The structure in front of the Galactic bar traced by red clump stars in the VVV survey

Dylan Paterson⋆, Brendan Coleman, Chris Gordon
School of Physical and Chemical Sciences, University of Canterbury, Christchurch, New Zealand

ABSTRACT
Studies of the red clump giant population in the inner Milky Way suggest the Galactic bulge/bar has a boxy/peanut/X-shaped structure as predicted by its formation via a disc buckling instability. We used a non-parametric method of estimating the Galactic bulge morphology that is based on maximum entropy regularisation. This enabled us to extract the three-dimensional distribution of the red giant stars in the bulge from deep photometric catalogues of the VISTA Variables in the Via Lactea (VVV) survey. Our high-resolution reconstruction confirms the well-known boxy/peanut/X-shaped structure of the bulge. We also find spiral arm structures that extend to around three kpc in front of and behind the bulge and are on different sides of the bulge major axis. Additionally, at distances over 500 pc above or below the Galactic plane we find that the X-arms are on the same side of the bulge major axis as their corresponding spiral arm structure. Closer to the Galactic plane, the X-arms are on opposite sides of the major axis of the bulge. This implies a twisting in the X-shape of the bulge.

Key words: Galaxy: bulge – Galaxy: structure – Galaxy: formation

1 INTRODUCTION
Extragalactic studies of disc galaxies have found bulges/bars with boxy/peanut/X-shaped (B/P/X) components are relatively common in nearby early type (S0 - Sd) disc galaxies (Laurikainen et al. 2014; Ciambur & Graham 2016). Theories of secular evolution suggest that these bulges grow through instability processes in the disc, which cause the inner parts to expand vertically (Sellwood 2014). These buckling processes have been observed in N-body simulations of disc galaxies, where the resulting bar exhibits a strong B/P/X geometry (Bureau & Athanassoula 2005; Debattista et al. 2006). When viewed at an inclined angle, the bars of N-body simulated galaxies in the post buckling phase appear to have offset spurs at the end, which have also been observed in other galaxies (Erwin & Debattista 2013, 2016). These spurs are associated with the presence of multiple bar components, similar to the long bar and B/P/X bar in the Milky Way. However, a B/P/X geometry does not automatically mean that the Galactic bar has buckled as the B/P/X bulge may have formed via an orbital resonance (Quillen et al. 2014).

Studying the bar of the Milky Way is challenging due to our viewing angle of the Galactic bulge. One promising method to uncover the shape of the inner Milky Way is to use the red clump (RC) stars, which have a narrow intrinsic luminosity range, as standard candles (Girardi 2016).

The RC has been the focus of several studies characterising the three-dimensional density structure of the Galactic bulge. The most common class of parametric model used to describe the bulge is the triaxial ellipsoid (Stanek et al. 1997; Rattenbury et al. 2007; Cao et al. 2013; Simion et al. 2017). Although these triaxial models do a reasonable job of describing the general structure of the bulge, some of the studies, e.g Simion et al. (2017) (hereafter S17), show evidence of non-triaxial structures in the residual star-counts. Non-parametric methods have also been used involving deconvolution or constant intrinsic RC magnitude assumptions as in Wegg & Gerhard (2013) and Saito et al. (2011) respectively. The Galactic RC was found to produce a double photometric peak by Nataf et al. (2010) using OGLE-III data and McWilliam & Zoccali (2010) using 2MASS. The deconvolution results of Wegg & Gerhard (2013) (from here on WG13) showed a B/P/X bulge using the RC stars in the VVV survey.

We developed and applied our fully non-parametric deprojection of the Galactic bulge stars, relying only on choice of luminosity function and smoothness regularisation. In applying the principle of maximum entropy for statistical inference (Jaynes 1957), we aimed to produce a smooth density estimate of the Galactic bulge region and explore potential features of interest. In particular, we are interested in features such as the X-shape of the bulge and spiral arm struc-

⋆ E-mail: dnp16@uclive.ac.nz

© The Authors
turers on the ends of the bar, which are difficult to model with existing methods.

Our article is arranged as follows: In Sections 2 and 3 we outline our data selection and we rationalise the choice of data cuts and masks. Our semi-analytic luminosity function is constructed and its foundations presented in Section 3. In Section 4 we present our non-parametric deconvolution method for inverting stellar statistics to recover the three-dimensional stellar density distribution. We perform maximum entropy deconvolution on the VVV data and interpret the results in Section 5. In a concurrent article (Coleman et al. (2019), hereafter Paper II), we test the robustness of the maximum entropy method presented in this article on both a simulated population and the VVV catalogue.

2 VVV DATA

We used data from the MW-BULGE-PSFPHOT compilation (Surot et al. 2019), an ultra deep, infra-red, photometric catalogue of almost 600 million stars in the Milky Way bulge. Included in the catalogue are $K_s$ and $J$ apparent magnitudes from PSF fitting VVV images (Minniti et al. 2010), completeness for most stars from artificial star tests, extinction corrected $K_s$ and $J$ magnitudes, combined photometric + systematic uncertainties for $K_s$ and $J$, and a variety of quality metrics.

From this catalogue we constructed binned star counts on a $(80 \times 100 \times 75)$ linear grid in extinction corrected magnitude ($K_s$), Galactic latitude ($l$), and Galactic longitude ($b$). The range of the grid was $11 < K_s < 15$, $-10^6 < l < 10^6$, and $-10^6 < b < 5^6$. To select mainly the Red Giant stars, we excluded sources with $0.4 < J - K_s < 1.0$. A few sources in the catalogue do not have completeness values, as the detectors on which they were observed were excluded from the completeness analysis, so we were unable to completeness correct our star counts on a star-by-star basis. Instead, we calculated the mean completeness in each $(K_s, l, b)$ voxel. We corrected for completeness by dividing the number count of stars in a voxel by the estimated completeness of that voxel.

The photometry in the MW-BULGE-PSFPHOT compilation was calibrated relative to the Cambridge Astronomical Survey Unit (CASU) aperture photometry catalogues (Saito et al. 2012), which are known to have field-to-field variations in $K_s$ zero-point of up to 0.1 mag. We corrected for this variation in zero-point by adding to the $K_s$ magnitudes, within each tile, the median difference between the 2MASS point source catalogue (Skrutskie et al. 2006) $K_s$ magnitude and the non extinction corrected MW-BULGE-PSFPHOT $K_s$ magnitude. We limited the cross matching to sources in 2MASS with $12 < K_s < 13$ to ensure good photometric quality in both source catalogues and used a cross matching threshold of 0.1". This limit was used to reduce to effect of crowding and source merging in the 2MASS catalogue (Hajdu et al. 2019). The photometric offsets are shown on the left panel of Fig. 1.

Even after extinction correction and completeness correction, some regions on the sky had residual effects in their star counts. We chose to exclude these regions from our analysis by masking where the crowding and extinction is high, using the combined systematic + photometric $K_s$ magnitude error as a proxy. In the right panel of Fig. 1, the mean $K_s$ magnitude error ($\langle \sigma_{K_s} \rangle$) of stars with $12.975 < K_s < 13.025$ is shown. The value of the exclusion boundary, $\langle \sigma_{K_s} \rangle = 0.06$, was chosen to visually match the $E(J - K) = 0.99$ boundary in the less crowded $|l| > 5^\circ$ region. We can see that in the left panel of Fig. 1, our $\langle \sigma_{K_s} \rangle$ based mask mask excluded from the analysis nearly all the tiles with a significant positive photometric correction. Pixels that contain globular clusters in the GLOBCLUDST (Harris 2010) globular cluster catalogue were also excluded from the analysis.

3 ISOCHRONES, BULGE METALLICITY AND LUMINOSITY FUNCTIONS

Previous studies (e.g. WG13 and S17) have produced luminosity functions by fitting a parametric model to simulated populations of stars, with masses randomly drawn from initial mass functions (IMFs) and absolute magnitudes assigned by interpolation of mass-absolute magnitude relations from isochrones. In this framework, the absolute magnitude and mass are treated as random variables, where the luminosity function is the probability density function of absolute magnitude, and the IMF is the probability density function of mass. Instead of simulating the luminosity function, we adopted a more analytic approach. The luminosity function for a specific age $\tau$ and metallicity $z$ is determined by

$$
\phi(M_{K_s}, z, \tau) = \sum_i \xi(\theta^{-1}(M_{K_s}, z, \tau)) \left| d\theta^{-1}(M_{K_s}, z, \tau) \right|
$$

where $\xi$ is the IMF and $\theta$ is the mass-absolute magnitude relation

$$
M_{K_s} = \theta(m, z, \tau).
$$

In mass ranges where $\theta$ is not uniquely invertible, the luminosity function is summed over all possible solutions to the inversion of $\theta$. To get the luminosity function for the full population, we took the expected value of Eq. (1)

$$
\Phi(M_{K_s}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(M_{K_s}, z, \tau) f(z, \tau) dz dr.
$$

where $f$ is the metallicity distribution function. We assumed a bulge age of 10 Gyr with metallicity normally distributed with solar mean metallicity $\mu_{[Fe/H]} = 0.0$ and metallicity dispersion $\sigma_{[Fe/H]} = 0.4$ (Zoccali et al. 2008).

We constructed our bulge luminosity function using mass-absolute magnitude relations from the PARSEC+COLIBRI 10 Gyr isochrone sets Marigo et al. (2017) using 39 metallicity bins linearly spaced in the range $-2.279 < [Fe/H] < 0.198$. These isochrones are tabulated at fixed mass and metallicity values. The magnitude values between the fixed points were interpolated using a linear univariate spline in mass along a single metallicity isochrone. Attempting to interpolate between evolutionary stages where there are large changes in luminosity, e.g. first ascent red giant to helium core burning giant, introduced artefacts in the resulting luminosity function, so we used the evolutionary stage flags in the isochrones to separate them; 0-3 red giant branch, 4-6 RC and $> 6$ asymptotic giant
The structure in front of the Galactic bar traced by red clump stars in the VVV survey

branch. Fig. 2 shows the luminosity function calculated using Eq. (3) with mass-absolute magnitude relations from PARSEC+COLIBRI isochrones and a Chabrier (2003) log-normal IMF. Fitting a Gaussian to the RC component gave a mean absolute magnitude \( \mu_{MB} \approx -1.53 \) with standard deviation \( \sigma_{RC} \approx 0.05 \) which is consistent with the luminosity function of S17.

Observational effects such as residual extinction and crowding introduce uncertainty in measuring the \( K_s \) apparent magnitude, which effectively broadens the observed luminosity function. We accounted for this by convolving our semi-analytic luminosity, described above, with a Gaussian with dispersion \( \sigma_{K_s} \). As \( \sigma_{K_s} \) was different for each voxel, effectively every voxel had a slightly different luminosity function.

4 DECONVOLUTION METHOD

The stellar density (\( \rho \)) of the Galactic bulge can be reconstructed by inverting the equation of stellar statistics (e.g. López-Corredoira et al. (2000), WG13)

\[
N(K_s, l, b) = N_{thin}(K_s, l, b) + N_{thick}(K_s, l, b) + \Delta \Omega \Delta K_s \int_{-\infty}^{\infty} \rho(s, l, b) \Phi(K_s - 5 \log s - 10) s^2 ds
\]

where \( N \) denotes total number of stars in a voxel centred at \((K_s, l, b)\). This is made up of contributions from the thin disc \((N_{thin})\), thick disc \((N_{thick})\), and a contribution given by a weighted integral of the Galactic bulge number density \( \rho \). The \( \Delta \Omega \Delta K_s \) denotes the solid angle subtended by the line-of-sight, \( \Delta K_s \) denotes the width of a \( K_s \) magnitude bin, and \( s \) denotes the distance from the Sun measured in kpc. An example of the the broadened luminosity function (\( \Phi \)) is shown in Fig. 2. We chose the integration range 4 kpc \( \leq s \leq 13 \) kpc as the Galactic bulge density is negligible outside that region.

Following S17, we modelled the thick and thin discs using the description for the Besançon galaxy model (Robin et al. 2003). The thin disc was constructed from seven sub-populations which have different ages spanning 0-10 Gyr, where the star formation rate was assumed constant for each sub-population. All sub-populations were assumed to have relaxed into isothermal distributions, where the density distribution is a cylindrically symmetric holed ellipsoid,

\[
\langle \sigma_{K_s} \rangle = 0.05
\]

\[
\langle \sigma_{K_s} \rangle = 0.05
\]

\[
\langle \sigma_{K_s} \rangle = 0.05
\]

\[
\langle \sigma_{K_s} \rangle = 0.05
\]
The disc parameters we used are listed in Tables 1 and 2. Both the thick and thin discs were modelled as having a warp and a flare, where (5) and (8) at $Z$, is instead evaluated at $Z + Z_{\text{warp}}$ when $R > R_{\text{warp}}$: $\phi_{\text{warp}}$ is the direction in which the warp is maximum. The flare was modelled by linearly increasing the scale height by

$$h_{\text{flare}} = \gamma_{\text{flare}}(R - R_{\text{flare}})$$

(11)

when $R > R_{\text{flare}}$. We used the same parameters for the flare and warp as Robin et al. (2003); $\gamma_{\text{warp}} = 0.18$, $R_{\text{warp}} = 0.98 R_0$; $\phi_{\text{warp}} = 90.0^\circ$, $\gamma_{\text{flare}} = 0.0054$ and $R_{\text{flare}} = 1.12 R_0$. The disc parameters we used are listed in Tables 1 and 2.

### Table 1. Density distribution parameters for the Besançon thick and thin discs used for our simulation

| Component    | Age (Gyr) | $h_{\text{flare}}/Z$ | $\epsilon/h_{\text{flare}}$ | $h_{\text{flare}}$ |
|--------------|-----------|-----------------------|-------------------------------|---------------------|
| Thin Disc    | 0.0-0.15  | 0.0140                | 3.00                          |
|              | 0.15-1    | 0.0268                | 1.32                          |
|              | 1-2       | 0.0375                | 1.32                          |
|              | 2-3       | 0.0551                | 1.32                          |
|              | 3-5       | 0.0696                | 1.32                          |
|              | 5-7       | 0.0785                | 1.32                          |
|              | 7-10      | 0.0791                | 1.32                          |

### Table 2. Metallicity distribution parameters for the Besançon thick and thin discs used for our simulation

| Component    | Age (Gyr) | $\mu_{[\text{Fe/H}]}$ | $\sigma_{[\text{Fe/H}]}$ |
|--------------|-----------|-----------------------|---------------------------|
| Thin Disc    | 0.0-0.15  | -0.01                 | 0.12                      |
|              | 0.15-1    | -0.03                 | 0.12                      |
|              | 1-2       | -0.03                 | 0.10                      |
|              | 2-3       | -0.01                 | 0.11                      |
|              | 3-5       | -0.07                 | 0.18                      |
|              | 5-7       | -0.14                 | 0.17                      |
|              | 7-10      | -0.37                 | 0.20                      |

### 4.1 Maximum Entropy Deconvolution

Maximum entropy methods (MEMs) have been used in applications such as image reconstruction in radio interferometry (Cornwell & Evans 1985). They have also been used in foreground/background modelling of diffuse emission processes, e.g. cosmic microwave background studies with WMAP (Bennett et al. 2003) and diffuse gamma-ray studies with Fermi-LAT (Storm et al. 2017).

We used penalised likelihoods with penalties which come in two general forms: the first is maximum entropy regularisation which is defined for a field $\kappa$,

$$-2 \ln L_{\text{MEM}} = 2\lambda \sum_{h,j,k} (1 - \kappa_{h,j,k} + \kappa_{h,j,k} \ln \kappa_{h,j,k})$$

(12)

where $i$, $j$, and $k$ are the grid points for $s$, $l$, and $b$ respectively. As shown in the Appendix, $L_{\text{MEM}}$ has an extremum at $\kappa_{i,j,k} = 1$. We used a parameterisation where $\kappa$ is the ratio between a modelled quantity of interest and a smooth prior estimate of the quantity. Where there is little information in the data about the modelled quantity, such as noisy or low count regions, the model will tend towards the prior. As shown in the Appendix, the prior uncertainty or dispersion of $\kappa$ is $\lambda^{-1/2}$. So for example, if we expected deviations of around 10% from the smooth prior estimation of the quantity, we would set $\lambda = 100$. The larger the value of $\lambda$ chosen, the smaller the prior uncertainty assumed and so the more regularisation of the solution is applied.

The second form of likelihood penalty is $\ell_2$-norm regularisation of the second derivative of the logarithm. If $\rho$ varied over one dimension, we would use the usual second order...
The structure in front of the Galactic bar traced by red clump stars in the VVV survey 5
central difference equation approximation of curvature:
\[-2 \ln L_{\text{smooth}} = \eta \sum_i (\ln \rho_{i-1} + \ln \rho_{i+1} - 2 \ln \rho_i)^2. \tag{13}\]
This penalty enforces smoothness as it has a minimum when \(\ln \rho\) has zero curvature which occurs for \(\ln \rho\) which is constant or varies linearly with \(i\). In which case \(\rho\) will either be constant or vary as the exponential of a linear function. Therefore, where there is no data to influence the fit, such as masked regions, this regularisation will tend to give exponential behaviour. As shown in Appendix, the prior relative standard deviation from an exponential of a linear function is approximately \(1/\sqrt{6}\). So, the larger the value chosen for \(\eta\) the more smoothness regularisation is applied. A similar smoothness regularizing term was used by Bissantz & Gerhard (2002) to estimate the morphology of the bulge from the COBE DIRBE data.

Our maximum entropy method constructs a model for predicting the binned star counts, using a non-parametric description of the density. It maximises the penalised log likelihood:
\[\ln L = \sum_{i,j,k} (n_{i,j,k} \ln N_{i,j,k} - N_{i,j,k}) - \sum_{i,j,k} \left( \lambda (1 - \kappa_{i,j,k} + \kappa_{i,j,k} \ln \kappa_{i,j,k}) + \eta \left( \ln \rho_{i-1,j,k} + \ln \rho_{i+1,j,k} - 2 \ln \rho_{i,j,k} \right)^2 / 2 \right. \tag{14}\] and the predicted counts \(N\). The second line has the maximum entropy regularisation term of the form given in Eq. (12), where \(\kappa\) is the ratio between the fitted stellar density, \(\rho\), and a prior estimate of the density, \(\rho_{\text{prior}}\):
\[\kappa \equiv \frac{\rho}{\rho_{\text{prior}}}. \tag{15}\]

The last three lines of Eq. (14) are the smoothness regularisation for the density field, of the form given in Eq. (13), in the \(s\), \(l\), and \(b\) directions. Including the maximum entropy term in the likelihood discourages the modelled density from over-fitting to regions of the data that are dominated by noise, where it will instead favour the smooth prior density. Addition of the smoothness terms discourages spurious high frequency variations in the modelled density by minimising curvature in the logarithm of the density. The smoothness term also has the added benefit of inpainting the density in lines of sight which have been masked out. We set \(\lambda = 0\) in masked regions so as they are only affected by the smoothness term and the values of the model at the edge of the mask.

For a smooth prior density of the bulge, we used a parametric S-model (Freudenreich 1998; Simion et al. 2017)
\[\rho_{\text{prior}} = \rho_0 \sech^2 (r_s) \tag{16}\]
where,
\[r_s = \left( \left( \frac{||X||}{x_0} \right)^{c_1} + \left( \frac{||Y||}{y_0} \right)^{c_2} + \left( \frac{||Z||}{z_0} \right)^{c_3} \right)^{1/4} \tag{17}\]
and \(X, Y,\) and \(Z\) are distances measured along a coordinate system that is centered in the bulge and aligned with the bulge axes.

In Paper II, we perform extensive tests of our maximum entropy method on a simulated Milky Way population. From those tests, we found that a suitable choice of regularisation parameters to reconstruct the stellar bulge density is \(\lambda = 0.01, \eta = 400.0, \eta_s = 200.0,\) and \(\eta_b = 100.0\). We found that the results were insensitive to small changes in these values.

5 DECONVOLUTION RESULTS

We applied the maximum entropy deconvolution process to the VVV data by maximizing the \(\ln L\) in Eq. (14). We used the Python implementation PYLIFEC3 of the Limited Memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm. The density was modelled non-parametrically on a \((257, 100, 75)\) grid of \((s, l, b)\), in the range \(4 < (s/kpc) < 13, -10^5 < b < 5^5\) and \(-10^5 < l < 10^5\), for a total of \(1.9275 \times 10^6\) free parameters (the grid spacing is \(\Delta s, \Delta l, \Delta b = 35\) pc, 0.2°, 0.2°). The normalisation of the thin and thick discs were also left as free parameters. To make the optimization of so many parameters feasible, we evaluated the gradients of \(\ln L\) in Eq. (14) analytically (see Appendix). We used a parametric S model (Eq. (16)) that was fitted to the VVV data in Paper II as the prior density \(\rho_{\text{prior}}\).

Fig. 3 shows examples of our model fit for two different lines of sight. At high latitudes, as in the top panel of Fig. 3, the VVV data is very noisy due to low number counts. The maximum entropy method is able to predict the splitting of the RC, even though it is not immediately apparent in the data. Due to the smoothing regularisation, the fitted model for the displayed line of sight is influenced by data in all of the neighbouring voxels. As a result, the model can appear poorly constrained by the data in a single line of sight, as in the bottom panel of Fig. 3. When viewed as a slice of the data, as in Fig. 4, the model is well constrained across multiple neighbouring line of sights. Due to the narrowness of the RC in the luminosity function, the morphology of the bulge is mainly constrained by the stars in the magnitude range \(12 < K_s < 14\). Therefore, the slight bias of the fit at \(K_s > 14\) is not of particular concern as it is not directly influencing our inferences about the morphology of the bulge region.

We plot a Cartesian projection of the reconstructed bulge density in Fig. 5, where the origin is centred on the maximum density of the bulge. From this we infer that the Sun is at \((x, y, z) = (-8.0, 0.0, 0.0)\) kpc. The X-arms are visible at \(|z| > 0.319\) kpc. Although WG13 had a similar result for \(z < 0\), they had significant gaps in their reconstruction for \(z > 0.263\) kpc. However, they filled in these gaps by assuming eight-fold symmetry. As we did not have this problem, we did not need to make any symmetry assumptions.

A result of this is that our final reconstruction of the X-arm nearest to the Sun is less dense than our reconstruction of the X-arm which is farthest from the Sun. In contrast, WG13’s final reconstruction had identical X-arms due to their symmetry assumptions. Our less restrictive symmetry

https://github.com/dedupeio/pylbfgs
assumptions have also allowed us to uncover the presence of a spiral arm structure in front of the bar, which is visible in the deconvolved density (left panels of Fig. 5) at $|z|<500$ pc at $x \sim -3$ kpc. We also found a spiral arm structure behind the bar which is visible at all $|z|<1$ kpc at $x \sim 3$ kpc.

Figure 3. Demonstration of the fitted model for a line of sight which displays a splitting in the RC (Top panel) and a line of sight which is near the edge of the masked midplane region (Bottom panel).

At $z = \pm 0.319$ kpc where, the X-arms have started to merge, the residuals in the right column of Fig. 5 show that the maxima of the X-arm at positive $x$ has migrated to $(x, y) = (0, 0)$ kpc. The maxima of the X-arm at negative $x$ has migrated to $(x, y) = (-0.8, 0.0)$ kpc, so that the midpoint of the two X-arms is at $(x, y) = (-0.4, 0.0)$ kpc. At $z = -1.001$ kpc, the midpoint of the X-arms is $(x, y) = (0.0, 0.0)$. This is consistent with the observed $R_0$ shift of 400 pc in WG13, which they interpreted as an intrinsic brightening of the RC due to a metallicity gradient. Additionally, a 400 pc shift forward is also consistent with the 5 Gyr old (0.1 mag brighter RC) E-model component of the S+E model in S17. The bottom of the X-arms merge into a structure that is consistent with the orientation of the E component of S17 S+E model.

Shown in the top left panel of Fig. 6 are the positions of the maximum density in the X-arms, where the X-arms maxima appear to move closer to the Sun as $z$ approaches zero. We assume that the shifting towards the Sun is an intrinsic brightening of the RC, so that at each $z$ the X-arm maxima positions should be shifted so that the midpoint between them is $(x, y) = (0.0, 0.0)$. After correcting for the apparent intrinsic brightening, the X-arms exhibit a clear twisting structure. The twist is especially obvious in the bottom-right panel of Fig. 6, which shows there is no axis in the $x - y$ plane in which the X-arms would exhibit axial symmetry. The X-arms have a $180^\circ$ rotational symmetry. Interestingly, the maxima of the X-arms appear to be nearly aligned with the major axis of the long bar of Wegg et al. (2015) (purple shading in bottom-right panel of Fig. 6). This lends support to the assumption made by Ciambur et al. (2017) that the long bar and X-bulge are aligned. Overall,
The structure in front of the Galactic bar traced by red clump stars in the VVV survey

The structure in front of the Galactic bar traced by red clump stars in the VVV survey.

...Non Parametric
Non Parametric - Parametric S Model
z = 0.495
Non Parametric - Parametric S Model
z = 0.319
Non Parametric - Parametric S Model
z = 0.0
Non Parametric - Parametric S Model
z = 0.319
Non Parametric - Parametric S Model
z = 0.495
Non Parametric - Parametric S Model
z = 0.66
Non Parametric - Parametric S Model
z = 0.5
Non Parametric - Parametric S Model
z = 0.836
Non Parametric - Parametric S Model
z = 1.0

Figure 5. Cartesian projections of the bulge density from the maximum entropy deconvolution (left column) and the parametric prior density of model (middle column). The Sun is located at \((x, y, z) = (-8.0, 0.0, 0.0)\). The dashed black line indicates \(l = 0^\circ\) and the solid black line is the major axis of the bar in the parametric model which is at an angle of 19.8° from the \(l = 0^\circ\) line. The \(z\) coordinate is measured in kpc. At \(x \sim \pm 3\) kpc the spiral arm structures at both ends of the bulge are visible, most clearly in the residuals (right column), which has had the colourbar clipped at \(\pm 10\%\). The pink crosses indicate the maximum density of the X-arms, and the pink circle is the midpoint between the two arms.

The results in Fig. 6 demonstrate that the X-bulge cannot be described by an eight-fold symmetric structure.

The positions of the spiral arm structures in Fig. 7 are consistent with the inner galaxy of the simulated gas distribution of Renaud et al. (2013)\(^2\). The location of the spiral arm structure from Gonzalez et al. (2018) (white triangles in Fig. 7) are closer to the Sun than predicted by our model. This is likely because we have only considered fields which are not heavily affected by extinction and crowding. Additionally, the \(K_s\) magnitudes of the VVV stars in Gonzalez et al. (2018) have the photometric zero-point calibrated to the CASU aperture photometry catalogues, which is not consistent with the corrected zero-point magnitudes we have used.

In order to evaluate the apparent magnitude of the spiral arm structure in front of the bar, we note that as illustrated in Fig. 2, \(M_{K_s} \approx -1.53\) mag. Also, as can be seen from Fig. 7, the distance of this feature from the Sun is 5 kpc. We can then use the standard relation

\[
K_s - M_{K_s} = 5 \log_{10} s + 10
\]  

where \(s\) is in kpc. Substituting in the above values and solving gives \(K_s \approx 11.96\) mag. As Gonzalez et al. (2018) only used data with \(K_s \geq 12\) mag, they would not have been sensitive to this feature.

As can be seen in Fig. 7, the spiral arm structure at positive \(x\) connects to the bar below the major axis, while the spiral arm structure at negative \(x\) connects to the bar above the major axis. As can be seen by comparing, for example, the \(z = 0\) kpc left and \(z = -0.8361\) kpc right hand side panels of Fig 5, the spiral arm structures are offset from the bar major axis on the same side as the high \(|z|\) X-arm maxima.

As noted by Gonzalez et al. (2018), the red giant branch bump (RGBB) of the bulge has a similar \(K_s\) to the feature behind the bar. A mismodelling of the RGBB might explain some of the signal at \(z\) far from the Galactic Midplane seen in...
6 CONCLUSIONS

We have used a non-parametric method incorporating maximum entropy and smoothness regularisation to deconvolve the density distribution of bulge stars in the VVV MW-BULGE-PSFPHOT catalogue. We have observed two new morphological features in the inner region of the Milky Way: a feature \( \sim 3 \) kpc behind the Galactic centre. The spiral arm structures are connected on opposite sides of the major axis. We also observed that the X-arms were offset in the same way as the spiral arm structures. It would be interesting to check with N-body simulations whether or not this is a general feature of a galaxy in the post buckling phase.

ACKNOWLEDGEMENTS

We thank Iulia Simion and Victor Debattista for helpful discussions and comments. Also, we are grateful to Elena Valenti for giving us early access to the MW-BULGE-PSFPHOT VVV catalogue. This work was made possible by the use of the Research Compute Cluster (RCC) facilities at the University of Canterbury. The following software packages were used: ASTROPY, Matplotlib, NUMPY, PYLBFGS, and SCIPY.

REFERENCES

Bennett C. L., et al., 2003, ApJS, 148, 97
Bissantz N., Gerhard O., 2002, MNRAS, 330, 591
Bureau M., Athanassoula E., 2005, ApJ, 626, 159
Cao L., Mao S., Natali D., Rattenbury N. J., Gould A., 2013, MNRAS, 434, 595
Chabrier G., 2003, PASP, 115, 763
Ciambur B. C., Graham A. W., 2016, MNRAS, 459, 1276
Ciambur B. C., Graham A. W., Bland-Hawthorn J., 2017, MNRAS, 471, 3988
Coleman B., Paterson D., Gordon C., Macias O., Ploeg H., 2019, arXiv e-prints, p. arXiv:1911.04714
Cornwell T. J., Evans K. F., 1985, A&A, 143, 77
Debattista V. P., Mayer L., Carollo C. M., Moore B., Wadsley J., Quinn T., 2006, ApJ, 645, 209
Enaisto J., 1979, in Burton W. B., ed., IAU Symposium Vol. 84, The Large-Scale Characteristics of the Galaxy. pp 451–458
Erwin P., Debattista V. P., 2013, MNRAS, 431, 3060
Erwin P., Debattista V. P., 2016, ApJ, 825, L30
Freudenreich H. T., 1998, ApJ, 492, 495
Girardi L., 2016, ARA&A, 54, 95
Gonzalez O. A., et al., 2018, MNRAS, 481, L130
Hajdu G., Dekany I., Catelan M., Grebel E. K., 2019, arXiv e-prints, p. arXiv:1908.06160
Harris W. E., 2010, arXiv e-prints, p. arXiv:1012.3224
Jaynes E. T., 1957, Physical Review, 106, 620
Laurikainen E., Salo H., Athanassoula E., Bosma A., Herrero-Enoqui M., 2014, MNRAS, 444, L80
López-Corredoira M., Hammersley P. L., Garzón F., Simonneau E., Mahoney T. J., 2000, MNRAS, 313, 392
Marigo P., et al., 2017, ApJ, 835, 77
McWilliam A., Zoccali M., 2010, ApJ, 724, 1491
Minniti D., et al., 2010, New Astron., 15, 433
Natali D. M., Udalski A., Gould A., Fouqué P., Stanek K. Z., 2010, ApJ, 721, L28
Quillen A. C., Minchev I., Sharma S., Qin Y.-J., Di Matteo P., 2014, MNRAS, 437, 1284
Rattenbury N. J., Mao S., Sumi T., Smith M. C., 2007, MNRAS, 378, 1064
Renaud F., et al., 2013, MNRAS, 436, 1836
Robin A. C., Reylé C., Derrière S., Picaud S., 2003, A&A, 409, 523
Robin A. C., Marshall D. J., Schultheis M., Reylé C., 2012, A&A, 538, A106
The analytic gradient of ln $\rho$ is then

$$\frac{\partial \ln L_p}{\partial \rho_0} = \Delta \Omega \Delta K_s \left( \sum_{i,j,k} \left( \frac{\partial \rho}{\partial \rho_0} \right)_{h,i,j,k} \Phi_{h,i,j,k} \delta_{h,s}^2 \right).$$

Where $\Phi_{h,i,j,k}$ is the discretised version of the luminosity function which needs an index for $s, K_s, l, \text{ and } b$. Since $\rho_0$ is single value at $(h,j,k) = (h',j',k')$ in the field $\rho$, then $\frac{\partial \rho}{\partial \rho_0}$ is $\delta_{h,h'} \delta_{j,j'} \delta_{k,k'}$, where the $\delta$ are the Kronecker delta. Substituting this into Eq. (A4) and simplifying gives

$$\frac{\partial \ln L_p}{\partial \rho_0} = \Delta \Omega \Delta K_s \Delta s \delta_{j,j'} \delta_{k,k'} \Phi_{h',i,j,k} \delta_{h,s}^2.$$

Substituting Eq. (A5) into Eq. (A2) gives the final form of the Poisson component of the gradient as

$$\frac{\partial \ln L_p}{\partial \rho_0} = \Delta \Omega \Delta K_s \Delta s \sum_{i \in K_s} \left( \frac{m_{i,j,k'}}{N_{i,j,k'}} - 1 \right) \Phi_{h',i,j,k'} \delta_{h,s}^2.$$
The gradient for the Poisson component is equal to zero when \( n = N \), so that the log-likelihood has an extremum when the model, \( N \), and the data, \( n \), are equal.

The gradient of the maximum entropy penalty was determined by taking the derivative of Eq. (12) with respect to \( \rho = \delta \) and applying the chain rule

\[
\frac{\partial}{\partial \rho} \ln L_{\text{MEM}} = -\lambda \sum_{(h,j,k) \in \{s,l,b\}} (1 - \kappa_{i,j,k} + \kappa_{i,j,k} \ln \kappa_{i,j,k}) \frac{\partial \kappa}{\partial \rho} \delta
\]

(A7)

where \( \kappa = \frac{\rho}{\rho_{\text{prior}}} \) is the ratio between the density field and a smooth prior estimate, \( \rho_{\text{prior}} \). Evaluating the derivative in Eq. (A7) and substituting in \( \frac{\partial \kappa}{\partial \rho} = 1 \left( \frac{\rho_{\text{prior}}}{\rho} \right) ^{\delta} \) gives

\[
\frac{\partial}{\partial \rho} \ln L_{\text{MEM}} = -\lambda \left( \frac{\rho_{\text{prior}}}{\rho} \right) ^{\delta} \ln \left[ \frac{\rho_{\text{prior}}}{\rho} \right], \quad \text{(A8)}
\]

which equals zero (thus giving the extremum) when \( \rho = (\rho_{\text{prior}})^{\delta} \). It follows from this equation that

\[
\frac{\partial^2}{\partial \kappa^2} \ln L_{\text{MEM}} \bigg|_{\kappa = 1} = -\lambda. \quad \text{(A9)}
\]

We can then evaluate the expected deviation from the prior using the standard Gaussian approximation for estimating errors in maximum likelihood:

\[
\sigma_{\text{MEM}} \equiv \sigma = \left( -\frac{\partial^2}{\partial \rho^2} \ln L_{\text{MEM}} \bigg|_{\rho = (\rho_{\text{prior}})^{\delta}} \right)^{-1/2} = \frac{1}{\sqrt{\lambda}} \quad \text{(A10)}
\]

The gradient of the smoothing term was obtained by direct differentiation of the last three lines of Eq. (14) so that

\[
\frac{\partial}{\partial \rho} \ln L_{\text{smooth}} = \frac{-\eta}{\rho^2} \left[ \ln \rho_{h-2,j,k} - 4 \ln \rho_{h-1,j,k} + 6 \ln \rho_{h,j,k} \right. + \left. -4 \ln \rho_{h+1,j,k} + \ln \rho_{h+2,j,k} \right]
\]

\[
-\frac{\eta}{\rho^2} \left[ \ln \rho_{h,j-2,k} - 4 \ln \rho_{h,j-1,k} + 6 \ln \rho_{h,j,k} \right. + \left. -4 \ln \rho_{h,j+1,k} + \ln \rho_{h,j+2,k} \right]
\]

\[
-\frac{\eta}{\rho^2} \left[ \ln \rho_{h,j-2,k} - 4 \ln \rho_{h,j-1,k} - 1 + 6 \ln \rho_{h,j,k} \right. + \left. -4 \ln \rho_{h,j+1,k} + \ln \rho_{h,j+2,k} \right]
\]

(A11)

It is easy to check by that this equation is equal to zero when \( \rho_{h,j,k} \) is an exponential function of the form:

\[
\rho_{h,j,k} = A \exp(A_h h + A_j j + A_k k)
\]

(A12)

where \( A_h, A_j, \) and \( A_k \) are constants. Similarly to the MEM case, it follows from differentiating Eq. (A11) and evaluating the result using (A12) that the relative deviation from an exponential function is given by:

\[
\sigma_{\text{curvature}} \equiv \sigma = \frac{1}{\sqrt{6 \eta}} \quad \text{(A13)}
\]