Vol. 62

Boletín de la Asociación Argentina de Astronomía

2021

Boletín de artículos científicos 62ª Reunión Anual
Asociación Argentina de Astronomía
Rosario, Provincia de Santa Fe, octubre de 2020

62ª Reunión Anual de la AAA
Rosario, Provincia de Santa Fe
13 al 16 de octubre de 2020
Organizada por la Universidad Nacional de Rosario y el Complejo Astronómico Municipal "Galileo Galilei"

La organización agradece el auspicio y ayuda financiera del Consejo Nacional de Investigaciones Científicas y Técnicas.

Publicado por
Asociación Argentina de Astronomía
Paseo del Bosque s/n, La Plata, Buenos Aires, Argentina

ISSN 1669-9521 (versión digital)
ISSN 0571-3285 (versión impresa)

Créditos:
Diseño de portada: Adrián Rovero y Andrea León
Confeccionado con la clase "confproc" en LaTeX

Impreso en Buenos Aires, julio de 2021

www.astronomiaargentina.org.ar

Bayesian estimation of King's profile parameters for ten open clusters in the Milky Way

M.S. Pera[1], G.I. Perren[1], H.D. Navone[2] & R.A. Vázquez[1]

[1] *Instituto de Física de La Plata, CONICET–UNLP, Argentina*
[2] *Instituto de Física de Rosario, CONICET–UNR, Argentina*

Contact / msolpera@gmail.com

Resumen / Aplicamos un método de inferencia bayesiana diseñado para estimar los parámetros de un perfil de King sobre diez cúmulos abiertos seleccionados con datos obtenidos de la segunda publicación de datos de la misión *GAIA*. Los cúmulos están ubicados a distintas distancias, esparcidos por todo el disco galáctico y abarcando un amplio rango de edad. Mediante la aplicación de nuestro algoritmo de estimación de membresía pyUPMASK, los datos de entrada se limpian de estrellas de campo contaminantes antes de su procesamiento con el marco bayesiano. El método de ajuste de perfil King permite ajustar un perfil elíptico girado en datos espaciales, lo que da como resultado una solución de cuatro parámetros: ángulo de rotación, excentricidad, radio del núcleo y radio de marea. Planeamos extender los resultados obtenidos aquí a tantos cúmulos abiertos como sea posible, creando así una base de datos homogénea de parámetros estructurales estimados mediante inferencia bayesiana.

Abstract / We apply a Bayesian inference method designed to estimate the parameters of a King profile on ten selected open clusters with data obtained from the Gaia DR2 survey. The clusters are located at various distances, scattered throughout the galactic disk, and spanning a wide age range. Through the application of our membership estimation algorithm pyUPMASK, the input data is cleaned from contaminating field stars previous its processing with the Bayesian framework. The King profile fitting method method allows for an elliptical rotated profile to be fitted on spatial data, resulting in a four parameters solution: rotation angle, eccentricity, core radius, and tidal radius. We plan on extending the results obtained here to as many open clusters as possible, thus creating a homogeneous database of structural parameters estimated through Bayesian inference.

Keywords / methods: statistical — galaxies: star clusters: general — open clusters and associations: general — techniques: photometric — parallaxes — proper motions

1. Introduction

Open clusters are valuable laboratories not only to study stellar evolution, but also stellar dynamics. Over 2000 open clusters have been catalogued to date in the database compiled by Dias et al. (2002)(DAML02). Most of the catalogued clusters have their fundamental parameters estimated through rather basic analysis methods applied to data sets, very often, quite limited in terms of cluster area coverage. This is particularly true for structural properties such as their centers and radii, which are simply estimated by eye in most cases. A cluster's radius is related to its dynamical relaxation and its age, two very important characteristics. We thus carry out an analysis of ten selected open clusters using our two packages: pyUPMASK (Pera et al., 2021) and ASteCA (Perren et al., 2015). The first one is employed to estimate membership probabilities, while the second one takes care of the center and radius estimation.

2. Methods

The first part of this work consists in estimating membership probabilities using our pyUPMASK membership estimation algorithm (Pera et al., 2021), an improved version of the UPMASK algorithm by Krone-Martins & Moitinho (2014), to clean the input data from contaminating stars.

Briefly, the pyUPMASK algorithm consist of an outer loop enclosing an inner loop that iteratively applies a cluster identification method and a cluster rejection method. The algorithm allows the implementation of a large number of cluster identification methods, included in the Python library `scikit-learn`* (Pedregosa et al., 2011).

Once chosen, the clustering method processes the non-positional features (photometry, proper motions, etc) at the beginning of the inner loop. For each identified cluster, its coordinates distribution is compared with that of a two-dimensional uniform distribution in the same range. Those that are similar to a random uniform distribution of elements are rejected. Clusters that survive are saved for a later iteration of the inner loop. When no more clusters are rejected, the inner loop ends. All stars within surviving clusters are assigned a value of 1 and all stars in rejected clusters are assigned a value of 0. After this, a new iteration of the outer loop is initiated. The final probabilities assigned to each star are simply the averages of the (0, 1) values assigned by the inner loop, at each run of the outer loop.

*`https://scikit-learn.org/`

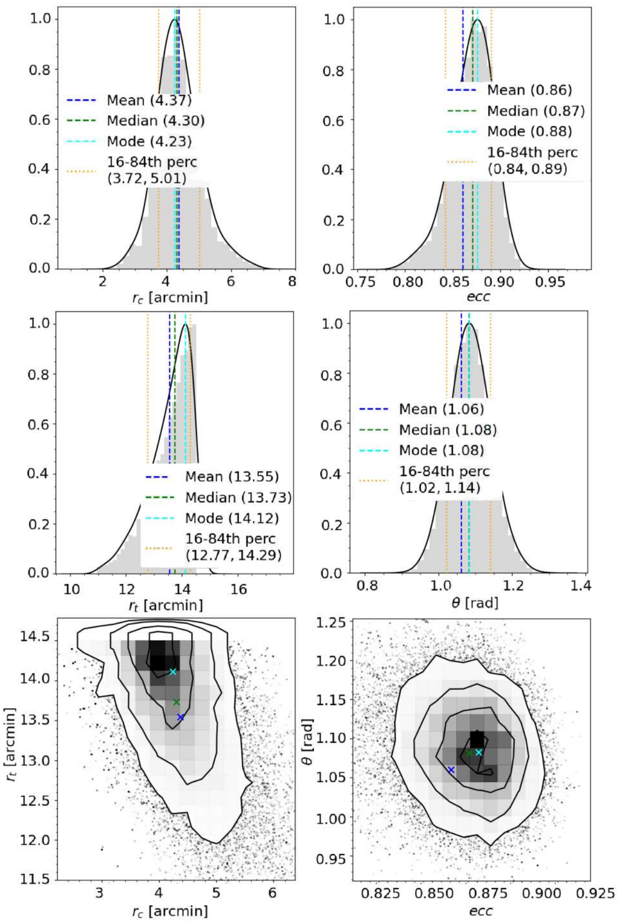

Figure 1: Distribution of the four parameters adjusted for NGC1893

For this work, we chose a clustering method from the `scikit-learn` library: Gaussian Mixture Model. We applied this method to the three-dimensional space formed by the proper motions and parallaxes (μ_α, μ_δ, Plx) and run 20 iterations of the outer loop.

The most probable members are then used to estimate King's profile parameters (King (1962)) on ten selected open clusters from the DAML02 database. The main parameters from this profile are the core (r_c, measures the degree of concentration at the core of the cluster) and tidal radius (r_t, measures the edge of the cluster beyond which stars are lost due to the gravitational pull of the host galaxy). For this we use the maximum likelihood estimation method employed in Pieres et al. (2016) and extended by us to process rotated and elliptical clusters. In the mentioned article, the likelihood that star i belongs to the full model (King profile) with radii r_c and rt and centred at (x_c, y_c) is written as

$$l_i = \rho_{cl}(r_i) + \rho_{fl}, \tag{1}$$

where ρ_{fl} is the field density value and $\rho_{cl}(r_i)$ is the surface density profile at a distance r_i from the cluster's center:

$$\rho_{cl}(r_i) = k \left(\frac{1}{[1+(r_i/r_c)^2]^{1/2}} + \frac{1}{[1+(r_t/r_c)^2]^{1/2}} \right), \tag{2}$$

where k is the central surface density and, for elliptical clusters, r_i is equivalent to the semi-major axis of a rotated ellipse with eccentricity ecc and rotation angle θ:

$$\left\{ \frac{[(x_i - x_c)cos(\theta) + (y_i - y_c)sin(\theta)]^2}{1 - ecc^2} + \frac{[(x_i - x_c)sin(\theta) - (y_i - y_c)cos(\theta)]^2}{1 - ecc^2} \right\}^{1/2}. \tag{3}$$

To estimate the parameters r_c, r_t, ecc and θ we use Bayesian inference on the model represented by the log-likelihood sum over all stars:

$$log(\mathscr{L})(r_c, r_t, ecc, \theta) = \sum_{i=1}^{N} log(l_i). \tag{4}$$

The ASteCA package first uses a two-dimensional Gaussian kernel density estimator to determine the center coordinates of the cluster. After this, the emcee package (Foreman-Mackey et al., 2013) is employed to explore the [r_c, r_t, ecc, θ] parameters space and estimate the distribution of each parameter.

We repeated this process two times: first using the subset of probable members estimated with pyUPMASK, and then using all the stars in the frame with no previous field star removal applied. We found the using the subset of most probable members impacts negatively on the results. This is a rather unexpected outcome, as we aimed at obtaining a smoother fit of King's profile using the sample of stars cleaned from field star contamination via pyUPMASK. The reason behind this appears to be the strong dependence of the fit on the field density parameter. The value of this parameter is estimated a priori and fixed during the King's profile fitting. When using the complete field, the value is obtained from the star density in field itself. When using the sampled cleaned by pyUPMASK this value is zero, as all the field stars are assumed to have been removed by the decontamination process. Although the pyUPMASK algorithm does a great job at assigning large probabilities to the most likely cluster members (as demonstrated in Pera et al., 2021), the selection of which stars with large membership values to keep as members is done by the user. A wrong selection can lead to an inaccurate representation of the cluster's true stellar density, which we believe is what's causing the poor performance in this case. This is a surprising result that we will investigate further in a future more in depth analysis.

3. Results

The results shown in this work are those that do not use the clusters cleaned by removing stars with low membership probabilities found by pyUPMASK. In general we found much better King profile fits using

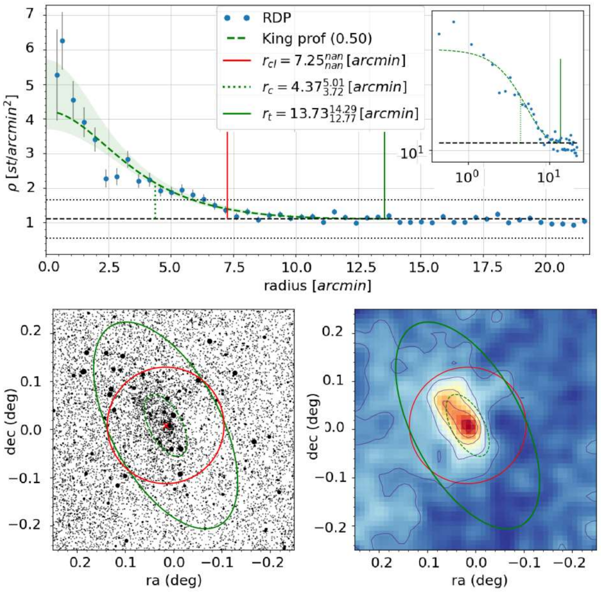

Figure 2: Elliptical radial density profile, r_c and r_t for NGC1893

simply the complete observed frame, as explained in Sect. 2.

We show the results for one of the ten clusters analyzed. Fig 1 and Fig 2 show the results given by ASTECA for NGC 1893. The six plots of the Fig 1 (three in the first column, three in the second column) show the resulting distribution of the four fitted parameters (r_c, r_t, ecc and θ) after the Bayesian inference method is applied. The plot in the first row of Fig 2 shows the elliptical radial density profile of the cluster region. The King profile fit is indicated with the dashed green curve with r_c and r_t shown as a vertical dashed and solid green line, respectively. The red vertical line is an estimated circular cluster found by another method in ASTECA. In the second row we show the cluster's coordinate space and a density map. The green ellipse corresponds to the ellipse of the adjusted parameters ecc and θ with a semi-major axis equal to r_t.

Table 1 shows the radii found in the DAML02 database for the 10 analyzed open clusters and the values r_c and r_t found by us through the Bayesian elliptical and rotated King profile fit.

It can be seen that in most cases the radius found in the DAML02 database are smaller than ours. For example, the estimated radius reported in the DAML02 database for Berkeley 31 is 2.5 [arcmin] which is four times smaller than the tidal radius of almost 10 [arcmin] estimated here.

This shows that the radii reported in DAML02, being in most cases a by-eye estimate, are underestimated by a substantial amount.

We also found that the rotation angle for ellipticities below ~ 0.5 has little value, as it is almost completely degenerated for clusters that do not have a clear ellipticity.

Table 1: Results obtained by us of the r_t and r_c values for the ten clusters, and the values found in the DAML02 database.

Name	α_{J2000}	δ_{J2000}	r_{cat} [arcmin]	r_c [arcmin]	r_t [arcmin]
Auner 1	07 04 16	−19 45 00	3.0	$0.78^{0.89}_{0.68}$	$6.86^{7.66}_{5.88}$
Berkeley 29	06 53 18	+16 55 00	3.0	$0.51^{0.57}_{0.45}$	$8.66^{9.37}_{7.64}$
Berkeley 31	06 57 36	+08 16 00	2.5	$0.91^{1.04}_{0.79}$	$8.66^{9.86}_{7.42}$
Berkeley 71	05 40 55	+32 16 40	3.1	$2.20^{3.02}_{1.44}$	$4.12^{4.89}_{3.52}$
IC 166	01 52 30	+61 50 00	3.5	$2.68^{2.85}_{2.50}$	$13.44^{14.05}_{12.70}$
IC 1311	20 10 18	+41 13 00	3.5	$1.97^{2.16}_{1.79}$	$9.77^{10.42}_{9.00}$
NGC 2671	08 46 12	-41 52 42	3.0	$13.56^{16.19}_{10.87}$	$17.71^{18.84}_{16.56}$
NGC 3603	11 15 07	-61 15 36	2.0	$0.67^{0.74}_{0.60}$	$4.21^{4.49}_{3.88}$
NGC 1893	05 22 44	+33 24 42	12.5	$4.37^{5.01}_{3.72}$	$13.73^{14.29}_{12.77}$
PISMIS 5	08 37 38	-39 35 00	6.0	$0.83^{1.17}_{0.52}$	$2.23^{2.70}_{1.82}$

4. Conclusions

Our method showed to work for rotated clusters even with a marked ellipticities, providing better fits that the simple circular (non-rotated) classic fit. The values found for r_t are, on average, two to five times larger than the values found in the DAML02 database. There are a few notable exceptions such as NGC 1893 where the DAML02 radius is comparable to the one found by us. Given the rather extreme ellipticity associated to this cluster (~ 0.86) it is easy to see that using a circular fit would overestimate the real area of the cluster by a large amount. This also means that a large number of non-members (field stars) would be included in the cluster region, further complicating its analysis.

Interestingly, we also found that the application of a decontamination method prior to the King profile fit generates unfavorable results for obtaining the desired parameters. This is true for both the classical non-elliptical non-rotated profile, as well as our more general elliptical and rotated profile.

We plan on extending this analysis to largest possible amount of open clusters from the DAML02 database, using Gaia DR2 data. We will also investigate in more depth analysis how the previous application of the decontamination method affects the King profile.

References

Dias W.S., et al., 2002, VizieR Online Data Catalog, VII/229A
Foreman-Mackey D., et al., 2013, PASP, 125, 306
King I., 1962, AJ, 67, 471
Krone-Martins A., Moitinho A., 2014, A&A, 561, A57
Pedregosa F., et al., 2011, J. Mach. Learn. Res., 12, 2825
Pera M.S., et al., 2021, arXiv e-prints, arXiv:2101.01660
Perren G.I., Vázquez R.A., Piatti A.E., 2015, A&A, 576, A6
Pieres A., et al., 2016, MNRAS, 461, 519