



# Inferring Stellar Densities with Flexible Models. I. The Distribution of RR Lyrae in the Milky Way with Gaia DR3

Madeline Lucey<sup>1</sup>, Cecilia Mateu<sup>2</sup>, Adrian M. Price-Whelan<sup>3</sup>, David W. Hogg<sup>3,4,5</sup>, Hans-Walter Rix<sup>4</sup>, and Robyn E. Sanderson<sup>1</sup>

<sup>1</sup> Department of Physics & Astronomy, University of Pennsylvania, 209 S 33rd Street, Philadelphia, PA 19104, USA

<sup>2</sup> Departamento de Astronomía, Instituto de Física, Universidad de la República Iguá 4225, CP 11400 Montevideo, Uruguay

<sup>3</sup> Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA

<sup>4</sup> Max Planck Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany

<sup>5</sup> Center for Cosmology and Particle Physics, Department of Physics, New York University, 726 Broadway, New York, NY 10003, USA

Received 2025 October 3; revised 2026 February 11; accepted 2026 February 23; published 2026 March 25

## Abstract

Understanding the formation and evolutionary history of the Milky Way requires detailed mapping of its stellar components, which preserve fossil records of the Galaxy’s assembly through cosmic time. RR Lyrae stars are particularly well suited for this endeavor, as they are old, standard candle variables that probe the Galaxy’s earliest formation epochs. In this work, we employ a hierarchical Bayesian Gaussian Mixture Model (GMM) to characterize the three-dimensional density distribution of RR Lyrae stars in the Milky Way over the galactocentric radius ( $R$ ) of  $\approx 0.2$ –120 kpc. This approach provides a flexible framework for modeling complex stellar distributions, particularly in the inner Galaxy where the bulge, disk, and halo components overlap. Our analysis reveals that the inner Galaxy ( $R \lesssim 10$  kpc) is dominated by a distinct prolate stellar population with axis ratio  $q = 1.31$ . Consistent with previous work, we find the halo follows a  $r^{-4}$  power-law profile that flattens within 12 kpc of the Galactic center. We also confirm the outer halo ( $R \gtrsim 10$  kpc) is oblate  $q = 0.70$  with a tilt angle of  $18^\circ$ . We report for the first time that this tilt aligns the halo major axis in the direction of the Sagittarius dwarf galaxy. These results establish GMMs as an effective and flexible tool for modeling Galactic structure and provide new constraints on the distribution of old stars in the inner Galaxy.

*Unified Astronomy Thesaurus concepts:* Milky Way Galaxy (1054); RR Lyrae variable stars (1410); Milky Way stellar halo (1060)

*Materials only available in the online version of record: machine-readable table*

## 1. Introduction

With the billions of resolvable stars in the Milky Way, we can study the formation and evolutionary history of our Galaxy in exquisite detail. The Galaxy’s past structure is chronicled in the distribution of its resolved stellar populations (E. K. Grebel 1999; M. W. Feast & P. A. Whitelock 2000; K. Freeman & J. Bland-Hawthorn 2002; A. Sandage & G. A. Tammann 2006; R. M. Catchpole et al. 2016). By studying the structure of a stellar population with a constrained age distribution we obtain a glimpse into the Galaxy’s dynamical history since that population’s birth. As different galaxy formation processes lead to different correlations between stellar age and dynamics, we can use these observed patterns to distinguish between formation scenarios (O. J. Eggen et al. 1962; L. Searle & R. Zinn 1978; K. C. Freeman 1987; M. Molla et al. 1996; I. R. Walker et al. 1996; J. Bland-Hawthorn & O. Gerhard 2016; A. Helmi 2020; A. J. Deason & V. Belokurov 2024).

Among the most valuable tracers for this purpose, RR Lyrae (RRL) are pulsating horizontal branch stars that are predominantly old ( $\geq 10$  Gyr) and metal-poor ( $[\text{Fe}/\text{H}] \lesssim -0.5$ ; H. A. Smith 1995), making them an effective probe for the Galaxy’s earliest evolution. Furthermore, they are an ideal population for studying Galactic structure because their

standard-candle nature gives precise distance estimates as low as 1%–2% when both infrared and optical photometry is available (B. F. Madore & W. L. Freedman 2012; J. R. Neeley et al. 2017; B. Sesar et al. 2017a; J. P. Mullen et al. 2023). Even in high-extinction regions, RRL are still excellent distance indicators because they are color standards; they have the same effective temperature and therefore intrinsic color during their minimum light phase (C. Sturch 1966; K. A. Guldenschuh et al. 2005; A. Kunder et al. 2010; A. K. Vivas et al. 2017).

These exceptional properties have made RRL valuable for studying many aspects of the Galaxy’s structure, including the Galactic disk. The disk is typically modeled with two components: the older thick disk and the younger thin disk (G. Gilmore & N. Reid 1983; R. F. G. Wyse & G. Gilmore 1995; B. E. Reddy et al. 2006; K. Fuhrmann 2008; T. Bensby et al. 2011). Given that they are generally old, RRL have been primarily used to study the Galactic thick disk (A. C. Layden 1995; S. Amrose & T. McKay 2001; K. Kinemuchi et al. 2006; C. Mateu et al. 2012). Overall, the thick-disk red-giant-branch and red-clump density distribution is consistent with results for RRL with most estimates converging around an exponential scale height ( $h_z$ ) of  $\approx 600$ –800 pc and scale length ( $h_R$ ) of  $\approx 2.1$  kpc (J. Bovy et al. 2012a, 2012b; A. C. Robin et al. 2014; J. Bovy et al. 2016b; C. Mateu & A. K. Vivas 2018; R. Tkachenko et al. 2025).

The relationship between RRL and the thin disk, however, has proven more complex. Although, metal-rich RRL stars with thin disk kinematics were found by A. C. Layden (1995)

and as early as P. Strugnell et al. (1986), other studies were unable to confirm the association of these stars to the thin disk and suggested they might belong to the metal-rich tail of the thick disk (J. C. Martin & H. L. Morrison 1998). Recent work has revealed a significant fraction of disk RRL with thin disk kinematics (G. Iorio & V. Belokurov 2021), sparking a debate about their origins (M. Chadid et al. 2017; V. A. Marsakov et al. 2018; V. A. Marsakov et al. 2019; Z. Prudil et al. 2020; R. Zinn et al. 2020). One hypothesis is that these stars have an accreted origin (D. K. Feuillet et al. 2022). Another possibility is that the thin disk RRL are actually an intermediate-age population and that it could be the result of binary evolution (G. Iorio & V. Belokurov 2021; A. Bobrick et al. 2024; M. Cabrera-Gadea et al. 2025; B. Cuevas-Otahola et al. 2025; H. Zhang et al. 2025).

Beyond the Galactic disk, the metal-poor RRL serve as ideal tracers for the stellar density distribution of the Galactic halo, which is thought to consist of ancient accreted galaxies (G. De Lucia & A. Helmi 2008; A. J. Deason et al. 2013). Generally, results for the halo density distribution of RRL are consistent with the total stellar distribution which are both most commonly modeled as a broken power law with a slope between  $\alpha \approx -1.5$  to  $-3$  for the inner halo and a slope of  $\alpha \approx -4$  for the outer halo with the break radius around  $\approx 20$ – $30$  kpc (B. Sesar et al. 2007, 2013; L. L. Watkins et al. 2009; A. J. Deason et al. 2011; L. Faccioli et al. 2014; R. Zinn et al. 2014; P. Das & J. Binney 2016; N. Hernitschek et al. 2018; G. Iorio et al. 2018; A. Chen et al. 2023; J. M. M. Lane et al. 2023). However, other works have found doubly broken power-law slopes with the second break radius of  $\approx 10$  kpc (J. J. Han et al. 2022; C. Yang et al. 2022). Recent works have also found the slope may be distinct along different lines of sight in the Galaxy (N. Hernitschek et al. 2018; J. A. S. Amarante et al. 2024; G. E. Medina et al. 2024). Works focusing on the very distant outer halo ( $R \gtrsim 100$  kpc) have found similar slopes of  $\alpha \approx -4$  (G. E. Medina et al. 2018; G. F. Thomas et al. 2018; K. M. Stringer et al. 2021; Y. Feng et al. 2024), although some have found a break radius at  $R \approx 160$ – $210$  kpc (T. Fukushima et al. 2018, 2019). None of these previous works include data within 5 kpc of the Galactic center with only a few of them including data with Galactocentric radii  $< 10$  kpc. There is also an interesting trend in the literature where works that include data at smaller radii tend to measure flatter slopes for the inner halo. Furthermore, previous works find that the halo is flattened with  $q \approx 0.6$ – $0.7$  where  $R = \sqrt{X^2 + Y^2 + (Z/q)^2}$  (A. J. Deason et al. 2011; G. Iorio et al. 2018; J. J. Han et al. 2022) with most recent work also measuring a tilt in the halo with respect to the Galactic disk plane (G. Iorio & V. Belokurov 2019; J. J. Han et al. 2022). There is also work which finds the flattening varies with Galactocentric radius (X.-X. Xue et al. 2015; N. Hernitschek et al. 2018; G. Iorio et al. 2018; Y. Xu et al. 2018).

Perhaps the most challenging region to characterize, the inner Galaxy contains many overlapping components (halo, disk, and bulge) and complex structures (the Galactic bar and X-shape). Whether the bulge is a distinct component from the inner halo is currently debated (M. Lucey et al. 2021; A. Ardern-Arentsen et al. 2024). P. Pietrukowicz et al. (2015) found the RRL density distribution in the inner Galaxy (0.2–3 kpc) is consistent with the measurements of the inner halo with a power law slope of  $\alpha \approx -3$  (A. Pérez-Villegas et al. 2017). There is also evidence that a population of RRL in

the inner Galaxy trace the bar structure (A. Kunder et al. 2020).

In this work, we simultaneously model the RRL density distribution of the Milky Way bulge, halo, and disk. In order to flexibly model all these structures, we utilize a Gaussian Mixture Model (GMM). GMMs have previously been used to model stellar density distributions of external galaxies, including to approximate power laws (M. Cappellari 2002; D. W. Hogg & D. Lang 2013; T. B. Miller & P. van Dokkum 2021). However, this work is the first time this method has been applied to measure the stellar density distribution of the Milky Way.

This publication is laid out as follows. In Section 2 we describe the sample of RRL, including parameterizing the selection function and calculating the distance modulus. We introduce our hierarchical Bayesian model of Gaussian mixtures in Section 3. We present our resulting density distribution of RRL and a comparison to previous work in Section 4. Last, in Section 5, we summarize this work and the conclusions.

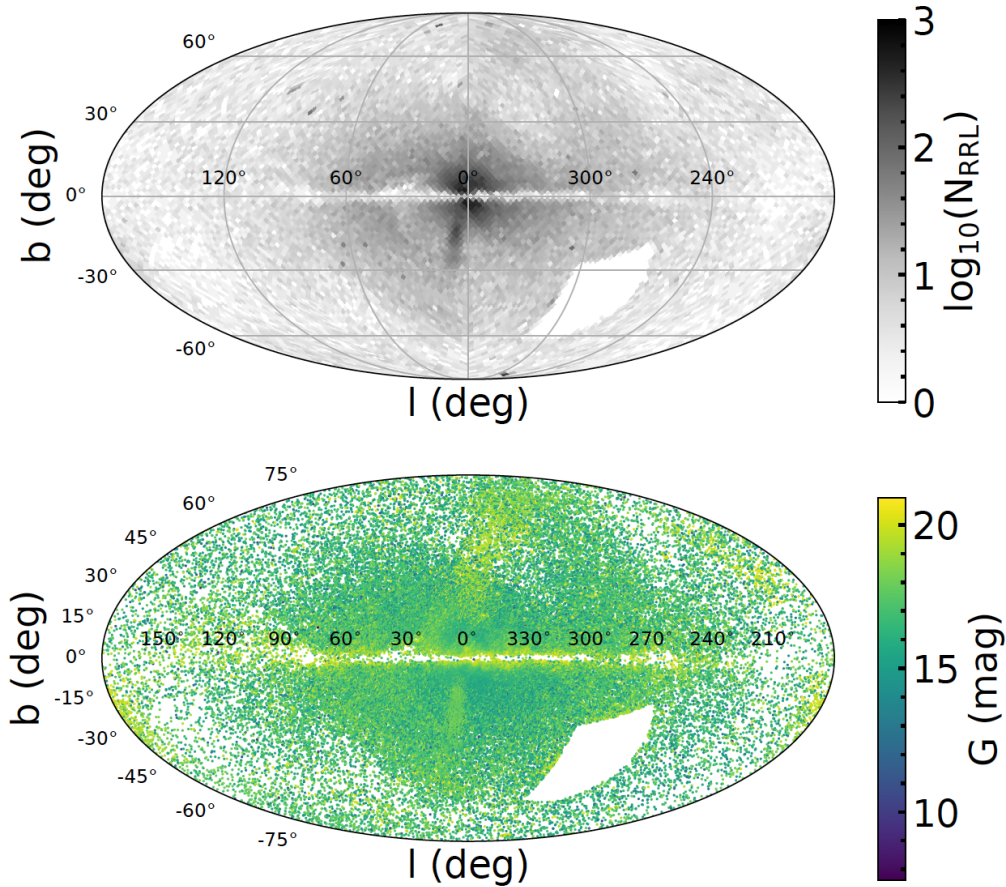
## 2. RR Lyrae Sample

Large samples of RRL stars are effective tools for studying Galactic structure given that their standard-candle nature provides precise distance estimates. In this work, we use a catalog compiled from the three largest public surveys of RRL stars: Gaia DR3 Specific Objects Study (SOS; G. Clementini et al. 2023) ASAS-SN (T. Jayasinghe et al. 2019) and PanSTARRS1 (PS1; B. Sesar et al. 2017b). The quality cuts applied to each catalog are described in detail in Section 2 of C. Mateu et al. (2020), as well as the cross-matching strategy used to match the known RRL from each catalog to a source in the Gaia DR3 `gaia_source` table. Duplicate sources were removed based on Gaia’s `source_id`, resulting in a base catalog with a total of 288,928 RRL.

Distances to the RRL stars were computed using the period–Wesenheit–metallicity relation from J. R. Neeley et al. (2019) for the Gaia reddening-free Wesenheit index  $W_{G, BP-RP}$ , and photometric metallicities from X.-Y. Li et al. (2023). These photometric metallicities were computed from the  $G$ -band light curve  $\phi_{31}$  Fourier coefficient and recalibrated by X.-Y. Li et al. (2023) to correct for systematic effects present in the photometric metallicities reported in the SOS catalog. Finally, the catalog was restricted to RRL stars with  $RUWE < 1.4$  and  $phot_{\{b\}\{p\}}_{rp\_excess\_factor} < 3$  to ensure, respectively, the quality of the astrometric parameters and of the BP/RP photometry used in the distance calculation; and filtering the catalog to stars with photometric metallicities available results in a final catalog of 128,353 RRL stars, comprised of 108,295, 19,986 and 72 RRL of types  $ab$  (fundamental mode),  $c$  (first overtone) and  $d$  (double mode). These stars have a heliocentric distance range of 0.2–140 kpc, corresponding to a galactocentric distance range of 0.2–120 kpc.

We present the sky distribution of our sample in Galactic coordinates ( $l, b$ ) in Figure 1. To remove contamination from the two most massive dwarf galaxies orbiting the Milky Way, the LMC, and SMC, we remove stars with R.A.  $< 110^\circ$  and decl.  $< -55^\circ$ . This is the simplest way to remove these major structures from the sample without complicated selection function effects from distance uncertainties. The sky distribution of the resulting 102,761 RRL is shown in the top panel of Figure 1.

The bottom panel of Figure 1 shows the sky distribution of our final sample colored by the observed  $G$  magnitude. The



**Figure 1.** The sky distribution in Galactic coordinates ( $l$ ,  $b$ ) of our RRL sample. The top panel shows the log density for the 105,325 RRL we use in this work, after removing the LMC and SMC with a cut in R.A. and decl. The bottom panel shows the sky distribution of our final sample colored by the Gaia  $G$  magnitude, which, because of their standard-candle nature, is strongly correlated with the distance.

Sagittarius dwarf galaxy stream is visible as a streak of stars with high  $G$  magnitudes passing vertically through the Galactic center at  $(l, b) = (0, 0)^\circ$ . However, as this is much lower density than the LMC and SMC we do not find it necessary to remove from the sample. Along the disk plane at  $b = 0^\circ$ , high levels of dust extinction cause higher  $G$  magnitudes as well as gaps in the RRL distribution. We account for this effect in our selection function which is described in the following Section.

### 2.1. Selection Function

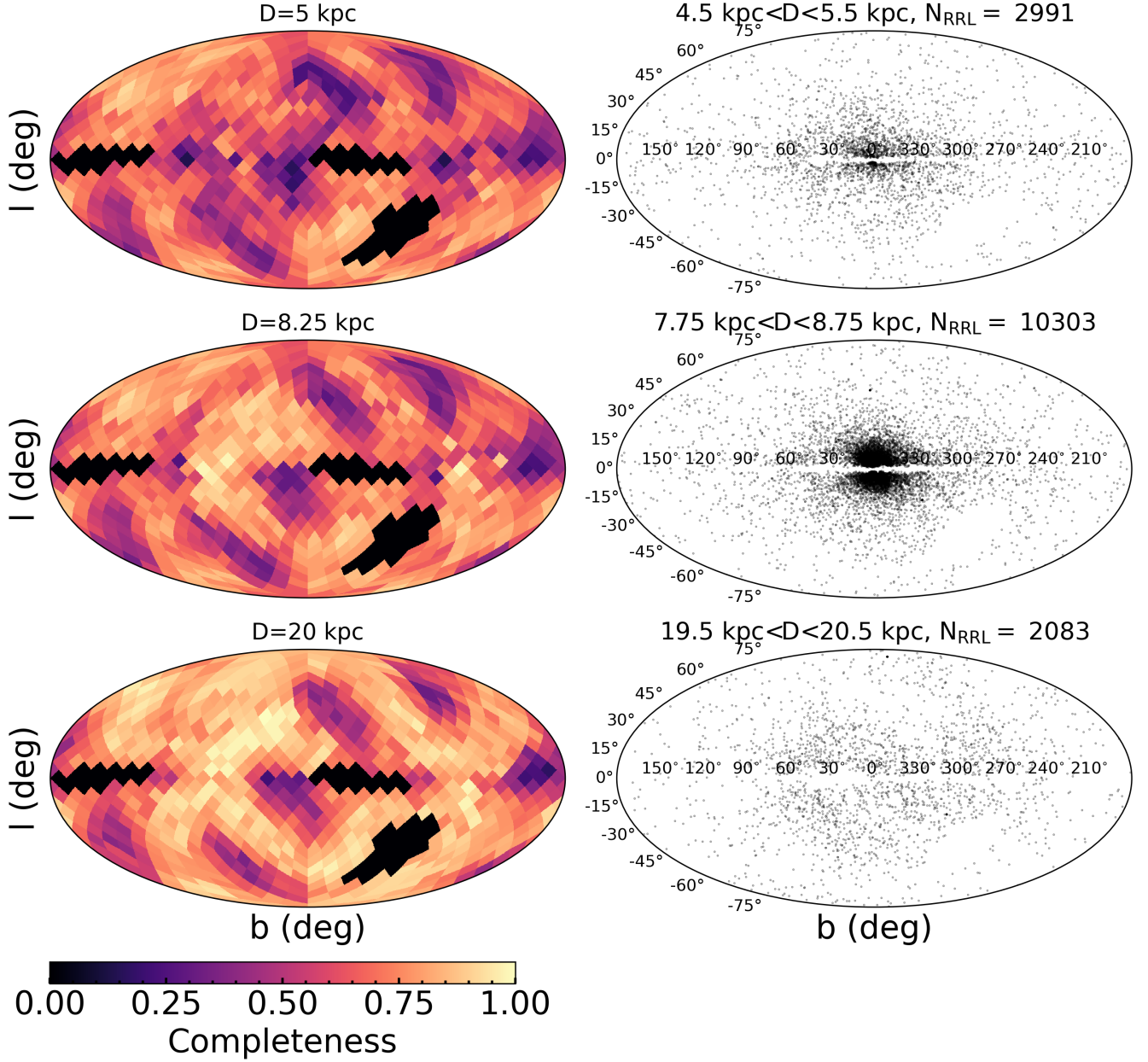
In order to account for the observational biases of our RRL sample, we need to quantify the selection function by estimating the completeness of our sample as a function of sky position and heliocentric distance. We estimate the completeness following the procedure described in detail in C. Mateu et al. (2020) and introduced by J. Rybizki & R. Drimmel (2018). The method uses two independent catalogs, in our case Gaia and ASAS+PS1, to empirically determine the completeness for each catalog based on the number of stars in both catalogs as well as in each. Completeness is estimated separately for RRL of types  $ab$  and  $c$ <sup>6</sup>, since RRLc samples are well known to be more affected by incompleteness as these stars have lower

photometric amplitudes and are more prone to misclassification (see, e.g., B. Sesar et al. 2017b; C. Mateu et al. 2020). In C. Mateu et al. (2020) completeness maps were presented for the Gaia DR2 SOS, ASAS-SN and PS1 RRL catalogs, and an update presented in C. Mateu (2024) for Gaia DR3 and the combined Gaia+ASAS+PS1. In the latter, completeness maps using only the RUWE and `phot_{b}{p}_rp_excess_factor` (BEP) filters were reported. Here we compute the completeness (or selection function) maps including also the more stringent availability filter on the X.-Y. Li et al. (2023) photometric metallicity (or equivalently Gaia DR3's). The resulting selection function maps with RUWE, BEP, and photometric metallicities filters are publicly available in C. Mateu (2025). The tools to generate this and other similar selection functions are also available at the `rll_completeness` GitHub repository.<sup>7</sup>

In the left panels of Figure 2 we present images of the resulting selection function in Galactic coordinates evaluated at three heliocentric distances (5, 10, and 20 kpc from top to bottom). The right panels show corresponding 1 kpc slices of the data. The dominant structure is a result of the Gaia scanning law (see Figure 10 in T. Cantat-Gaudin et al. 2023) where the completeness is lowest in regions with the fewest

<sup>6</sup> For these purposes we consider the very few RRLc as RRLab since they are too few for a completeness map to be estimated.

<sup>7</sup> [https://github.com/cmateru/rll\\_completeness](https://github.com/cmateru/rll_completeness)



**Figure 2.** Empirical Completeness Maps of the RRL sample in Galactic Coordinates at heliocentric distances of 5 kpc (top left), 10 kpc (middle left), and 20 kpc (bottom left). We show corresponding slices of the data in heliocentric distance bins of 1 kpc in the right panels. At small distances, the completeness is degraded because of the bright limits of surveys, while the faint limits impact the completeness at large heliocentric distances. The dominant structure in the completeness map is consistent with the Gaia scanning law.

scans. At small heliocentric distances, the completeness is lower due to the bright limits of the surveys, while the faint limits impact the completeness at large heliocentric distances.

### 3. Hierarchical Bayesian Model of Multi-Gaussian Galaxy

The aim of this work is to find the maximum likelihood estimate for a given RRL density model described by parameters,  $\theta$ , given our RRL distance catalog and selection function ( $S(l, b, d)$ ). Specifically, we use a hierarchical Bayesian model to infer the true underlying stellar density distribution  $n(\mathbf{x}; \theta)$ , where  $\theta$  are the parameters of our density model. We transform true 3D Galactocentric positions  $\mathbf{x}$  to observed distances ( $d$ ) and sky positions ( $l, b$ ) and marginalize

over the distance uncertainty as follows:

$$\begin{aligned} \mathcal{L}(\theta) &\propto \prod_{i=1}^N P(l_i, b_i, d_i | \theta, \mathbf{x}_i) S(l_i, b_i, d_i) \\ &\propto \prod_{i=1}^N \int_{d_{min}}^{d_{max}} dx_i P(l_i, b_i, d_i | \mathbf{x}_i) P(\mathbf{x}_i | \theta) S(l_i, b_i, d_i) \end{aligned} \quad (1)$$

where

$$P(\mathbf{x}_i | \theta) = \frac{1}{N} n(\mathbf{x}_i; \theta) \quad (2)$$

is a normalized version of our density function, and the index  $i$  tracks each of the  $N$  RRL stars we use as data.

To perform the transformation to and from Galactocentric positions, we assume a right-handed coordinate system with the Sun’s position at  $(-8.25 \text{ kpc}, 0, 20.8 \text{ pc})$ ; GRAVITY Collaboration et al. 2021; H. W. Leung et al. 2023). The measured sky positions  $(l, b)$  of are significantly more precise than the distances  $(d)$  so the integral over 3D position simply becomes a 1D integral over the distance. For computational efficiency, this integral is calculated via a Riemann sum from  $d_{min}$  to  $d_{max}$ , which are defined as follows:

$$d_{min} = \min \left\{ \begin{array}{l} d - 3\sigma_d \text{ [kpc]} \\ 10^{-3} \text{ kpc} \end{array} \right. \quad (3)$$

$$d_{max} = d + 3\sigma_d \text{ [kpc]} \quad (4)$$

where  $\sigma_d$  is the uncertainty on the distance. The choice of  $3\sigma_d$  in the integration bounds is arbitrary, but sufficient given the Gaussian distance uncertainties, as this captures 99.7% of the probability density. To prevent negative distance values in the case of  $3\sigma_d > d$ , a minimum distance of  $10^{-3} \text{ kpc}$  is chosen as an adequately small bound given that it is two orders of magnitude less than the smallest distance in our sample (0.28 kpc).

The  $P(\mathbf{x}_i|\boldsymbol{\theta})$  term in Equation 1 is our density model of the Galaxy. In this work, we use a GMM to describe the Galactic stellar density because it is a normalized probability density distribution unlike many of the power law models typically used to describe stellar halos (e.g., A. J. Deason et al. 2011; G. Iorio et al. 2018; J. J. Han et al. 2022). Furthermore, GMMs can accurately approximate many power law functions with only a few components (M. Cappellari 2002; D. W. Hogg & D. Lang 2013; T. B. Miller & P. van Dokkum 2021). While the choice of GMM is not physically motivated, it offers numerous mathematical advantages (O. Bendinelli 1991; J. Bovy et al. 2011), including an analytical solution to the Jean’s equations (M. Cappellari 2008), and better fits to real data where simple parametric models fail to recover the full complexity of galactic structure (E. Emsellem et al. 1994).

We fit two GMMs which are then combined with weights that sum to 1 as our complete density model. All Gaussian components of our model are centered on  $(0,0,0) \text{ kpc}$  corresponding to the Galactic center. One of the GMMs is set to model the inner Galaxy, while the other models the halo. For both, we set the covariance matrices so that they increase by a factor of two between components. The difference between the inner Galaxy and halo model lies in the scaling factor  $(k)$  applied to the covariance matrices. For the inner Galaxy GMM  $k = 0.1 \text{ kpc}$ , while  $k = 1.5 \text{ kpc}$  for the halo GMM. The inner Galaxy GMM has five components while the halo GMM has 6. Using 6–11 total GMM components with logarithmically scaled widths is standard for modeling galaxy profiles (M. Cappellari 2002; D. W. Hogg & D. Lang 2013; T. B. Miller & P. van Dokkum 2021). Our method adds additional flexibility in allowing the halo to be flattened in a direction different than in the inner Galaxy. To allow flexibility in the model for exactly where the halo GMM begins to dominate the density profile, we ensure overlap in that the smallest halo GMM component has  $\sigma_{xy}=1.5 \text{ kpc}$  and the largest inner Galaxy component has  $\sigma_{xy} = 1.6 \text{ kpc}$ . Each GMM also includes its own flattening parameter  $(q)$  applied to the  $z$ -coordinate, which is nominally perpendicular to the Galactic disk. Therefore, for both GMMs, the nonrotated

covariance matrix for the  $j$ th component is as follows:

$$cov_j = \begin{bmatrix} \sigma_{xy}^2 & 0 & 0 \\ 0 & \sigma_{xy}^2 & 0 \\ 0 & 0 & \sigma_z^2 \end{bmatrix} = \begin{bmatrix} k2^j & 0 & 0 \\ 0 & k2^j & 0 \\ 0 & 0 & kq2^j \end{bmatrix}. \quad (5)$$

However, there is increasing evidence the the Galactic halo of the Milky Way may be rotated with respect to the Galactic disk (G. Iorio & V. Belokurov 2019; J. J. Han et al. 2022; J. Nibauer & A. Bonaca 2025). In other words, the Galactic halo may not be flattened in a direction directly perpendicular to the Galactic disk ( $z$ ) but instead along another direction. To investigate this, we develop a rotated halo model. Specifically, we transform the halo components’ covariance matrices with the following rotation matrix:

$$R_{rot} = \begin{bmatrix} \cos \phi \cos \alpha & -\sin \phi & \cos \phi \sin \alpha \\ \sin \phi \cos \alpha & \cos \phi & \sin \phi \sin \alpha \\ -\sin \alpha & 0 & \cos \alpha \end{bmatrix} \quad (6)$$

where the new  $z$ -axis now points in the direction of the vector  $[\sin \alpha \cos \phi, \sin \alpha \sin \phi, \cos \alpha]$  in the nonrotated coordinate frame.

In total, we use the following function form for the GMM of the density of RRL:

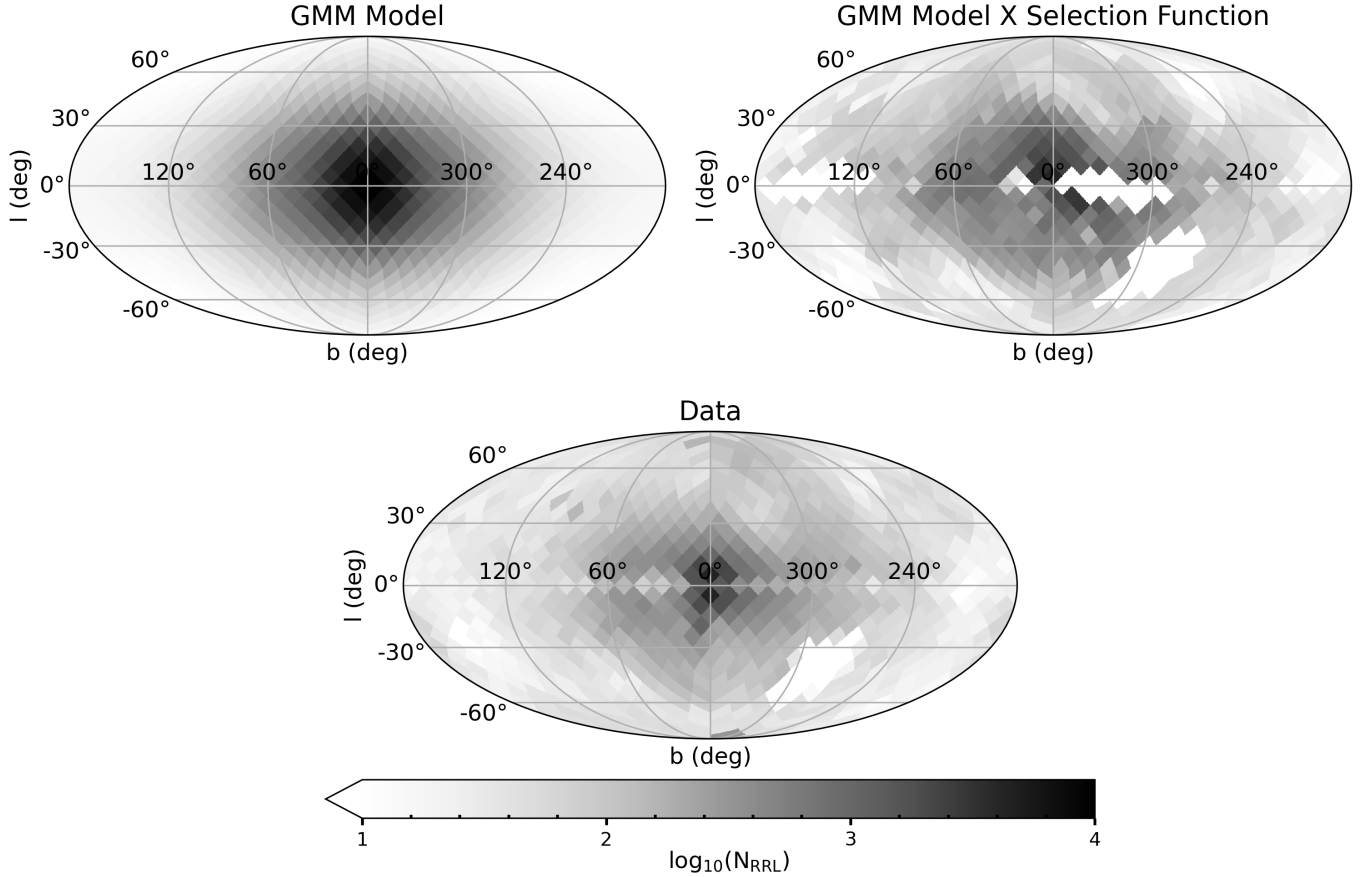
$$\begin{aligned} n(\mathbf{x}_i|\boldsymbol{\theta}) &= (1 - A) \sum_{j=1}^6 A_{H,j} \mathcal{N}(\mathbf{x}_i|\boldsymbol{\theta}, R_{rot} cov_j(k = 2.0 \text{ kpc}) R_{rot}^T) \\ &+ A \sum_{j=1}^5 A_{IG,j} \mathcal{N}(\mathbf{x}_i|\boldsymbol{\theta}, cov_j(k = 0.1 \text{ kpc})) \end{aligned} \quad (7)$$

where  $A$  is the amplitude for the inner Galaxy GMM,  $A_{H,j}$  is the amplitude for each halo component and  $A_{IG,j}$  is the amplitude for each inner Galaxy component.

We utilize the stick breaking transform to fit the amplitudes of the GMM components. For complete details of the transform we refer the reader to Stan Development Team (2024). In short, it transforms  $J$  GMM amplitudes to  $J - 1$  unconstrained parameters. This leads to better convergence because it reduces the dimensionality and ensures the amplitudes sum to 1. With this transform, our model has a total of 14 free parameters (ten amplitudes, two flattening parameters  $(q)$ , and two angles) to fit.

### 3.1. Inference

As our hierarchical Bayesian model is computationally expensive with high dimensionality (14 free parameters), we require fast, scalable optimizers and posterior estimators. For fast maximum likelihood estimates, we use an the Adafactor optimizer to find the minimum of the negative log-likelihood. We choose Adafactor because it has similar benefits to the Adam optimizer (efficient stochastic gradient descent) with lower memory cost (N. Shazeer & M. Stern 2018), which is crucial in this work given that all of the data has to be held in memory at once in order to calculate the full likelihood. To estimate the uncertainties, we calculate the covariance matrix at the maximum likelihood estimate (MLE) using the negative inverse Hessian.



**Figure 3.** Comparison of the sky projected densities of our model to the data in Galactic coordinates ( $l, b$ ). The top left panel is sky projection of our model where the darkest patches indicate the highest density of RRL. The top-right panel shows the same but convolved with our model for the selection function. The bottom panel shows the equivalent sky projection of the RRL sample used in this work. In general, the sky projected model is similar to the data when convolved with the selection.

**Table 1**  
Inferred Parameters of GMM Model

	Inner Galaxy	Halo
$q$	$1.311 \pm 0.002$	$0.697 \pm 0.002$
$\alpha$ ( $^\circ$ )	N/A	$17.7 \pm 0.7$
$\phi$ ( $^\circ$ )	N/A	$-27.2 \pm 2.4$
Amplitudes for GMM		
A	0.37	
$j =$	$A_{IG,j}$	$A_{H,j}$
1	$7.00 \times 10^{-2}$	$9.99 \times 10^{-5}$
2	$6.53 \times 10^{-2}$	$1.14 \times 10^{-4}$
3	$6.03 \times 10^{-2}$	$5.60 \times 10^{-2}$
4	$5.05 \times 10^{-2}$	$7.70 \times 10^{-1}$
5	$7.53 \times 10^{-1}$	$1.32 \times 10^{-1}$
6	...	$4.25 \times 10^{-2}$

**Note.** The parameters of our inferred GMM model, including the flattening ( $q$ ) and rotation angles ( $\alpha, \phi$ ) as well as the amplitudes for each GMM component. The amplitude for the inner Galaxy GMM ( $A$ ; see Equation 7) and its five components ( $A_{IG,j}$ ) along with the six halo amplitudes ( $A_{H,j}$ ) are all given.

#### 4. The Galactic Density Distribution of RR Lyrae

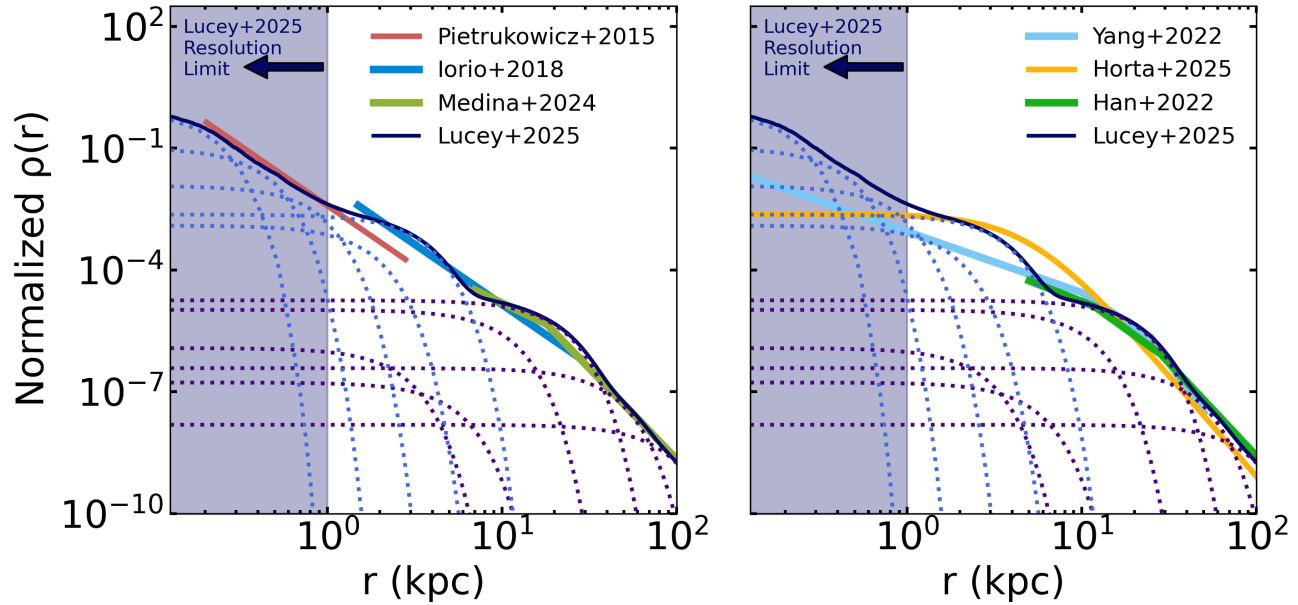
We present the main result of our work, the Galactic distribution of RRL, in the form of the MLE of our 16-parameter

GMM. The MLE and associated  $1\sigma$  uncertainties for the halo flattening and rotation are shown in Table 1 as the parameters that are directly comparable to previous works (e.g., G. Iorio et al. 2018; J. J. Han et al. 2022; J. Nibauer & A. Bonaca 2025).

In order to visually assess the accuracy of our model, we compare the predicted sky projected density of RRL from our model to the observed sample. Figure 3 shows three panels of sky projected density distribution with darker shading corresponding to higher densities. The top-left panel is calculated by simply integrating our model along each line of sight. The top-right panel shows the full forward model of the data by convolving the GMM with the empirical selection function. When compared to the sky projection of the observed data (bottom panel), we find generally qualitative agreement. As the precision of the empirical selection function estimate depends on the sample size in each healpix and distance bin, it is worst close to the Galactic plane where dust extinction severely limits the RRL sample as shown in Figures 1 and 2. Larger RRL catalogs are required to improve these results.

##### 4.1. Comparison to Literature Halo Models

Galactic halos are commonly described with power-law profiles, including cored, (doubly) broken, triaxial, and rotated adaptations (A. J. Deason et al. 2011; G. Iorio et al. 2018; J. J. Han et al. 2022). However, in this work, we use a GMM



**Figure 4.** The normalized halo density profile of RRL as a function of distance from the Galactic center ( $r$ ) for our GMM compared to literature results using RRL (left) and other tracers (right). In both panels, the light purple dotted lines indicate the five components of the inner Galaxy mixture model while the dark purple dotted lines show the halo components. The dark blue solid line is their sum. The literature RRL profiles included in the left panel are P. Pietrukowicz et al. (2015; red), G. Iorio et al. (2018; blue), and G. E. Medina et al. (2024; green). The works in the right panel, D. Horta & R. P. Schiavon (2025; yellow), J. J. Han et al. (2022; green), and C. Yang et al. (2022; light blue), use red-giant stars.

as an easily normalizable probability density and more flexible alternative to the power-law. This flexibility is especially crucial for modeling the inner Galaxy where the disk, halo, and bulge all overlap.

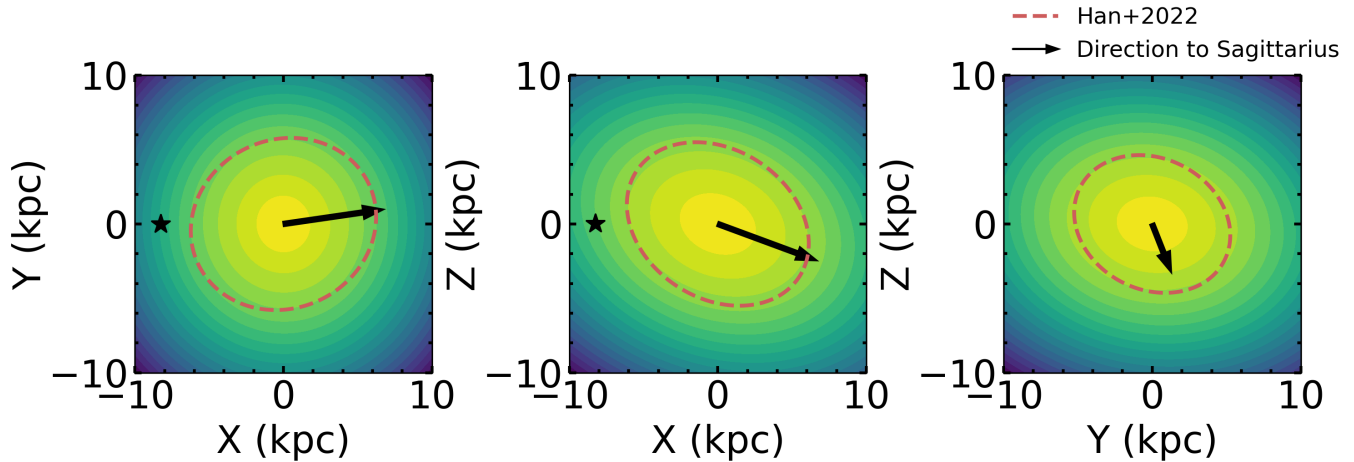
We show the normalized density as a function of Galactocentric radius ( $r$ ) for the halo components of our GMM Galaxy in Figure 4, compared to results from the literature using RRL in the left panel and other tracers in the right panel. In each panel, the five Gaussian halo components for the inner Galaxy GMM and the halo GMM are shown in light purple and dark purple dotted lines, respectively, with the sum shown in the solid dark blue line. In the left panel, the RRL profiles from the literature include results from P. Pietrukowicz et al. (2015; red), G. Iorio et al. (2018; blue), and G. E. Medina et al. (2024; green). These results are chosen to cover a wide range of radii, and includes both single and double power-law results, prioritizing recent work. The right panels includes the works of C. Yang et al. (2022; light blue), J. J. Han et al. (2022; green) and D. Horta & R. P. Schiavon (2025; yellow).

The literature results shown in the right panel are chosen to highlight the diversity of profiles measured using different stellar populations and methodologies to trace the stellar halo. For example, C. Yang et al. (2022) estimates the halo profile using direct orbit integration of K-giant stars with galactocentric distances between 5 and 50 kpc. They remove disk stars using a selection in  $[\text{Fe}/\text{H}]$  and cylindrical velocity. Similar to our method, J. J. Han et al. (2022) use a Bayesian model to infer the halo profile using a sample of stars selected to be from the Gaia–Enceladus–Sausage merger (V. Belokurov et al. 2018; A. Helmi et al. 2018) using  $[\alpha/\text{Fe}]$  and  $[\text{Fe}/\text{H}]$  abundances as well as orbital eccentricity. D. Horta & R. P. Schiavon (2025) uses a similar method to J. J. Han et al. (2022); however, they utilize a

selection in  $[\text{Mn}/\text{Mg}]$ ,  $[\text{Al}/\text{Fe}]$ ,  $[\text{Mg}/\text{Fe}]$ , and  $[\text{Fe}/\text{H}]$  to isolate the “Heracles/protogalaxy” structure that is generally confined to the inner Galaxy with apocenters  $< 12$  kpc. Our work improves upon these previous results by not performing a chemical or dynamical selection and instead directly modeling all of the available data including stars with Galactocentric radii as small as  $\approx 0.2$  kpc.

At Galactic radii  $> 12$  kpc, our results are generally consistent with previous results of a roughly  $r^{-4}$  power law, measured both with RRL and other tracers (A. J. Deason et al. 2011; G. Iorio et al. 2018; J. J. Han et al. 2022; C. Yang et al. 2022; G. E. Medina et al. 2024). Within  $r \approx 12$  kpc, previous works have also found a flattening in the halo profile, with the power law decreasing to  $r^{-1.5}$  (J. J. Han et al. 2022; C. Yang et al. 2022; G. E. Medina et al. 2024). Within  $\approx 5$  kpc of the Galactic center, we find that the inner Galaxy GMM begins to dominate the stellar density. We note that the amplitudes of the halo GMM components at these radii are essentially zero, indicating that the dominant RRL population within 5 kpc of the Galactic center is not consistent with the rotated halo population.

At Galactic radii  $> 1$  kpc and  $< 5$  kpc, we find a bump in the RRL density distribution. While the halo RRL slope has flattened between 5 and 12 kpc, the onset of the inner Galaxy RRL population causes a jump in the RRL distribution which again flattens approaching 1 kpc. Within 1 kpc of the Galactic center our model agrees with results from P. Pietrukowicz et al. (2015). However, our work is limited by selection function resolution at very small Galactic radii. Given our RRL sample size, the highest angular resolution we can achieve for estimating the empirical selection function is  $7.32$ . This corresponds to  $\approx 1$  kpc at the Galactic center assuming  $R_{\odot} = 8.25$  kpc. From Figure 3, it is also clear that our selection



**Figure 5.** The inferred Galactic distribution of RRL as a GMM. The face-on distribution evaluated at  $Z = 0$  kpc is shown in the left panel. The middle and right panels show the edge-on distributions evaluated at  $X = 0$  kpc and  $Y = 0$  kpc, respectively. The location of the Sun is shown as a black star in the left and middle panels. In all panels, we include an ellipse (red-dashed line) which demonstrates the shape and tilt of the halo model from J. J. Han et al. (2022). We also show the direction to Sagittarius which is located at (17.9, 2.6,  $-6.6$ ) kpc (E. Vasiliev et al. 2021).

function is least accurate closest to the Galactic plane. However, our results robustly disagree with extrapolating this profile out to 5 kpc as suggested by A. Pérez-Villegas et al. (2017).

In addition to comparing the slope of our halo model to previous works, we can also compare the rotation. Specifically, we compare our model to the tilted halo model of J. J. Han et al. (2022), which was recently confirmed by an independent analysis in J. Nibauer & A. Bonaca (2025). In these works, they use a triaxial halo model, and relatedly, a different rotation matrix since their model does not have  $XY$ -symmetry, unlike ours. However, we can directly compare the pitch angles which both describe the tilt of the  $z$ -axis toward the positive  $x$ -axis. We find a pitch angle of  $18^\circ$ , in agreement with the halo mass distribution of J. Nibauer & A. Bonaca (2025) and stellar density distribution of J. J. Han et al. (2022). It is interesting to note that J. J. Han et al. (2022) specifically measure the stellar distribution of the Gaia–Sausage–Enceladus system, which makes up the bulk of the local halo (V. Belokurov et al. 2018; A. Helmi et al. 2018).

To visually compare our rotated halo with that of J. J. Han et al. (2022), we plot density contours of the halo GMM in Figure 5, with the highest density in yellow and the lowest density in dark blue. The left panel shows the face-on slice evaluated at  $Z = 0$  kpc, while the middle and right panel show edge-on slices at  $X = 0$  kpc and  $Y = 0$  kpc, respectively. To guide the reader, we show the Sun’s position (black star) in the left and middle panels at  $X = -8.25$  kpc. The shape and tilt of our halo model agrees well with results from J. J. Han et al. (2022) shown as a red-dashed line, especially in the middle panel, which shows the  $X$ – $Z$  slice. The left and right panels show slight disagreement with the J. J. Han et al. (2022) results, but this is likely because of the difference in rotation matrices, where our model preserves  $X$ – $Y$  symmetry while J. J. Han et al. (2022) does not.

In Figure 5, we also show an arrow that points to the location of the Sagittarius dwarf galaxy at (17.9, 2.6,  $-6.6$ ) kpc (E. Vasiliev et al. 2021). This aligns well with the rotated major axis of the halo models. Sagittarius is a clear non-axisymmetric feature in our observed data (see Figure 2) so it is not unexpected that it could cause the tilt in our inferred

halo model. J. J. Han et al. (2022) targets modeling the halo contribution of only the Gaia–Enceladus–Sausage. Whether the tilt is caused by contamination from Sagittarius stars, or is a strange coincidence requires further investigation, including ongoing theoretical work testing the impact of assuming a single halo model when a galaxy is undergoing a Sagittarius-like merger versus explicitly modeling the second, less massive galaxy. J. Nibauer & A. Bonaca (2025) models the dark matter halo tilt, however they assume an axisymmetric subtraction of stellar mass that does not include Sagittarius. This could cause the elongation of the “missing” mass model toward the Sagittarius system.

For the inner Galaxy component, we measure an oblate shape with  $q = 1.31$ . This is consistent with works measuring the shape of the dark matter halo within 20 kpc of the Galactic center using kinematic tracers (J. Bovy et al. 2016a; A. Bowden et al. 2016), including the work of L. Posti & A. Helmi (2019), which also found  $q = 1.3$ . However, these works assume an oblate stellar halo or bulge for the nondisk component of the inner Galaxy. Further work to simultaneously study the stellar and dark matter mass distribution in the inner Galaxy is required to determine whether the RRL are tracing the distribution of dark matter as predicted by cosmological zoom-in simulations (M. Lucey et al. 2025).

## 5. Summary and Conclusions

The stellar density distribution of the Milky Way provide clues to its formation and evolutionary history. Among the most challenging regions to characterize is the inner Galaxy where there are high levels of crowding and dust extinction. Yet, as the oldest component, disentangling its structure reveals crucial insight to the earliest epochs of galaxy evolution, including the role of dark matter (K. El-Badry et al. 2018; H.-W. Rix et al. 2022; M. Lucey et al. 2025).

To address these observational challenges, we use the standard-candle nature of RRL stars to map the structure of the Milky Way with a hierarchical Bayesian GMM. This is the first time a GMM has been used to describe the stellar density distribution of the Milky Way, building on the success of this method in characterizing external galaxies including the power laws typically used to describe halos (M. Cappellari 2002;

D. W. Hogg & D. Lang 2013; T. B. Miller & P. van Dokkum 2021). Crucially, we find the GMM provides the increased flexibility needed to model the inner Galaxy where the bulge, halo, and disk all intersect.

We use a total of 11 Gaussian components to model the Galactic distribution of RRL. All of the components are centered on the Galactic center. Five of the components are set to model the inner Galaxy with XY-covariances between 0.1 and 1.6 kpc and allowing the covariance in the z-direction to differ by a factor of  $q$ . We use a similar setup for the six components that model the halo, but with larger covariances (1.5–48 kpc) and allowing the z-axis to be rotated with respect to the Galactic disk. This framework yields three key findings about the Galactic distribution of RRL:

1. The inner Galaxy ( $R \lesssim 10$  kpc) is dominated by distinct prolate population with  $q = 1.31$ , consistent with results for the shape of the dark matter distribution in the inner Galaxy (A. Bowden et al. 2016; L. Posti & A. Helmi 2019).
2. The halo distribution is best described by a  $r^{-4}$  power-law that flattens within 12 kpc of the Galactic center.
3. We infer a flattening ( $q = 0.70$ ) and halo tilt in the RRL population ( $\alpha = 18^\circ$ ) consistent with J. J. Han et al. (2022) and J. Nibauer & A. Bonaca (2025). We note for the first time that the rotated major-axis points to the Sagittarius dwarf galaxy.

These results demonstrate that GMMs offer a powerful and flexible framework for modeling complex stellar distributions in the Milky Way. The main limit to our work is the sample size, especially close to the Galactic plane where we are severely limited by dust extinction. This especially impacts the selection function resolution and precision. Larger samples of RRL with well-calibrated distances, especially in high-extinction regions will greatly improve this work, which the future Roman Galactic Plane Survey should provide (R. E. Sanderson et al. 2024). Looking ahead, the next work in this series will apply this method to study how the halo structure varies as a function of stellar metallicity. As the relationship between the stellar halo’s metallicity and structure is indicative of its origins (O. J. Eggen et al. 1962; L. Searle & R. Zinn 1978), such studies will provide insight into the formation history and chemical evolution of our Galaxy.

### Acknowledgments

This material is based upon work supported by the National Science Foundation under Award No. 2303831.

This work used Bridges-2 at Pittsburgh Supercomputing Center through allocation PHY250006 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program, which is supported by U.S. National Science Foundation grants #2138259, #2138286, #2138307, #2137603, and #2138296. This work has made use of data from the European Space Agency (ESA) mission

Gaia (<https://www.cosmos.esa.int/gaia>), processed by the Gaia Data Processing and Analysis Consortium (DPAC; <https://www.cosmos.esa.int/web/gaia/dpac/consortium>).

Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement.

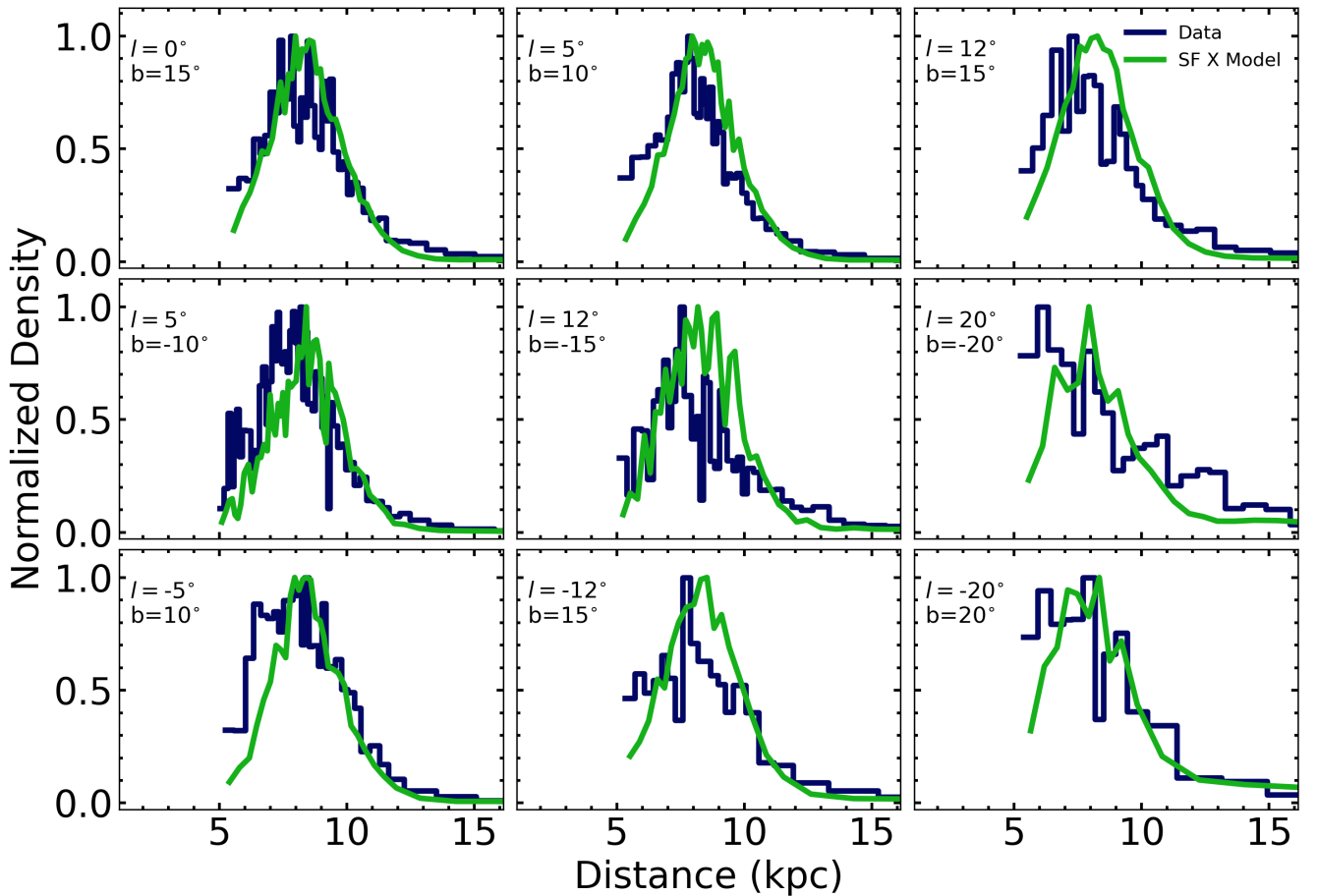
*Software:* Astropy (Astropy Collaboration et al. 2013, 2018), Matplotlib (J. D. Hunter 2007), IPython (F. Pérez & B. E. Granger 2007), Numpy (C. R. Harris et al. 2020), Scipy (P. Virtanen et al. 2020), JAX (J. Bradbury et al. 2018), OTT-JAX (M. Cuturi et al. 2022), Optax (DeepMind et al. 2020), JAXopt (M. Blondel et al. 2021), BlackJAX (A. Cabezas et al. 2024), Corner (D. Foreman-Mackey 2016), luas (M. Fortune et al. 2024).

## Appendix A

### Comparing the Data to the Inferred Model

To further assess the quality of our model fit to the observed data, we investigate the observed and predicted distribution of distances for a given line of sight in Figure 6. Given the complex selection function effects and distance uncertainties which are accounted for in our statistical model, it can be hard to directly compare the observed data distribution to our inferred model. This figure does not include the effect of our distance uncertainties, but we do show the predicted distribution of distances given our selection function and the model (green). For both the observed data (dark blue) and the model prediction, the binning matches that of the selection function calculation which is chosen so that for each healpix each distance bin has the same number of stars for the completeness calculation. The lines of sight chosen are all toward the Galactic center where the majority of our data lies. However, we do not show any lines of sight where  $l < 0^\circ$  and  $b < 0^\circ$ , because the completeness is so low (see Figure 2) and there are two few stars.

In general, the selection function modified model fits the data well, especially given that the distance uncertainties. However, this figure also demonstrates how the accuracy and precision of our model is limited by the data and selection function calculation. While our model marginalizes over the distance uncertainties, we assume symmetric distance uncertainties and therefore in cases where we may be systematically underestimating the distances because of dust extinction toward the Galactic center, this will bias our results. In many of the panels, the peak in the observed data is at lower distances than the model and our data distribution contains more RRL at small distances than expected given the data and selection function. This is likely due to either the selection function underestimating the completeness at these distances or these distances being underestimated due to dust extinction. Larger RRL catalogs with improved distance estimates are required to improve these results.



**Figure 6.** The number density of the data (dark blue) compared to the selection function modified model (green) as a function of distance from the Sun along various lines of sight. The healpix and distance bins are set by the calculation of the selection function. Specifically, the distance bins are chosen so that for each line of sight each distance bin has the same number of stars used in the selection function calculation.

## Appendix B RRL Catalog

The RR Lyrae distance catalog is shown in Table 2.

**Table 2**  
RRL Distance Catalog

Gaia DR3 ID	R.A. (deg)	Decl. (deg)	$l$ (deg)	$b$ (deg)	D (kpc)	$\sigma_D$ (kpc)
2125982599343482624	291.365640	42.783489	74.951410	12.304166	0.2841	0.0209
6005656897473385600	230.605734	-40.926801	331.330985	13.473541	0.3422	0.0236
524535426630702464	12.971012	65.180555	122.978775	2.308846	0.4263	0.0294
2211629018927324288	339.805873	64.859351	109.470470	5.501704	0.4482	0.0310
2857456211775108480	5.929506	29.400923	115.724916	-33.082618	0.4831	0.0337

(This table is available in its entirety in machine-readable form in the [online article](#).)

### ORCID iDs

Madeline Lucey <https://orcid.org/0000-0001-7297-8508>  
 Cecilia Mateu <https://orcid.org/0000-0002-6330-2394>  
 Adrian M. Price-Whelan <https://orcid.org/0000-0003-0872-7098>  
 David W. Hogg <https://orcid.org/0000-0003-2866-9403>  
 Hans-Walter Rix <https://orcid.org/0000-0003-4996-9069>  
 Robyn E. Sanderson <https://orcid.org/0000-0003-3939-3297>

### References

Amarante, J. A. S., Koposov, S. E., & Laporte, C. F. P. 2024, *A&A*, **690**, A166  
 Amrose, S., & McKay, T. 2001, *ApJL*, **560**, L151  
 Ardern-Arentsen, A., Monari, G., Queiroz, A. B. A., et al. 2024, *MNRAS*, **530**, 3391  
 Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, **156**, 123  
 Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, **558**, A33  
 Belokurov, V., Erkal, D., Evans, N. W., Koposov, S. E., & Deason, A. J. 2018, *MNRAS*, **478**, 611

- Bendinelli, O. 1991, *ApJ*, 366, 599
- Bensby, T., Alves-Brito, A., Oey, M. S., Yong, D., & Meléndez, J. 2011, *ApJL*, 735, L46
- Bland-Hawthorn, J., & Gerhard, O. 2016, *ARA&A*, 54, 529
- Blondel, M., Berthet, Q., Cuturi, M., et al. 2021, arXiv:2105.15183
- Bobrick, A., Iorio, G., Belokurov, V., et al. 2024, *MNRAS*, 527, 12196
- Bovy, J., Bahmanyar, A., Fritz, T. K., & Kallivayalil, N. 2016a, *ApJ*, 833, 31
- Bovy, J., Hogg, D. W., & Roweis, S. T. 2011, *AnApS*, 5, 1657
- Bovy, J., Rix, H.-W., & Hogg, D. W. 2012a, *ApJ*, 751, 131
- Bovy, J., Rix, H.-W., Liu, C., et al. 2012b, *ApJ*, 753, 148
- Bovy, J., Rix, H.-W., Schlafly, E. F., et al. 2016b, *ApJ*, 823, 30
- Bowden, A., Evans, N. W., & Williams, A. A. 2016, *MNRAS*, 460, 329
- Bradbury, J., Frostig, R., Hawkins, P., et al. 2018, JAX: composable transformations of Python+NumPy programs, v0.3.13, <http://github.com/jax-ml/jax>
- Cabezas, A., Corenflos, A., Lao, J., & Louf, R. 2024, arXiv:2402.10797
- Cabrera-Gadea, M., Mateu, C., & Ramos, P. 2025, *A&A*, 701, A136
- Cantat-Gaudin, T., Fouesneau, M., Rix, H.-W., et al. 2023, *A&A*, 669, A55
- Cappellari, M. 2002, *MNRAS*, 333, 400
- Cappellari, M. 2008, *MNRAS*, 390, 71
- Catchpole, R. M., Whitelock, P. A., Feast, M. W., et al. 2016, *MNRAS*, 455, 2216
- Chadid, M., Sneden, C., & Preston, G. W. 2017, *ApJ*, 835, 187
- Chen, A., Li, Z., Wang, Y., et al. 2023, *MNRAS*, 525, 3075
- Clementini, G., Ripepi, V., Garofalo, A., et al. 2023, *A&A*, 674, A18
- Cuevas-Otahola, B., Mateu, C., Cabrera-Ziri, I., et al. 2025, *MNRAS*, 541, 1434
- Cuturi, M., Meng-Papaxanthos, L., Tian, Y., et al. 2022, arXiv:2201.12324
- Das, P., & Binney, J. 2016, *MNRAS*, 460, 1725
- De Lucia, G., & Helmi, A. 2008, *MNRAS*, 391, 14
- Deason, A. J., & Belokurov, V. 2024, *NewAR*, 99, 101706
- Deason, A. J., Belokurov, V., & Evans, N. W. 2011, *MNRAS*, 416, 2903
- Deason, A. J., Belokurov, V., Evans, N. W., & Johnston, K. V. 2013, *ApJ*, 763, 113
- DeepMind, Babuschkin, I., Bauml, K., et al. 2020, The DeepMind JAX Ecosystem, <http://github.com/google-deeppmind>
- Eggen, O. J., Lynden-Bell, D., & Sandage, A. R. 1962, *ApJ*, 136, 748
- El-Badry, K., Bland-Hawthorn, J., Wetzell, A., et al. 2018, *MNRAS*, 480, 652
- Emsellem, E., Monnet, G., & Bacon, R. 1994, *A&A*, 285, 723
- Faccioli, L., Smith, M. C., Yuan, H. B., et al. 2014, *ApJ*, 788, 105
- Feast, M. W., & Whitelock, P. A. 2000, *MNRAS*, 317, 460
- Feng, Y., Guhathakurta, P., Peng, E. W., et al. 2024, *ApJ*, 966, 159
- Feillet, D. K., Feltzing, S., Sahlholdt, C., & Bensby, T. 2022, *ApJ*, 934, 21
- Foreman-Mackey, D. 2016, *JOSS*, 1, 24
- Fortune, M., Gibson, N. P., Foreman-Mackey, D., et al. 2024, *A&A*, 686, A89
- Freeman, K., & Bland-Hawthorn, J. 2002, *ARA&A*, 40, 487
- Freeman, K. C. 1987, *ARA&A*, 25, 603
- Fuhrmann, K. 2008, *MNRAS*, 384, 173
- Fukushima, T., Chiba, M., Homma, D., et al. 2018, *PASJ*, 70, 69
- Fukushima, T., Chiba, M., Tanaka, M., et al. 2019, *PASJ*, 71, 72
- Gilmore, G., & Reid, N. 1983, *MNRAS*, 202, 1025
- GRAVITY Collaboration, Abuter, R., Amorim, A., et al. 2021, *A&A*, 647, A59
- Grebel, E. K. 1999, *IAUS*, 192, 17
- Guldenshuh, K. A., Layden, A. C., Wan, Y., et al. 2005, *PASP*, 117, 721
- Han, J. J., Conroy, C., Johnson, B. D., et al. 2022, *AJ*, 164, 249
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Natur*, 585, 357
- Helmi, A. 2020, *ARA&A*, 58, 205
- Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, *Natur*, 563, 85
- Hernitschek, N., Cohen, J. G., Rix, H.-W., et al. 2018, *ApJ*, 859, 31
- Hogg, D. W., & Lang, D. 2013, *PASP*, 125, 719
- Horta, D., & Schiavon, R. P. 2025, *MNRAS*, 537, 3730
- Hunter, J. D. 2007, *CSE*, 9, 90
- Iorio, G., & Belokurov, V. 2019, *MNRAS*, 482, 3868
- Iorio, G., & Belokurov, V. 2021, *MNRAS*, 502, 5686
- Iorio, G., Belokurov, V., Erkal, D., et al. 2018, *MNRAS*, 474, 2142
- Jayasinghe, T., Stanek, K. Z., Kochanek, C. S., et al. 2019, *MNRAS*, 486, 1907
- Kinemuchi, K., Smith, H. A., Woźniak, P. R., McKay, T. A. & ROTSE Collaboration 2006, *AJ*, 132, 1202
- Kunder, A., Chaboyer, B., & Layden, A. 2010, *AJ*, 139, 415
- Kunder, A., Pérez-Villegas, A., Rich, R. M., et al. 2020, *AJ*, 159, 270
- Lane, J. M. M., Bovy, J., & Mackereth, J. T. 2023, *MNRAS*, 526, 1209
- Layden, A. C. 1995, *AJ*, 110, 2312
- Leung, H. W., Bovy, J., Mackereth, J. T., et al. 2023, *MNRAS*, 519, 948
- Li, X.-Y., Huang, Y., Liu, G.-C., Beers, T. C., & Zhang, H.-W. 2023, *ApJ*, 944, 88
- Lucey, M., Hawkins, K., Ness, M., et al. 2021, *MNRAS*, 501, 5981
- Lucey, M., Sanderson, R. E., Horta, D., et al. 2025, *ApJ*, 982, 87
- Madore, B. F., & Freedman, W. L. 2012, *ApJ*, 744, 132
- Marsakov, V. A., Gozha, M. L., & Koval, V. V. 2018, *ARep*, 62, 50
- Marsakov, V. A., Gozha, M. L., & Koval, V. V. 2019, *Ap*, 62, 467
- Martin, J. C., & Morrison, H. L. 1998, *AJ*, 116, 1724
- Mateu, C. 2024, *RNAAS*, 8, 85
- Mateu, C. 2025, *rrl\_completeness*, v2.1.1, Zenodo, doi:10.5281/zenodo.10835147
- Mateu, C., Holl, B., De Ridder, J., & Rimoldini, L. 2020, *MNRAS*, 496, 3291
- Mateu, C., & Vivas, A. K. 2018, *MNRAS*, 479, 211
- Mateu, C., Vivas, A. K., Downes, J. J., et al. 2012, *MNRAS*, 427, 3374
- Medina, G. E., Muñoz, R. R., Carlin, J. L., et al. 2024, *MNRAS*, 531, 4762
- Medina, G. E., Muñoz, R. R., Vivas, A. K., et al. 2018, *ApJ*, 855, 43
- Miller, T. B., & van Dokkum, P. 2021, *ApJ*, 923, 124
- Molla, M., Ferrini, F., & Diaz, A. I. 1996, *ApJ*, 466, 668
- Mullen, J. P., Marengo, M., Martínez-Vázquez, C. E., et al. 2023, *ApJ*, 945, 83
- Neeley, J. R., Marengo, M., Bono, G., et al. 2017, *ApJ*, 841, 84
- Neeley, J. R., Marengo, M., Freedman, W. L., et al. 2019, *MNRAS*, 490, 4254
- Nibauer, J., & Bonaca, A. 2025, *ApJL*, 985, L22
- Pérez, F., & Granger, B. E. 2007, *CSE*, 9, 21
- Pérez-Villegas, A., Portail, M., & Gerhard, O. 2017, *MNRAS*, 464, L80
- Pietrukowicz, P., Kozłowski, S., Skowron, J., et al. 2015, *ApJ*, 811, 113
- Posti, L., & Helmi, A. 2019, *A&A*, 621, A56
- Prudil, Z., Dékány, I., Grebel, E. K., & Kunder, A. 2020, *MNRAS*, 492, 3408
- Reddy, B. E., Lambert, D. L., & Allende Prieto, C. 2006, *MNRAS*, 367, 1329
- Rix, H.-W., Chandra, V., Andrae, R., et al. 2022, *ApJ*, 941, 45
- Robin, A. C., Reylé, C., Fliri, J., et al. 2014, *A&A*, 569, A13
- Rybicki, J., & Drimmel, R. 2018, *gdr2\_completeness*: GaiaDR2 data retrieval and manipulation, Astrophysics Source Code Library, ascl:1811.018
- Sandage, A., & Tammann, G. A. 2006, *ARA&A*, 44, 93
- Sanderson, R. E., Hickox, R., Hirata, C. M., et al. 2024, arXiv:2404.14342
- Searle, L., & Zinn, R. 1978, *ApJ*, 225, 357
- Sesar, B., Fouesneau, M., Price-Whelan, A. M., et al. 2017a, *ApJ*, 838, 107
- Sesar, B., Ivezić, Ž., Lupton, R. H., et al. 2007, *AJ*, 134, 2236
- Sesar, B., Ivezić, Ž., Stuart, J. S., et al. 2013, *AJ*, 146, 21
- Sesar, B., Hernitschek, N., Mitrović, S., et al. 2017b, *AJ*, 153, 204
- Shazeer, N., & Stern, M. 2018, arXiv:1804.04235
- Smith, H. A. 1995, *RR Lyrae Stars*, Vol. 27 (Cambridge Univ. Press)
- Stan Development Team 2024, *Stan Modeling Language Users Guide and Reference Manual*, v2.36, <https://mc-stan.org/docs/reference-manual/>
- Stringer, K. M., Drlica-Wagner, A., Macri, L., et al. 2021, *ApJ*, 911, 109
- Strugnell, P., Reid, N., & Murray, C. A. 1986, *MNRAS*, 220, 413
- Sturch, C. 1966, *ApJ*, 143, 774
- Thomas, G. F., McConnachie, A. W., Ibata, R. A., et al. 2018, *MNRAS*, 481, 5223
- Tkachenko, R., Vieira, K., Lutsenko, A., Korchagin, V., & Carraro, G. 2025, *Univ*, 11, 132
- Vasiliev, E., Belokurov, V., & Erkal, D. 2021, *MNRAS*, 501, 2279
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *NatMe*, 17, 261
- Vivas, A. K., Saha, A., Olsen, K., et al. 2017, *AJ*, 154, 85
- Walker, I. R., Mihos, J. C., & Hernquist, L. 1996, *ApJ*, 460, 121
- Watkins, L. L., Evans, N. W., Belokurov, V., et al. 2009, *MNRAS*, 398, 1757
- Wyse, R. F. G., & Gilmore, G. 1995, *AJ*, 110, 2771
- Xu, Y., Liu, C., Xue, X.-X., et al. 2018, *MNRAS*, 473, 1244
- Xue, X.-X., Rix, H.-W., Ma, Z., et al. 2015, *ApJ*, 809, 144
- Yang, C., Zhu, L., Tahmasebzadeh, B., Xue, X.-X., & Liu, C. 2022, *AJ*, 164, 241
- Zhang, H., Iorio, G., Belokurov, V., et al. 2025, *MNRAS*, 544, 2493
- Zinn, R., Chen, X., Layden, A. C., & Casetti-Dinescu, D. I. 2020, *MNRAS*, 492, 2161
- Zinn, R., Horowitz, B., Vivas, A. K., et al. 2014, *ApJ*, 781, 22