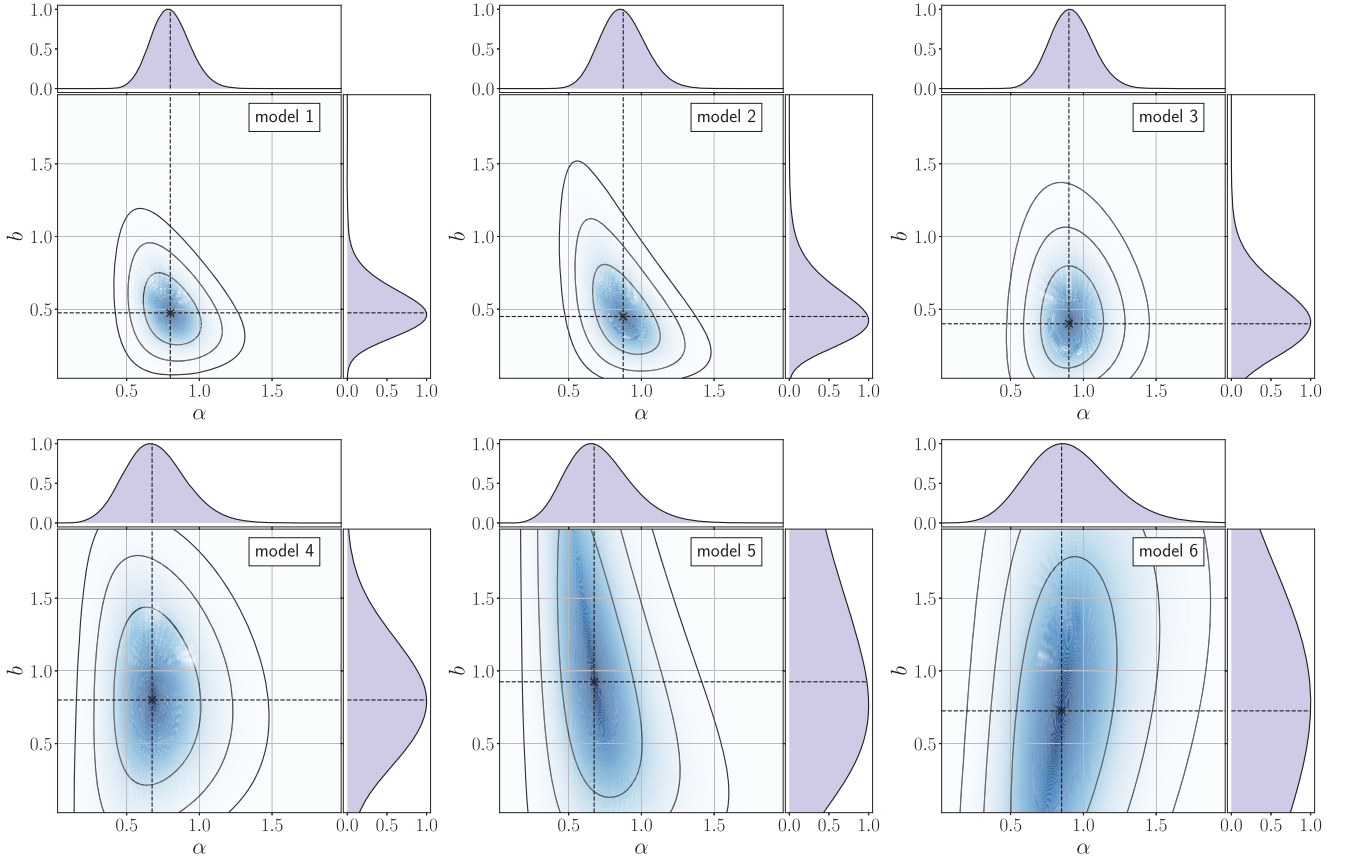


Figure 8. Probability distributions $P(\alpha, b)$ calculated using equation (20) for models 1–6 (see Table 1). Contours show the 1σ , 2σ , and 3σ contours that contain 68, 95, and 99.7 per cent of the total probability, respectively. The parameter α is the mass normalization of the NSD relative to the normalizations as given in Section 3.4.1. The parameter $b = \sigma_R^2/\sigma_z^2$ is the anisotropy parameter introduced by Cappellari (2008). The lateral panels represent the marginalized distributions $P(\alpha) = \int P(\alpha, b) db$ and $P(b) = \int P(\alpha, b) d\alpha$. The crosses mark the maximum of P in the 2D distributions, which correspond to the values reported in Table 1.

line-of-sight velocities ($v_{\text{los}} \geq 200 \text{ km s}^{-1}$), which are also most likely stars on x_1 -type bar orbits (Molloy et al. 2015; Habing 2016) and also have a gas counterpart as ‘high-velocity peaks’ in the (l, v) plane (Binney et al. 1991; Sormani et al. 2015b). In order to assess the potential impact of such contamination from the Galactic bar, models 10–12 repeat the fits excluding all the stars outside the blue parallelogram in the second panel of Fig. 1. This removes 54 APOGEE stars from the sample, leaving 219. Table 1 and Fig. A1 show that this does not affect the mass normalization significantly, but it tends to give even lower values for the anisotropy parameter b .

To assess the impact of including the NSC component, which is important only for $R \lesssim 30 \text{ pc}$ ($|l| \lesssim 0.2$), we now consider models that only include the NSD potential and density. Models 13–15 are identical to models 1–3 except that we exclude the NSC by setting its normalization to $\beta = 0$ (see equations 23 and 30). Table 1 and Fig. A2 show that this favours a slightly larger mass and anisotropy parameter b than the NSD+NSC models, but are consistent within the uncertainties. Thus, the inclusion of the NSC does not affect the results significantly, which is reasonable given the small number of data points at $|l| \lesssim 0.2$ (Figs 2 and 4). The best-fitting NSD-only model fits the data comparably well as the best NSD+NSC model. This confirms that our approach to keep the NSC mass fixed to the value determined by Chatzopoulos et al. (2015) in our fitting procedure is reasonable. The inclusion of the NSC will be more important when better data will be available in the future.

As mentioned in Section 3.4.2, the stellar populations traced by our data might be distributed differently than stars that make up most of

the mass of the NSD/NSC. In particular, given that both the selection functions of APOGEE and SiO maser stars tend to favour relatively young stellar populations with ages $\sim 1\text{--}2 \text{ Gyr}$ (see references in Section 3.4.2), we consider a ring-like density distribution that might reflect more closely the gas distribution in the CMZ. Table 1 and the bottom row of Fig. A2 show the result of fitting the ring models, which have the same potential as the NSD+NSC models but a ring-like density distribution given by equation (31), to APOGEE data. As for the NSD-only models, the ring models favour a slightly larger mass scaling parameter α and an anisotropy a bit higher and closer to $b \sim 1$. We continue the discussion on this point in Section 5.2.

Fig. 9 compares our fiducial model 3 to both APOGEE and SiO maser data. The model and data show a reasonably good agreement given the quality of the data. Models 1 and 2, which employ a different NSD mass distribution, offer comparably good representations of the data. A similar consideration applies to the NSD-only models. This makes clear that the limiting factor in our analysis is the quality of the data, and not the assumed potential/density distribution. Trying to refine the potential/density distribution would not make sense until better data become available.

5 DISCUSSION

5.1 The mass of the nuclear stellar disc

We have seen in Section 4 that our models favour a mass $M(r < 100 \text{ pc}) = (4 \pm 1) \times 10^8 M_\odot$ that is consistent with but lower than the

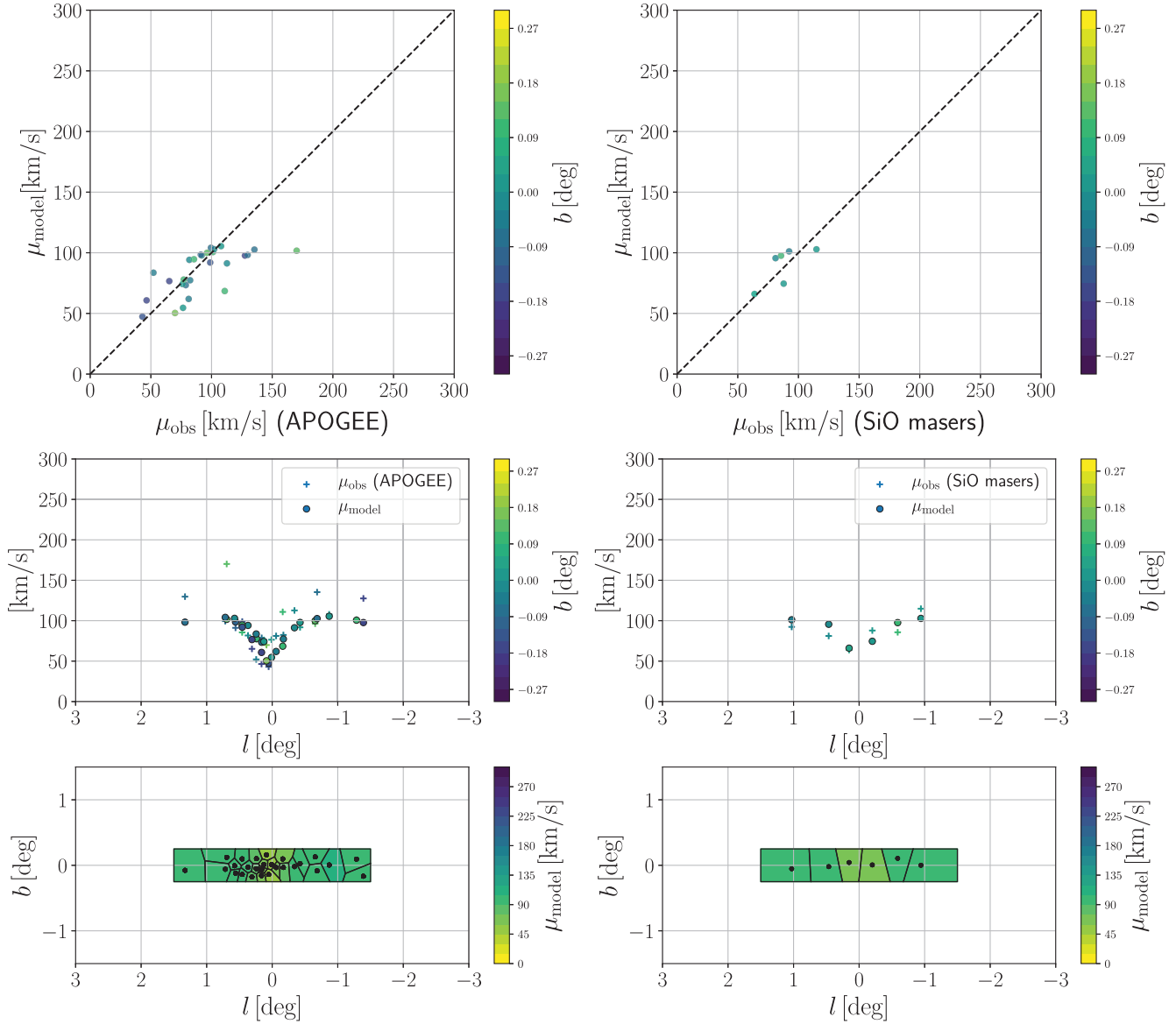


Figure 9. Comparison between observational data and model 3 (see Table 1). Left: APOGEE data. Right: SiO maser data. Top: observed second moment of the line-of-sight velocity (equation 18) versus predictions (equation 19). Each point on the left-hand (right-hand) panel represents one of the bins shown in Fig. 2 (Fig. 4). Middle: second moment of the line-of-sight velocity as a function of longitude for observational data (crosses) and best-fitting model (circles). Bottom: predictions for the second moment of the line-of-sight velocity (compare with bottom panels in Figs 2 and 4).

best-fitting value $M(r < 100 \text{ pc}) \simeq (6 \pm 2) \times 10^8 M_\odot$ of Launhardt et al. (2002). The two determinations are to a large extent independent since ours is based on the line-of-sight kinematics while the one of Launhardt et al. (2002) is purely based on the photometry.

As mentioned in Section 1, the size of the CMZ in simulations of gas flow in Milky Way-like barred potentials depends on the mass of the NSD. Li et al. (2020) use this fact to constrain the mass of the NSD. They run several simulations with different NSD mass until the size of the simulated CMZ matches the size of the observed CMZ. While there are several uncertainties in this approach related to the fact that the size of the simulated CMZ also depends on the assumed equation of state of the gas (e.g. Sormani, Binney & Magorrian 2015a) and on the details of the assumed Galactic bar potential (e.g. Sormani et al. 2015b), they also found a NSD mass that is on the lower side of the range indicated by Launhardt et al. (2002) (see fig. 7 in Li et al. 2020), consistent with our result.

Nogueras-Lara et al. (2020b) used the GALACTICNUCLEUS survey (Nogueras-Lara et al. 2019b) to create dereddened K_s luminosity functions and fit them using theoretical stellar evolution models, and estimated the mass contained in a cylinder of $R \lesssim 45 \text{ pc}$ and $|z| \lesssim 20 \text{ pc}$ to be $M = 6.5 \pm 0.4 \times 10^7 M_\odot$. As shown in fig. 7 of Li et al. (2020), this mass would also be consistent with a mass slightly lower than that of Launhardt et al. (2002).

The mass estimation of Launhardt et al. (2002) involves assuming a mass-to-infrared-light ratio, which carries rather large uncertainties (see their section 5.4). For the NSD, they assumed a rather large value of $\Upsilon = 2$. Assuming a value closer to the more common $\Upsilon = 0.6$ (e.g. Meidt et al. 2014; Schödel et al. 2014) would lower their mass estimate considerably.

Given that all our models consistently suggest a mass that is on the lower side of the mass estimated by Launhardt et al. (2002) for all the combination of potential/density/data set/filters we considered,

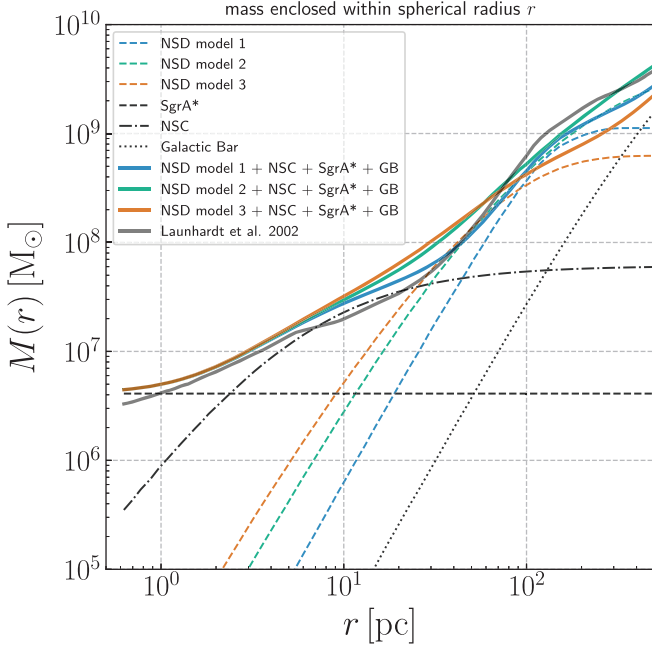


Figure 10. Mass enclosed within spherical radius r . Orange dashed line: best-fitting NSD of our (fiducial) model 3. Blue dashed line: best-fitting NSD of model 1. Green dashed line: best-fitting NSD of model 2. Black dashed line: mass of the central black hole Sgr A* (Gravity Collaboration et al. 2019). Dash-dotted line: NSC (equation 28). Dotted line line: Galactic bulge/bar (GB) model from section 4.2 of Launhardt et al. (2002). Orange solid line: sum NSD from model 3+NSC+Sgr A*+GB. Blue thick solid line: sum NSD from model 1+NSC+Sgr A*+GB. Green thick solid line: sum NSD from model 2+NSC+Sgr A*+GB. Grey thick solid line: best-fitting photometric mass from Launhardt et al. (2002).

and given the large uncertainties in the mass estimation of the latter stemming from the mass-to-infrared-light ratio, we conclude that it is likely the mass of the NSD is lower than the mass estimated by Launhardt et al. (2002).

Fig. 10 compares the mass enclosed within spherical radius r of models 1–3, which differ in the assumed NSD mass distributions. While at $r \lesssim 100$ pc the three models agree well with each other, they diverge at larger radii. This is because the three models have very different extensions as can be seen in Fig. 11. Model 3 is the least extended of the three, while model 2 is by far the most extended. As a result, the total mass of the NSD in models 1, 2, and 3 are $M_{\text{NSD1}} = 1.2 \times 10^9 M_\odot$, $M_{\text{NSD2}} = 5.3 \times 10^9 M_\odot$, and $M_{\text{NSD3}} = 0.69 \times 10^9 M_\odot$, respectively. The first is the easiest to compare with Launhardt et al. (2002) ($M_{\text{NSD}} = 1.4 \pm 0.6 \times 10^9 M_\odot$) since it assumes the same underlying NSD mass distribution. The model of Chatzopoulos et al. (2015) is most likely too extended and gives an unrealistically high total mass. This is not too surprising since these authors were mostly concerned with fitting the innermost few tens of parsecs and not the larger scales considered here. Our fiducial model 3 gives the lowest mass of the three, and is probably the most accurate at least out to $r \simeq 150$ pc given that it is based on the highest resolution data and that the subtraction of the Galactic bulge/bar is made with exactly the same model as Launhardt et al. (2002).

Fig. 12 shows the rotation curves for models 1–3. The rotation curves show significant differences. Since all the three models are all plausible models of the NSD, the scatter can be taken as a measure of the uncertainty in the rotation curve of the Galaxy in the innermost

few hundred parsec. Note however that all the rotation curves are significantly lower than the rotation curve implied by Launhardt et al. (2002).

5.2 A vertically biased disc?

All our models favour a value of the anisotropy parameter $1/\sqrt{b} = \sigma_z/\sigma_R > 1$ (see Table 1 and Figs 8, A1, and A2). This means that vertical oscillations are stronger than radial oscillations, which is unusual for a disc system. For example, the Galactic disc in the solar neighbourhood has values ranging from $\sigma_z/\sigma_R \simeq 0.4$ for the youngest populations to $\simeq 0.8$ for the oldest (e.g. Holmberg, Nordström & Andersen 2009; Martig et al. 2016; Mackereth et al. 2019; Nitschai, Cappellari & Neumayer 2020). Modelling of the kinematics of external galaxies hints at a loose correlation between σ_z/σ_R and Hubble type (e.g. van der Kruit & de Grijs 1999; Gerssen & Shapiro Griffin 2012; Pinna et al. 2018), with σ_z/σ_R decreasing from about 1.0 in early types (lenticulars) to about 0.4 in late types (Sd). Gentile et al. (2015) find $\sigma_z/\sigma_R = 1.2 \pm 0.2$ for the Sb galaxy NGC 3223, which is one of the highest values found in any other galaxy. Our value of $\sigma_z/\sigma_R \sim 1.5$ for the NSD is much larger than any of these. We note, however, that these other measurements are all for large, kpc-scale discs, not for a compact NSD. On much smaller scales, Brown & Magorrian (2013) fit unusually large vertical oscillations in their model of the eccentric disc at the centre of M31.

In order to test whether the finding that $\sigma_z/\sigma_R > 1$ depends on our assumption that the velocity ellipsoid is aligned on cylindrical coordinates, we have repeated our analysis assuming that the velocity ellipsoid is aligned on spherical coordinates (see section 2.4 of Cappellari 2020). We found that models with $\sigma_\theta/\sigma_r > 1$ are clearly favoured, which on the plane $z = 0$ corresponds to $\sigma_z/\sigma_R > 1$. Thus, the alignment of the velocity ellipsoid does not affect the results discussed here.

There are two questions in relation to our finding that $\sigma_z/\sigma_R > 1$. The first is why are such values favoured by our models? Comparison of Fig. 9 with Fig. 13 shows that the reason is that a small value of b (i.e. a large σ_z/σ_R) is needed to reproduce the drop in the observed μ_{obs} near the centre ($|l| \lesssim 0.5$), which is present both in the APOGEE data and the SiO maser data (see middle row in Fig. 9). In Fig. 9, which shows model 3 with $b = 0.45$, the drop is well reproduced, while in Fig. 13, which shows the same model but for $b = 1$, the drop is not well reproduced. The ‘NSD only’ models favour a slightly larger value of b compared to the ‘NSD+NSC’ models because the absence of the NSC component in the middle lowers the velocity dispersion in the central regions compared to the outer parts. The ring models also favour a larger value of b (consistent with $b \sim 1$, see Fig. A2) because the density is essentially zero for $R \lesssim 50$ pc, and therefore those regions do not contribute to the integrals in equations (11) and (12). Assuming that the drop in the data is real (which ought to be confirmed with better data), this suggests that indeed $b < 1$, or that our assumed tracer density population is not representative of the population from which the kinematics are drawn. We note that the observed kinematics directly constrain only the v_R and v_ϕ components of velocity: our constraints on the v_z component come from the integral (11) of ρ and Φ along z . If the potential or density were much flatter than we have assumed, then the σ_z given by (11) would increase and we could fit the observed kinematics with larger values of b . It is currently unclear whether such a strong flattening would be detectable given the extreme and strong differential extinction.

Assuming that $\sigma_z/\sigma_R > 1$ then the second question is how would stars get such large vertical oscillations? In the solar neighbourhood

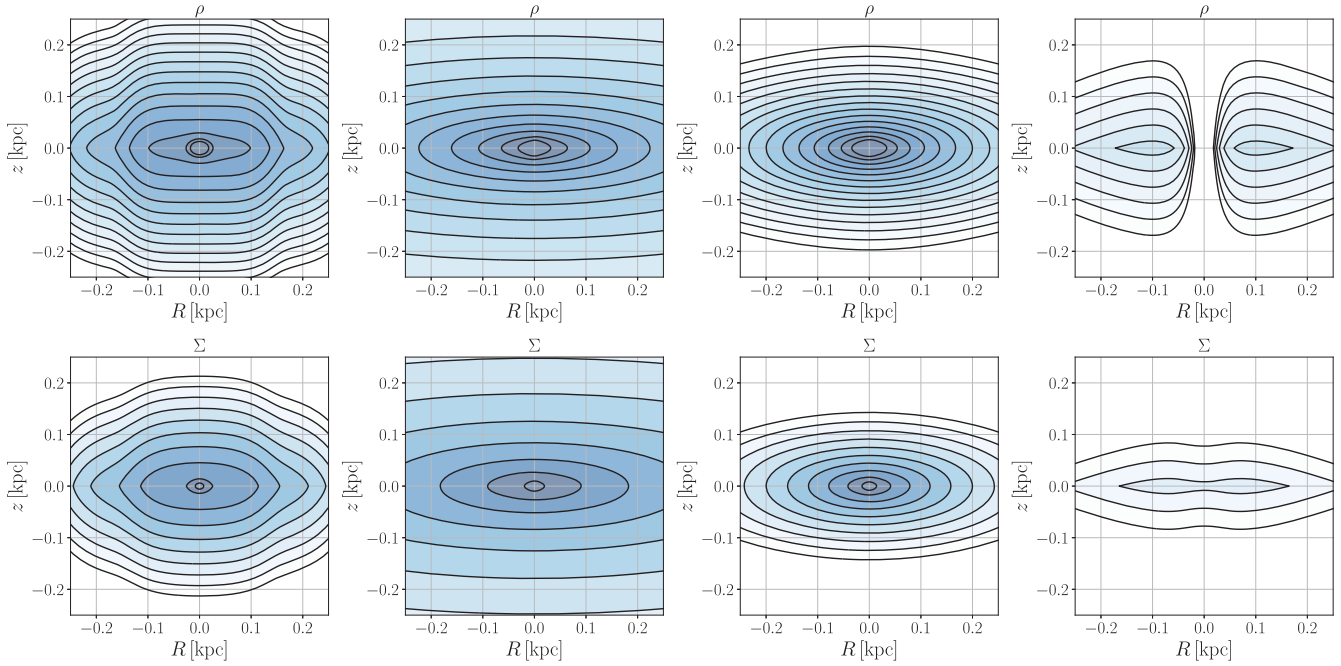


Figure 11. Contours of constant density $\rho(R, z)$ (top) and surface density $\Sigma(R, z)$ (bottom). From left to right: models 1, 2, 3 (see Table 1 and equation 30), and ring model (see equation 31). The ring model is normalized with an arbitrary density scaling $\rho_0 = 10^{10} \text{ M}_\odot \text{ kpc}^{-3}$ (this scaling does not enter the fitting procedure since it simplifies in the calculations of the second moments 19). The lowest contour in the top panels corresponds to a density of $\rho = 3.2 \times 10^7 \text{ M}_\odot \text{ kpc}^{-3}$ and contours are geometrically spaced every 0.3 dex. The lowest contour in the top panels corresponds to a surface density of $\Sigma = 10^8 \text{ M}_\odot \text{ kpc}^{-2}$ and contours are geometrically spaced every 0.33 dex.

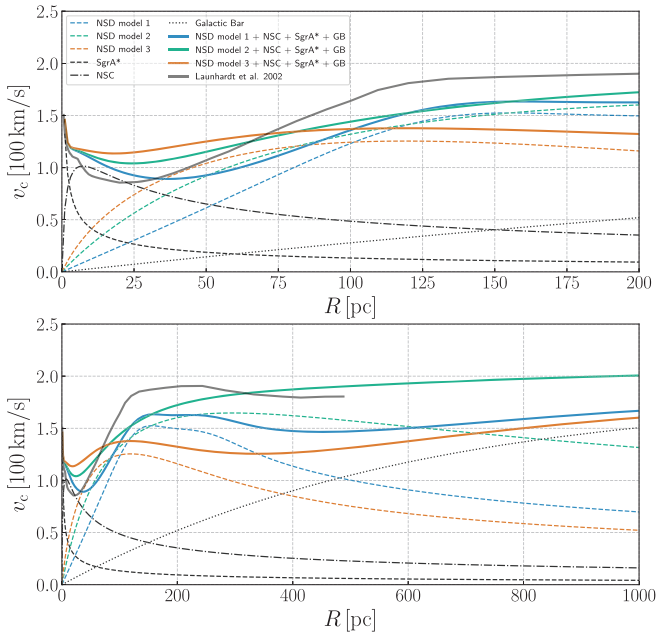


Figure 12. Circular velocity curves in the plane $z = 0$. The top panel is a zoom in the innermost 200 pc of the bottom panel. The style/color scheme is the same as in Fig. 10. The circular velocity curve of Launhardt et al. (2002) is calculated from the corresponding line in Fig. 10 assuming spherical symmetry. For all the other components the circular velocity curve is calculated as $v_{\text{circ}} = \sqrt{R d\Phi/dR}$. For the Galactic bar component, the velocity curve is calculated after azimuthally averaging the triaxial density profile given in Launhardt et al. (2002).

stars are formed from gas clouds that move on almost closed orbits, beginning their lives with random velocities of the order of a few km s^{-1} . There are a number of dynamical processes that inevitably cause these random velocities to increase over time (see Sellwood 2014 for a recent review). Each of these heating mechanisms has a different effect on the ratio σ_z/σ_R . For example, spiral density waves tend to increase σ_R , but have little effect on σ_z . We note, however, that the NSD is probably hot enough that any spiral waves are weak. Two-body scattering of stars by other stars or by giant molecular clouds produces more isotropic heating (Jenkins & Binney 1990; Aumer, Binney & Schönrich 2016), but still limited to $\sigma_z/\sigma_R \lesssim 0.6$ (Ida, Kokubo & Makino 1993), much smaller than we find in the NSD. The most promising mechanism for producing $\sigma_z/\sigma_R > 1$ from an initially cold stellar population is probably from bending instabilities caused by the presence of a counter-rotating population (Khoperskov & Bertin 2017).

An alternative explanation is to relax the assumption that NSD stars were born from a kinematically cold gas disc. Interestingly, observations show that the dense and star-forming molecular gas in the CMZ is currently concentrated into streams that possess strong vertical oscillations of the order of $\Delta z \simeq 30 \text{ pc}$ (see e.g. fig. 4 in Molinari et al. 2011 and fig. 5 in Purcell et al. 2012). This value is similar to the NSD scale height determined by Nishiyama et al. (2013). Moreover, Tress et al. (2020) argue that these large vertical oscillations in the CMZ gas are induced by the large-scale bar-driven accretion and are quite typical in that region based on a combination of observations and hydrodynamical simulations (see their section 6.4). This suggests that NSD stars might already possess large vertical oscillations at birth.

That leaves open the question of whether this mechanism would produce vertical oscillations that are so much stronger than the radial

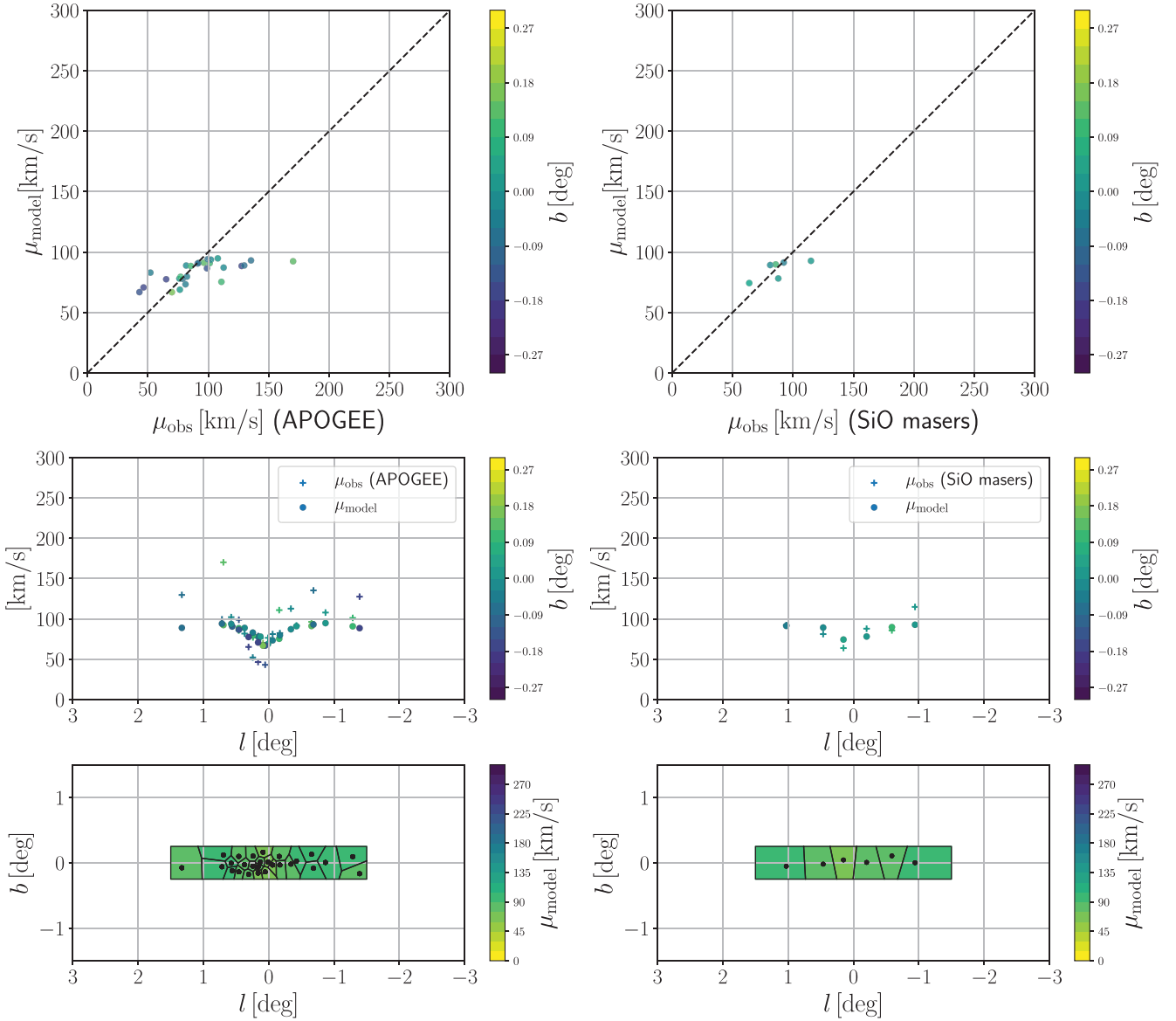


Figure 13. Same as Fig. 9, but for a model that is identical to model 3 except that $b = 1$. Note that the drop in second moment velocities at $|l| \lesssim 0.5$ is reproduced less well than in the model shown in Fig. 9.

ones. The currently observed scale height of the NSD is $\simeq 45$ pc. Assuming that this is similar to the typical vertical excursions of stars in the NSD, it implies that stars oscillate between $z_{\text{min}} \simeq -45$ pc and $z_{\text{max}} \simeq 45$ pc. Assuming $\sigma_z/\sigma_R = 1/\sqrt{b} \simeq 1.5$ (Table 1) and a mean radius of $R \simeq 120$ pc, it implies typical radial oscillations between $R_{\text{min}} \simeq 90$ pc and $R_{\text{max}} \simeq 150$ pc, which is roughly consistent with the expected eccentricities of the x_2 orbits on which the CMZ gas is believed to be flowing on (e.g. Binney et al. 1991; Englmaier & Gerhard 1999; Sormani et al. 2015a, 2020; Tress et al. 2020), and is also consistent with the eccentricity of the ballistic model of Kruijssen et al. (2015, see their table 1). However, large uncertainties remain, and the question should be readdressed in the future when better data become available.

We conclude that the NSD might be the first example of a vertically biased stellar disc ($\sigma_z/\sigma_R > 1$). We propose that the large vertical dispersion might be already imprinted at stellar birth by the star-forming molecular gas in the CMZ.

6 CONCLUSIONS

We have constructed axisymmetric Jeans models of the NSD and have fitted them to the line-of-sight kinematic data of APOGEE and SiO maser stars. We adopted three rather different mass distributions that have been previously shown to be consistent with near/mid-infrared photometry and star counts. Our main results are as follows.

(i) All our models indicate that the mass of the NSD is lower than, but consistent with, the value determined independently from near-infrared photometry by Launhardt et al. (2002) (see Fig. 10). Our fiducial model, based on the recent analysis of high-resolution mid-infrared *Spitzer* data by Gallego-Cano et al. (2020), has a mass contained within spherical radius $r = 100$ pc of $M(r < 100 \text{ pc}) = 3.9 \pm 1 \times 10^8 M_\odot$ and a total mass of $M_{\text{NSD}} = 6.9 \pm 2 \times 10^8 M_\odot$. If instead we assume the same underlying mass distribution of the NSD as Launhardt et al. (2002), which is more spatially extended than our fiducial model, we obtain $M_{\text{NSD}} = 1.2 \pm 2 \times 10^9 M_\odot$, still

lower side than the original determination of Launhardt et al. (2002). The absence/presence of the nuclear stellar cluster in our models and switching between a disc- or ring-like density distribution for the tracer population do not affect these results significantly.

(ii) We find evidence that the NSD is vertically biased, i.e. $\sigma_z/\sigma_R > 1$. If true, the NSD would be the first example of a vertically biased disc system. Observations and theoretical models of the dense star-forming molecular gas in the CMZ suggest that large vertical velocity dispersions may be already imprinted at stellar birth. However, we caution that the finding $\sigma_z/\sigma_R > 1$ depends on many assumptions, and in particular on the observed drop in the second moment of the line-of-sight velocity in the innermost parts, on our assumptions of axisymmetry that the anisotropy is spatially constant, on whether the stellar populations traced by APOGEE and SiO maser data follow a disc- or ring-like density distribution, and, more generally, on our assumption that the available corrected star count data provide good estimates of the underlying light and mass distribution. All of these need to be established by future observations and/or modelling.

(iii) The rotation curves implied by our models are shown in Fig. 12. The rotation curve of our fiducial model 3 is the most accurate to date for the innermost 500 pc of our Galaxy. The scatter between model 1, 2, and 3 can be taken as a measure of the current uncertainty of the rotation curve in this region.

While Jeans models provide useful constraints and insight into the dynamics of the NSD, they are limited as there is no guarantee that they correspond to a physical DF that is everywhere positive ($f > 0$). Therefore, a worthwhile direction of future investigation is to produce DF-based models that can overcome the shortcomings of Jeans modelling.

ACKNOWLEDGEMENTS

We thank the referee for a constructive report that improved the quality of the paper. MCS thanks James Binney, Francesca Fragkoudi, Dimitri Gadotti, Ryan Leaman, Zhi Li, Lorenzo Posti, and Robin Tress for useful comments and discussions. MCS, FN-L, NN, and RSK acknowledge financial support from the German Research Foundation (DFG) via the Collaborative Research Centres (SFB 881, Project-ID 138713538) ‘The Milky Way System’ (subprojects A1, B1, B2, and B8). RSK furthermore thanks for financial support from the European Research Council via the ERC Synergy Grant ‘ECOGAL – Understanding our Galactic ecosystem: from the disk of the Milky Way to the formation sites of stars and planets’ (grant 855130). JM acknowledges funding from the UK Science and Technology Facilities Council under grant number ST/S000488/1. RS acknowledges funding by the Royal Society via a University Research Fellowship.

DATA AVAILABILITY

The data sets used in this paper are publicly available. For APOGEE stars: <https://www.sdss.org/dr16/irspec/>. For SiO masers: <http://vizier.u-strasbg.fr/viz-bin/VizieR?-source=J/A+A/393/115>, <http://vizier.u-strasbg.fr/viz-bin/VizieR?-source=J/A+A/418/103>, and <http://vizier.u-strasbg.fr/viz-bin/VizieR?-source=J/A+A/435/575>.

REFERENCES

Ahumada R. et al., 2020, *ApJS*, 249, 3
Alard C., 2001, *A&A*, 379, L44
Aumer M., Schönrich R., 2015, *MNRAS*, 454, 3166

Aumer M., Binney J., Schönrich R., 2016, *MNRAS*, 462, 1697
Baba J., Kawata D., 2020, *MNRAS*, 492, 4500
Binney J., Tremaine S., 1987, *Galactic Dynamics*. Princeton Univ. Press, Princeton, NJ
Binney J., Tremaine S., 2008, *Galactic Dynamics*, 2nd edn. Princeton Univ. Press, Princeton, NJ
Binney J., Gerhard O. E., Stark A. A., Bally J., Uchida K. I., 1991, *MNRAS*, 252, 210
Binney J. et al., 2014, *MNRAS*, 439, 1231
Bland-Hawthorn J., Gerhard O., 2016, *ARA&A*, 54, 529
Bovy J. et al., 2014, *ApJ*, 790, 127
Bovy J., Rix H.-W., Green G. M., Schlafly E. F., Finkbeiner D. P., 2016, *ApJ*, 818, 130
Brown C. K., Magorrian J., 2013, *MNRAS*, 431, 80
Cappellari M., 2008, *MNRAS*, 390, 71
Cappellari M., 2020, *MNRAS*, 494, 4819
Cappellari M., Copin Y., 2003, *MNRAS*, 342, 345
Catchpole R. M., Whitlock P. A., Glass I. S., 1990, *MNRAS*, 247, 479
Chatzopoulos S., Fritz T. K., Gerhard O., Gillessen S., Wegg C., Genzel R., Pfuhl O., 2015, *MNRAS*, 447, 948
Cole D. R., Debattista V. P., Erwin P., Earp S. W. F., Roškar R., 2014, *MNRAS*, 445, 3352
Englmaier P., Gerhard O., 1999, *MNRAS*, 304, 512
Evans N. W., de Zeeuw P. T., 1994, *MNRAS*, 271, 202
Everall A., Evans N. W., Belokurov V., Schönrich R., 2019, *MNRAS*, 489, 910
Feldmeier A. et al., 2014, *A&A*, 570, A2
Fux R., 1999, *A&A*, 345, 787
Gadotti D. A. et al., 2019, *MNRAS*, 482, 506
Gadotti D. A. et al., 2020, *A&A*, in press ([arXiv:2009.01852](https://arxiv.org/abs/2009.01852))
Gallego-Cano E., Schödel R., Noguera-Lara F., Dong H., Shahzamanian B., Fritz T. K., Gallego-Calvente A. T., Neumayer N., 2020, *A&A*, 634, A71
Gentile G. et al., 2015, *A&A*, 576, A57
Gerhard O., Martínez-Valpuesta L., 2012, *ApJ*, 744, L8
Gerssen J., Shapiro Griffin K., 2012, *MNRAS*, 423, 2726
Gravity Collaboration et al., 2019, *A&A*, 625, L10
Habing H. J., 2016, *A&A*, 587, A140
Habing H. J., Sevenster M. N., Messineo M., van de Ven G., Kuijken K., 2006, *A&A*, 458, 151
Henshaw J. D. et al., 2016, *MNRAS*, 457, 2675
Holmberg J., Nordström B., Andersen J., 2009, *A&A*, 501, 941
Ida S., Kokubo E., Makino J., 1993, *MNRAS*, 263, 875
Jenkins A., Binney J., 1990, *MNRAS*, 245, 305
Khoperskov S., Bertin G., 2017, *A&A*, 597, A103
Kuijssens J. M. D., Dale J. E., Longmore S. N., 2015, *MNRAS*, 447, 1059
Launhardt R., Zylka R., Mezger P. G., 2002, *A&A*, 384, 112
Li Z., Shen J., Schive H.-Y., 2020, *ApJ*, 889, 88
Lindqvist M., Habing H. J., Winnberg A., 1992, *A&A*, 259, 118
Mackereth J. T. et al., 2019, *MNRAS*, 489, 176
Majewski S. R. et al., 2017, *AJ*, 154, 94
Mamon G. A., Boué G., 2010, *MNRAS*, 401, 2433
Martig M., Minchev I., Ness M., Funesneau M., Rix H.-W., 2016, *ApJ*, 831, 139
Matsunaga N. et al., 2015, *ApJ*, 799, 46
Meidt S. E. et al., 2014, *ApJ*, 788, 144
Messineo M., Habing H. J., Sjouwerman L. O., Omont A., Menten K. M., 2002, *A&A*, 393, 115
Messineo M., Habing H. J., Menten K. M., Omont A., Sjouwerman L. O., 2004, *A&A*, 418, 103
Messineo M., Habing H. J., Menten K. M., Omont A., Sjouwerman L. O., Bertoldi F., 2005, *A&A*, 435, 575
Molinari S. et al., 2011, *ApJ*, 735, L33
Molloy M., Smith M. C., Evans N. W., Shen J., 2015, *ApJ*, 812, 146
Nishiyama S., Nagata T., Tamura M., Kandori R., Hatano H., Sato S., Sugitani K., 2008, *ApJ*, 680, 1174
Nishiyama S. et al., 2013, *ApJ*, 769, L28
Nitschai M. S., Cappellari M., Neumayer N., 2020, *MNRAS*, 494, 6001
Noguera-Lara F. et al., 2018a, *A&A*, 610, A83

Nogueras-Lara F. et al., 2018b, *A&A*, 620, A83
 Nogueras-Lara F., Schödel R., Najarro F., Gallego-Calvente A. T., Gallego-Cano E., Shahzamanian B., Neumayer N., 2019a, *A&A*, 630, L3
 Nogueras-Lara F. et al., 2019b, *A&A*, 631, A20
 Nogueras-Lara F., Schödel R., Neumayer N., Gallego-Cano E., Shahzamanian B., Gallego-Calvente A. T., Najarro F., 2020a, *A&A*, in press ([arXiv:2007.04401](https://arxiv.org/abs/2007.04401))
 Nogueras-Lara F. et al., 2020b, *Nat. Astron.*, 4, 377
 Pinna F., Falcón-Barroso J., Martíg M., Martínez-Valpuesta I., Méndez-Abreu J., van de Ven G., Leaman R., Lyubenova M., 2018, *MNRAS*, 475, 2697
 Pizzella A., Corsini E. M., Morelli L., Sarzi M., Scarlata C., Stiavelli M., Bertola F., 2002, *ApJ*, 573, 131
 Purcell C. R. et al., 2012, *MNRAS*, 426, 1972
 Rodríguez-Fernández N. J., Combes F., 2008, *A&A*, 489, 115
 Schödel R., Najarro F., Muzic K., Eckart A., 2010, *A&A*, 511, A18
 Schödel R., Feldmeier A., Kunneriath D., Stolovy S., Neumayer N., Amaro-Seoane P., Nishiyama S., 2014, *A&A*, 566, A47
 Schönrich R., Aumer M., Sale S. E., 2015, *ApJ*, 812, L21
 Sellwood J. A., 2014, *Rev. Modern Phys.*, 86, 1
 Siebert A. et al., 2008, *MNRAS*, 391, 793

Skrutskie M. F. et al., 2006, *AJ*, 131, 1163
 Sormani M. C., Binney J., Magorrian J., 2015a, *MNRAS*, 449, 2421
 Sormani M. C., Binney J., Magorrian J., 2015b, *MNRAS*, 454, 1818
 Sormani M. C., Treß R. G., Ridley M., Glover S. C. O., Klessen R. S., Binney J., Magorrian J., Smith R., 2018, *MNRAS*, 475, 2383
 Sormani M. C., Treß R. G., Glover S. C. O., Klessen R. S., Battersby C. D., Clark P. C., Hatchfield H. P., Smith R. J., 2020, *MNRAS*, 497, 5024
 Treß R. G., Sormani M. C., Glover S. C. O., Klessen R. S., Battersby C. D., Clark P. C., Hatchfield H. P., Smith R. J., 2020, *MNRAS*, preprint ([arXiv:2004.06724](https://arxiv.org/abs/2004.06724))
 van der Kruit P. C., de Grijs R., 1999, *A&A*, 352, 129
 Wright E. L. et al., 2010, *AJ*, 140, 1868
 Zasowski G. et al., 2013, *AJ*, 146, 81

APPENDIX A: PROBABILITY DISTRIBUTIONS FOR MODELS 7–18

Figs A1 and A2 show the probability distributions for models 7–12 and 13–18, respectively.

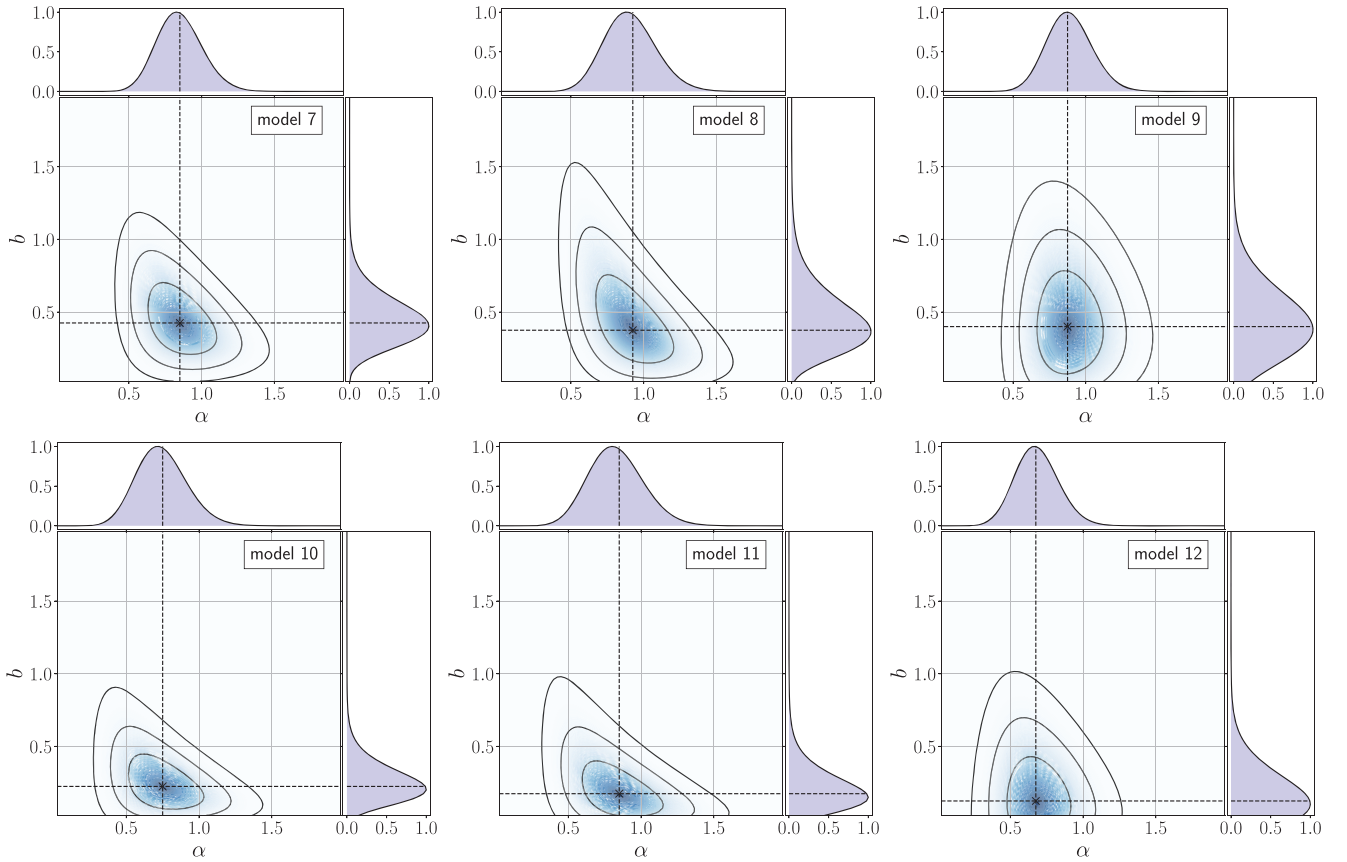


Figure A1. Same as Fig. 8, but for models 7–12.

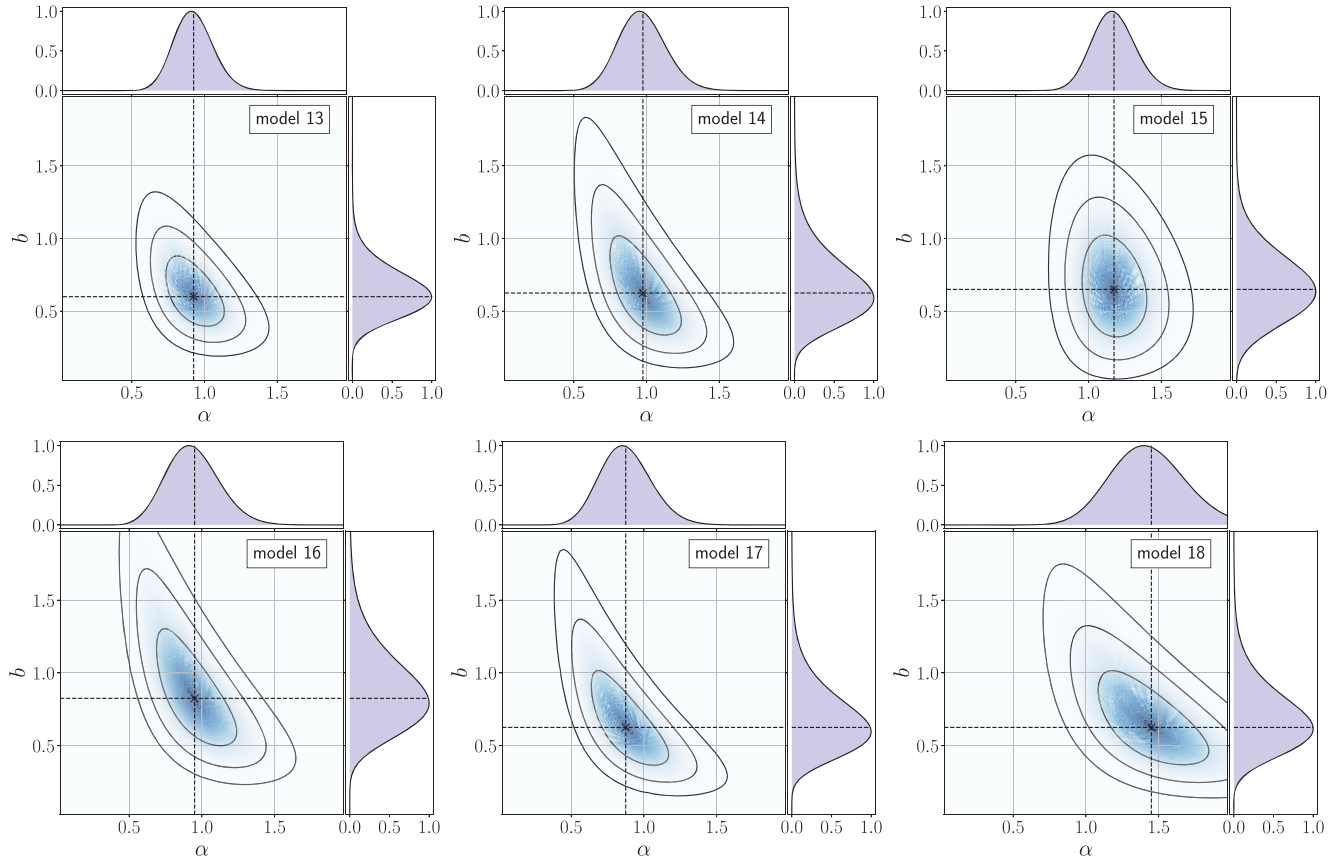


Figure A2. Same as Fig. 8, but for models 13–18.

APPENDIX B: COMPARISON BETWEEN DATA AND MODELS 1 AND 2

and 2.

Figs B1 and B2 show a comparison between the data and models 1

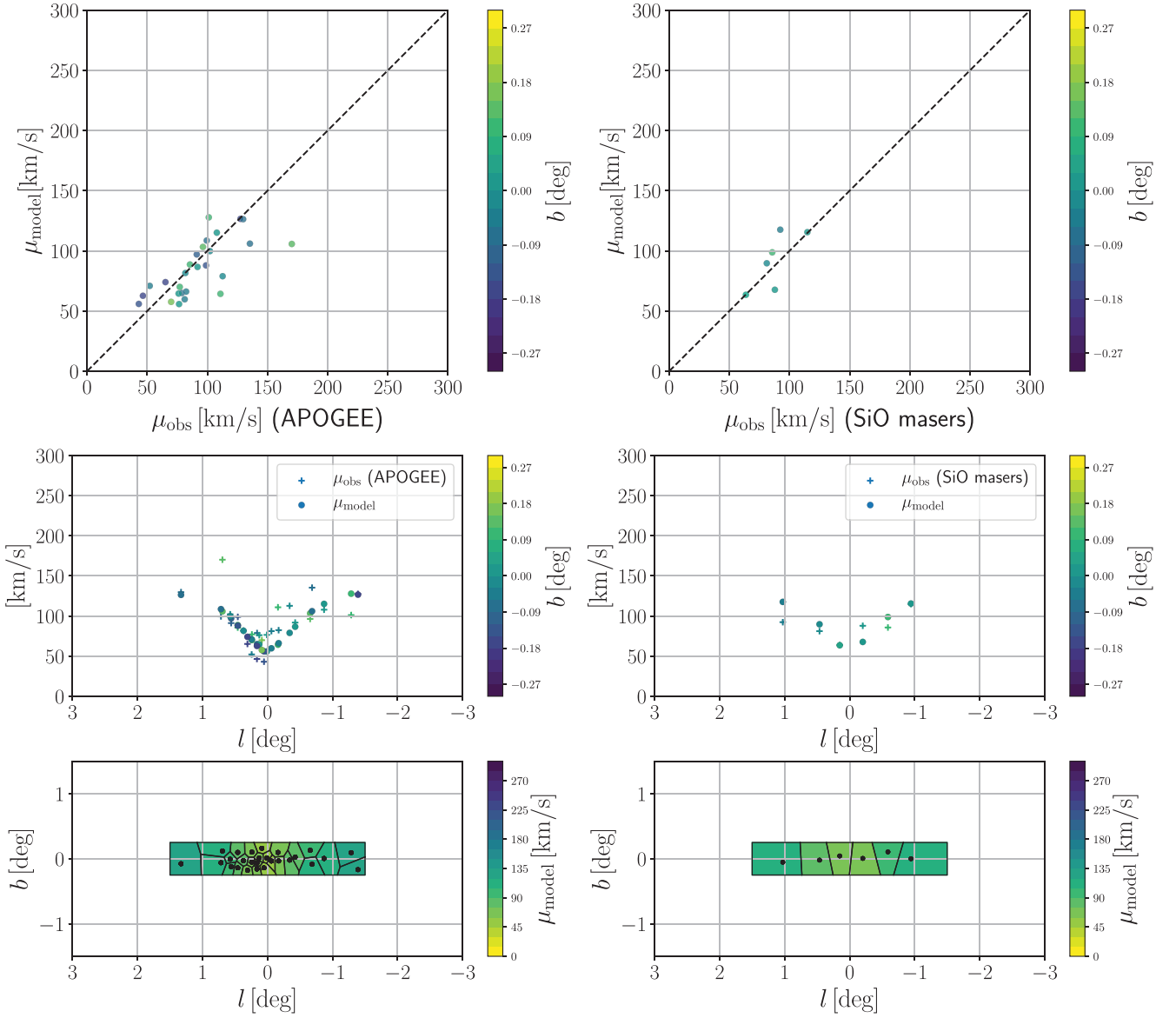


Figure B1. Same as Fig. 9, but for model 1.

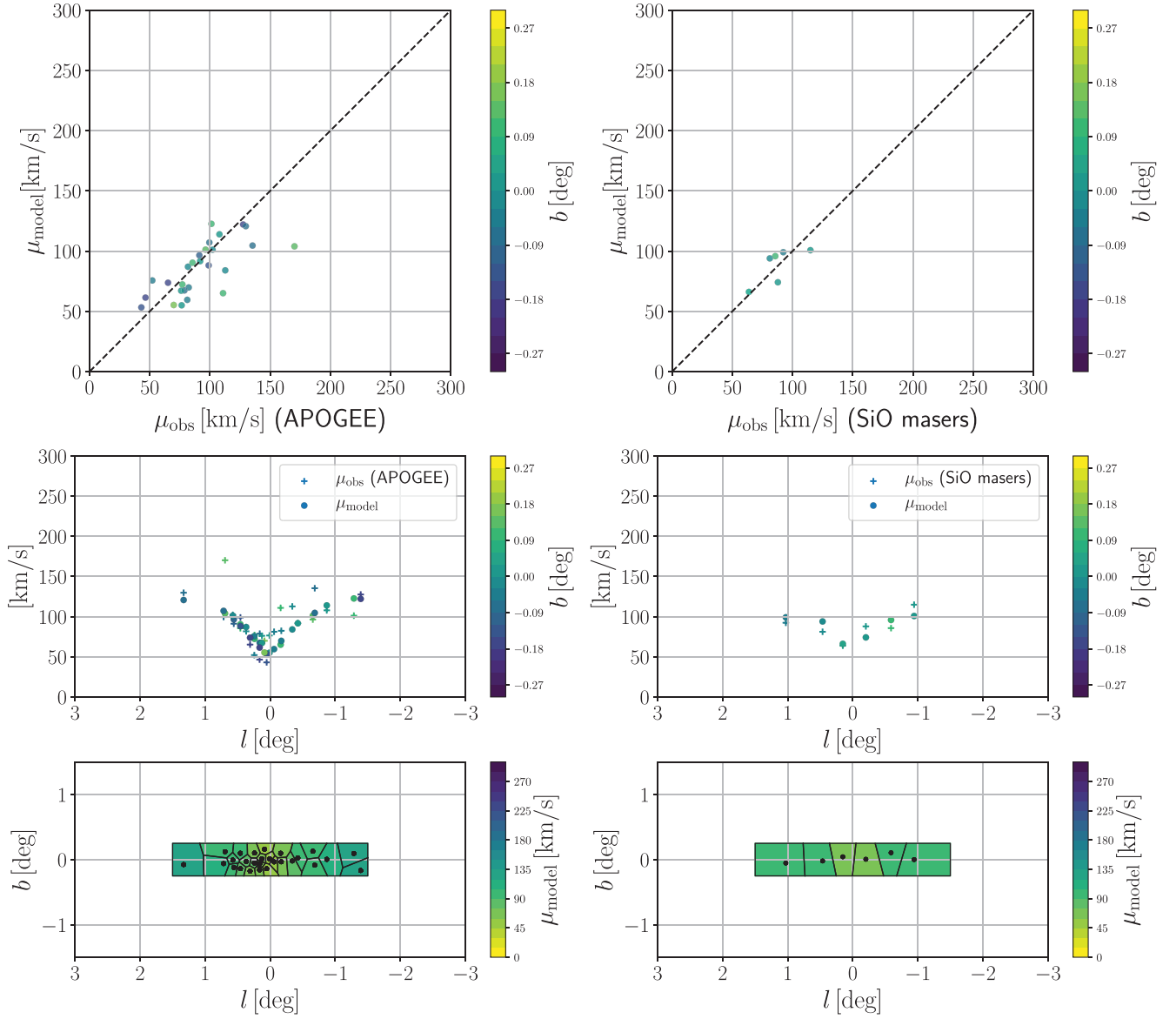


Figure B2. Same as Fig. 9, but for model 2.

This paper has been typeset from a \LaTeX file prepared by the author.