Rotational Velocity in the Milky Way

Colombo Gabriele, Maszkowski Michal, Seghezzi Daniele

March 31, 2025

Abstract

In this report we present a model for the rotation of stars in the Milky Way around its center. In a fisrt simpler model, we assume that the stars in the GC (Galactic Center) frame move in circular orbits around its origin with a constant velocity V_{rot}, as well as the LSR (Local Standard of Rest), and that the Sun frame is moving with a velocity $U_{\odot}$ in the x-direction and $V_{\odot}$ in the y-direction in the GC frame with respect to the LSR. Then, we present a second, more complex model, in which we add, for each star, a random component $v'_{\mathrm{rand}} \sim N(0, \sigma)$ to its radial velocity.

In this report, we use data provided by GAIA DR2 [1]. For each star on the galactic plane we consider its longitude, parallax, and radial velocity, with respect to GAIA, which we equate with the Sun frame of reference. We also consider the provided uncertainties of the measurements of parallax and radial velocity.

By means of Bayesian inference on the data we find the distributions of parameters of the two models, namely $\boldsymbol{\theta_1} = (V_{rot}, U_{\odot}, V_{\odot})$ and $\boldsymbol{\theta_2} = (V_{rot}, U_{\odot}, V_{\odot}, \sigma)$. We used MCMC (Monte Carlo Markov Chain) to estimate the non-normalized posterior of the two models.

In the first model, we estimate the parameters and their 95% confidence interval to be: $V_{\mathrm{rot}} = 211.45$, 95% CI = [211.42, 211.50] $\mathrm{km\,s^{-1}}$; $u_{\odot} = 11.638$, 95% CI = [11.633, 11.644] $\mathrm{km\,s^{-1}}$; $v_{\odot} = 21.604$, 95% CI = [21.599, 21.609] $\mathrm{km\,s^{-1}}$. With the second one instead, we got: $V_{\mathrm{rot}} = 204$, 95% CI = [202, 206] $\mathrm{km\,s^{-1}}$; $u_{\odot} = 11.7$, 95% CI = [11.4, 12.0] $\mathrm{km\,s^{-1}}$; $v_{\odot} = 21.7$, 95% CI = [21.4, 22.1] $\mathrm{km\,s^{-1}}$; $\sigma = 30.6$, 95% CI = [30.5, 30.8] $\mathrm{km\,s^{-1}}$.

1 Data

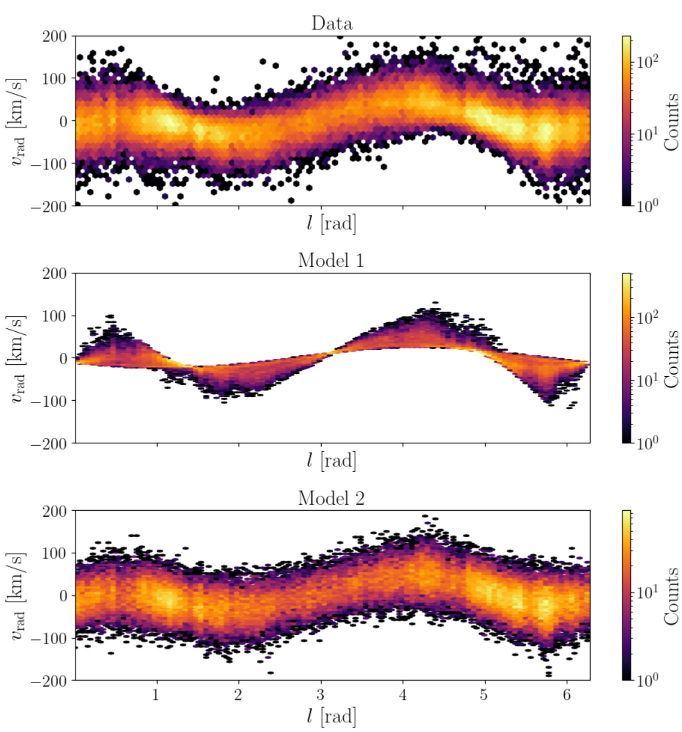

Figure 1: Upper: radial velocities in the interval [-200, 200] $\mathrm{km\,s^{-1}}$ of the used data. Middle: predictions of the radial velocities with the first model. Lower: simulation of the distribution of the radial velocities with the second model.

In this work we use data collected by GAIA DR2 [1]. From the vast dataset of stars analyzed by GAIA, we select only those for which radial velocities, v_{rad}, were measured relative to the Sun using the Doppler effect. To manage the dataset size efficiently, we then apply a random selection to significantly reduce the quantity of data (imposing the random index of data to be less than 100000000).

For each selected star, we extract key parameters from GAIA DR2, including parallax p and its associated error σ_{p}, radial velocity v_{rad} with its measurement uncertainty σ_{v}, and galactic coordinates, i.e. latitude b and longitude l.

To focus our analysis on stars located within the galactic plane, we impose a selection criterion of $|b| < 5°$. Additionally, to ensure the reliability of the data, we retain only stars with a relative parallax error smaller than 20% and a radial velocity error below 5 $\mathrm{km\,s^{-1}}$. Following these criteria, 75659 stars were selected. The distribution of their radial velocity with respect to the Sun is plotted in the upper panel of figure 1 as a function of their longitude. Only values in the interval [-200$\mathrm{km\,s^{-1}}$, 200$\mathrm{km\,s^{-1}}$] were reported for better visualisation. However, values up to 500 $\mathrm{km\,s^{-1}}$ were observed. These are likely outliers, and a more sophisticated data selection procedure may improve the quality of the following analysis.

2 Model

In this section we present the physical and the statistical models we developed for the study of the Galactic kinematics of the Milky Way. Our physical assumptions, presented in the paragraph 2.1, lead to a prediction for the radial motion of stars with respect to the Sun as a function of their longitude l and parallax p. The prediction, for each star, can then be compared to the direct measurement. In the paragraph 2.2 we formally present our statistical assumptions on the data and the parameters of the two models with their priors and likelihoods.

2.1 Physical Model

In our model, we only consider the motion on the galactic plane ($b \approx 0$). We assume that each star moves with a circular orbit around the Galactic Center GC with the same speed, V_{rot}. Fixing the frame of reference as in figure 2, a generic star s has a velocity $\boldsymbol{v}_{\text{s}}$ given by:

$$\boldsymbol{v}_{\text{s}} = V_{\text{rot}} \begin{pmatrix} \sin(\varphi) \\ -\cos(\varphi) \end{pmatrix} \tag{1}$$

where φ is the angle from the x-axis of the star with respect to the GC. In the following, primate vectors are in the frame of reference of the Sun, whereas unprimed ones are in the frame of reference of the GC. Angles are supposed to be expressed as radians.

In this frame of reference, the Sun has $\varphi = \pi$, and we assume it to be at a fixed distance $R = 8300\,\text{pc}$ with respect to the GC. In this model, the drift velocity of the LSR (Local Standard of Rest) is also given by equation 1. In addition, the Sun has its own peculiar motion with respect to the LSR, with components, $U_\odot$ along the x-axis, and $V_\odot$ along the y one. Therefore, the total velocity $v_\odot$ of the Sun is:

$$\boldsymbol{v}_\odot = \begin{pmatrix} 0 \\ V_{rot} \end{pmatrix} + \begin{pmatrix} U_\odot \\ V_\odot \end{pmatrix} \tag{2}$$

The velocity $\boldsymbol{v}'_s$ of a star s, in the frame of reference of the Sun, at distance d from the Sun is given by the equation:

$$\boldsymbol{v}'_s = \boldsymbol{v}_s - \boldsymbol{v}_\odot \tag{3}$$

We can now project the velocity $\boldsymbol{v}'_s$ onto radial direction with respect to the Sun $\hat{e}'_r$,

obtaining the radial component $v_{\text{rad}'}$ of the velocity of a star with longitude l in the Sun's frame of reference:

$$\begin{aligned} v_{\text{rad}'} &= \boldsymbol{v}'_s \cdot \hat{e}'_r = \\ &= V_{\text{rot}} \Big[\sin\varphi \cos l - (1 + \cos\varphi) \sin l \Big] - U_\odot \cos l - V_\odot \sin l \end{aligned} \tag{4}$$

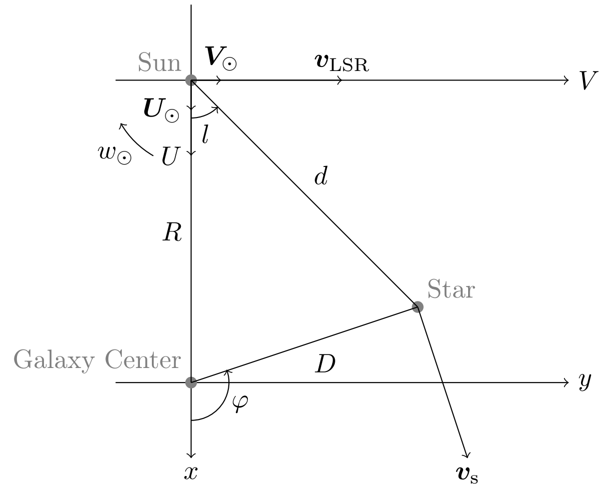

Figure 2: Frames of reference. The angular frequency $w_\odot$ only contributes to the tangential component of the velocities of the stars with in the Sun's frame of reference, and is therefore omitted in equation 3.

Eq.4 must be adapted to the actual data provided by GAIA, which means expressing $sin\varphi$ and $cos\varphi$ in terms of l and the parallax p, expressed in milliarcoseconds (mas). First of all, the distance in parsec can be computed as:

$$d[\text{pc}] = \frac{1000}{p[\text{mas}]} \tag{5}$$

Then, by geometric considerations we can evaluate $D, \sin\varphi$ and $\cos\varphi$ as a function of R, d and l (see figure 2), obtaining:

$$D = \sqrt{d^2 + R^2 - 2dR\cos l} \tag{6}$$

$$\sin\varphi = \pm\sqrt{1 - \cos^2\varphi} = \frac{d \sin l}{D} \tag{7}$$

$$\cos\varphi = \frac{d\cos l - R}{D} \tag{8}$$

By substituting eq.5-8 into eq.4, we get an expression for the prediction of the model for the radial component of the velocity of star i $\hat{v}^{(1)}_{rad,i}(l_i, p_i)$ as a function of the measurements of its longitude and parallax l_i, p_i:

$$\begin{aligned} \hat{v}^{(1)}_{rad,i}(l_i, p_i) =& V_{\text{rot}} \sin l_i \left(\frac{R}{\sqrt{(\frac{1000}{p_i})^2 + R^2 - 2(\frac{1000}{p_i}) R \cos l_i}} - 1 \right) - \\ & - U_\odot \cos l_i - V_\odot \sin l_i \end{aligned} \tag{9}$$

In a second, more sophisticated model, we also consider the random motion of all stars, adding to the right-hand-side of equation 9 a random variable $v_{\text{rand,i}} \sim N(0, \sigma)$, obtaining:

$$\begin{aligned} \hat{v}^{(2)}_{rad,i}(l_i, p_i) =& V_{\text{rot}} \sin l_i \left(\frac{R}{\sqrt{(\frac{1000}{p_i})^2 + R^2 - 2(\frac{1000}{p_i}) R \cos l_i}} - 1 \right) - \\ & - U_\odot \cos l_i - V_\odot \sin l_i + v_{\text{rand,i}} \end{aligned} \tag{10}$$

2.2 Statistical Model

In the following we describe the Bayesian inference we have made on the parameters of the two models. We used Monte Carlo Markov Chains (MCMC) provided by the emcee package in python to get an approximation of the non-normalized posteriors of our models $\mathbb{P}^*(\theta|\mathcal{D})$, considering only the numerator of Bayes' theorem (posterior):

$$\mathbb{P}^*(\theta|\mathcal{D}) = \mathscr{L}(\mathcal{D}|\theta)\mathbb{P}(\theta) \tag{11}$$

where $\mathscr{L}(\mathcal{D}|\theta)$ is the likelihood and $\mathbb{P}(\theta)$ is the prior. In practice, the logarithms of these quantities were used to achieve numerical stability.

For simplicity, in both models, we assumed each measurement and each parameter to be independent of all the others. Under this assumptions, the likelihood and the prior factorize in the product of individual terms, and their logarithms are given by the sum of those individual terms. Then, we assume the value m of each measure to be the sum of its true value, and a random error $\epsilon \sim N(0, \sigma_{\mathrm{m}})$, extracted from a normal distribution centered at 0, with standard deviation given by the statistical uncertainty provided by GAIA database.

For the first model we have a set of three parameters $\boldsymbol{\theta_1} = (V_{\mathrm{rot}}, U_\odot, V_\odot)$. Neglecting the uncertainties associated to the parallax measurements, the difference between the measure of the radial velocity $v_{\mathrm{rad},i}$ and its prediction $\hat{v}^{(1)}_{\mathrm{rad},i}$ is a random variable extracted from a normal distribution centered in 0, with standard deviation only given by the statistical uncertainty on the measurements of the radial velocity. The log-likelihood of this model therefore is given by the sum of independant terms as:

$$\log \mathbb{P}^{(1)}(\mathcal{D}|\theta_1) = -\frac{1}{2}\sum_i [\log(2\pi\sigma^2_{\mathrm{v},i}) + \frac{(v_{rad,i} - \hat{v}^{(1)}_{rad,i})^2}{\sigma^2_{\mathrm{v},i}}] \tag{12}$$

Then, we chose a flat prior for $V_{\mathrm{rot}} \in [0, 500\,\mathrm{km\,s^{-1}}]$, in order to include typical values of the rotational motion of stars in spiral barred (Sb) galaxies which are found in the range [144 $\mathrm{km\,s^{-1}}$, 330$\mathrm{km\,s^{-1}}$] [3]. For $U_\odot$ and $V_\odot$ we chose a gaussian prior centered in 0, $\log \mathbb{P}^{(1)}(U_\odot) + \log \mathbb{P}^{(1)}(U_\odot) \sim -\frac{U_\odot^2 + V_\odot^2}{v^2_{\mathrm{gal}}}$, assuming the peculiar motion of the sun to be analogous to a stochastic thermal motion. As a value for v_{gal}, we arbitrarily chose 200 $\mathrm{km\,s^{-1}}$ as a reference number for the random motion of stars in a Sb galaxy.

In our second model, there are four parameters: $\boldsymbol{\theta_2} = (V_{rot}, U_\odot, V_\odot, \sigma)$. Accounting for the errors on the parallax measurements, the model prediction $\hat{v}^{(2)}_{\mathrm{rad},i}$ (see equation 10) is a random variable extracted from a normal distribution centered in 0, and with variance $\sigma^(2)$ given by the sum of the variance of the random component σ^2, and the contribution $\sigma^2_{\mathrm{p}->(2)}$ originating from the error on the parralax measurement, which can be computed by error propagation as:

$$\sigma^2_{\mathrm{p}_i->(2)} = \left(\frac{\partial \hat{v}^{(2)}_{\mathrm{rad},i}}{\partial \mathrm{p}_i}\right)^2 \sigma^2_{\mathrm{p}_i} \tag{13}$$

where the derivative can be computed analytically from equation 10. Then, the difference between the measured value of the radial velocity $v_{\mathrm{rad},i}$ and the model prediction $\hat{v}^{(2)}_{\mathrm{rad},i}$ is a random variable extracted from a normal distribution centered in 0, and with total variance given by the sum of the variances of the error on the radial velocity measurements, and that on the model prediction, which results in:

$$\sigma^2_{\mathrm{tot},i} = \sigma^2_{\mathrm{v},i} + \sigma^2_{\mathrm{p}_i->(2)} + \sigma^2 \tag{14}$$

As the likelihood factors in independent terms, we get, for its logarithm, the expression:

$$\log \mathscr{L}^{(2)}(\mathcal{D}|\theta_2) = -\frac{1}{2}\sum_i \left\{ \log[2\pi(\sigma_i^2 + \sigma^2)] + \frac{(v_{rad,i} - \hat{v}^{(2)}_{rad,i})^2}{\sigma^2_{\mathrm{tot},i}} \right\} \tag{15}$$

We keep for the first three parameters the same priors decided in the first model, while we also assume the fourth parameter σ to be uncorrelated to the others. The log-prior of the parameter σ was chosen as the non-informative prior for the standard deviation of a gaussian [2], since, in our model, it has a similar role:

$$\log \mathbb{P}(\sigma) = -\log(\sigma) \tag{16}$$

3 DataAnalysis

4 Conclusions

References

[1] A. G. A. Brown et al. "Gaia Data Release 2: Summary of the contents and survey properties". In: *Astronomy amp; Astrophysics* 616 (Aug. 2018), A1. ISSN: 1432-0746. DOI: 10.1051/0004-6361/201833051. URL: http://dx.doi.org/10.1051/0004-6361/201833051.

[2] David J.C. MacKay. *Information Theory, Inference, and Learning Algorithms.* Cambridge, UK: Cambridge University Press, 2003. ISBN: 9780521642989. URL: http://www.inference.phy.cam.ac.uk/mackay/itila/.

[3] Peter Schneider. *Extragalactic Astronomy and Cosmology: An Introduction.* 2nd. See table 3.2 for the typical rotational speed of spiral galaxies. See appendix B for the estimation of the mass from measurements of luminosity. Heidelberg, Germany: Springer-Verlag Berlin Heidelberg, 2015. ISBN: 978-3-642-54082-0. DOI: 10.1007/978-3-642-54083-7.

A Posterior distributions

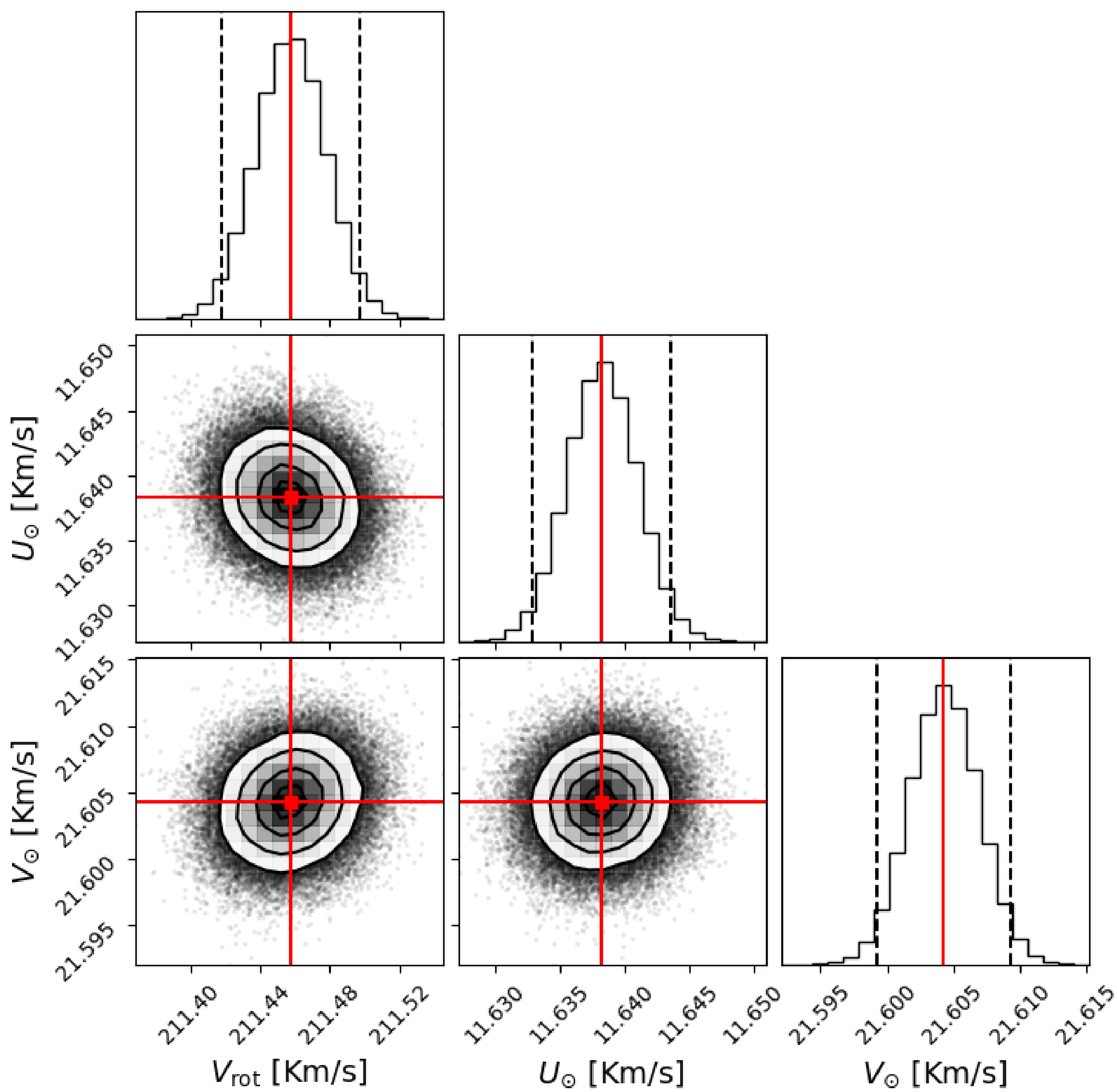

Figure 3: Posterior distribution for the parameters of the first model.

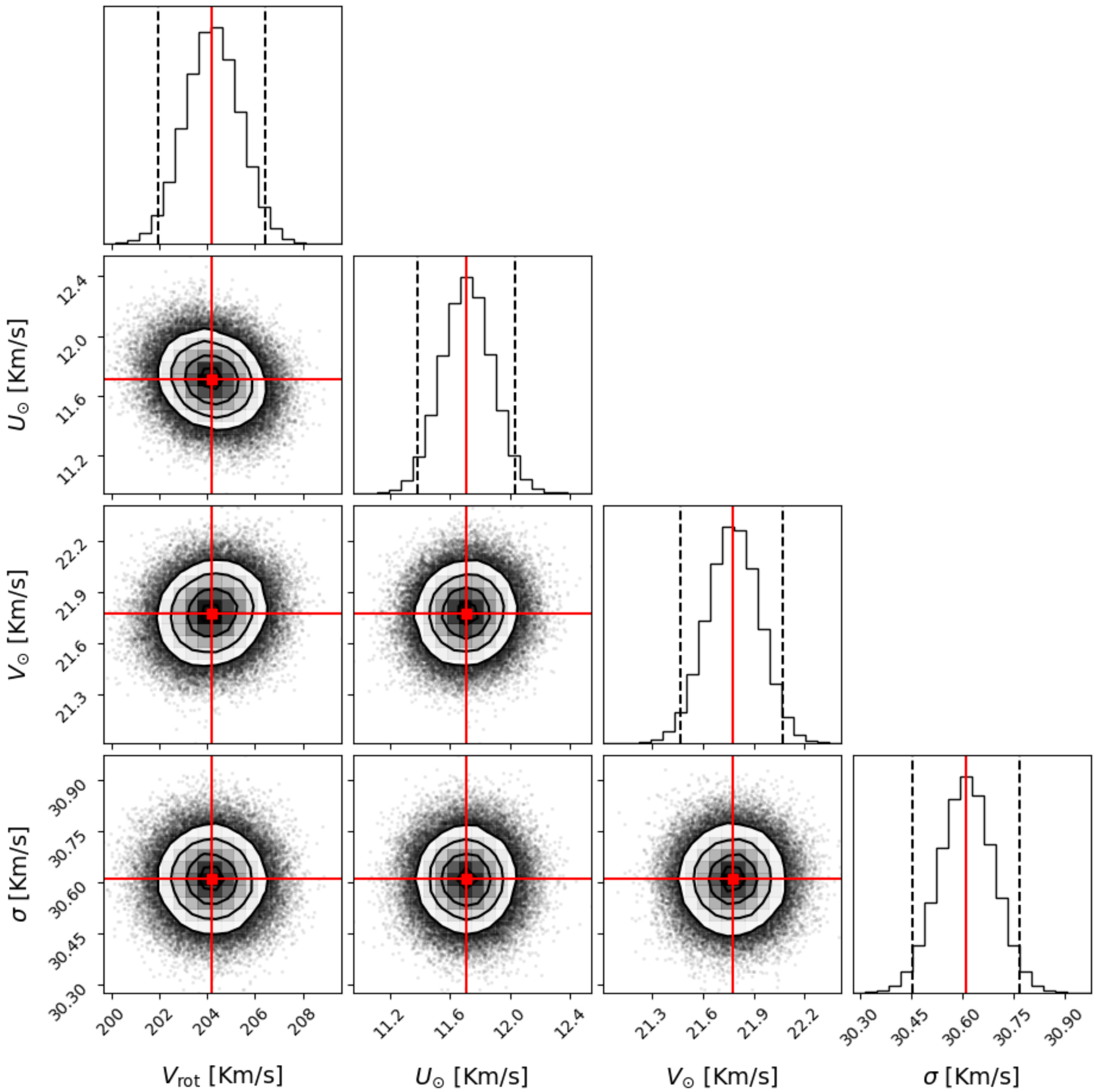

Figure 4: Posterior distribution for the parameters of the second model.