The Utilization of Bayesian Framework in Determining Projected Spatial Distributions of Star Clusters

Y A Hidayat, M I Arifyanto, and Aprilia

Department of Astronomy, Faculty of Mathematics and Natural Sciences, Institut Teknologi Bandung, Indonesia

email: yudarhd@gmail.com

Abstract. Star clusters are groups of stars that are bonded to each other by their gravitational forces. In general, star clusters are divided into the globular cluster and open cluster. There will always be some dynamic interactions in their lifetime, i.e. the process of relaxation and equipartition of interstellar encounter energy. The modeling is done by using King, Generalised King (GKing), Optimised Generalised King (OGKing), Elson Fall Freeman (EFF), Generalised Density Profile (GDP), and Restricted Generalised Density Profile (RGDP) profiles. Based on the Bayes factor, the results of a review of the density structure for star cluster Blanco 1 are dominated by King group profiles (King, GKing, and OGKing). Ellipticity values at the core radius ($\epsilon_{rc}$) and tidal radius ($\epsilon_{rt}$) obtained are $0.07^{+0.05}_{-0.01} + 0.10^{+0.06}_{-0.01}$, respectively. The estimation of the occurrence of mass segregation can be viewed from the value of $\kappa$. The $\kappa$ valued obtained for Blanco 1 is $0.06^{+0.10}_{-0.24}$ means that it undergo some mass segregation. The total mass obtained using the tidal force relationship of gravitational interactions with the potential Galaxy is $1623^{+373}_{-264} M_\odot$.

1. Introduction

The study of the dynamical evolution of the star cluster can be partially done by viewing the structure of a star cluster. The dynamical evolution is the interaction between the stars on a star cluster and it will change the structure of the cluster. This interaction, one of which can be caused by the process of the relaxation between the stars. This relaxation process can cause the deviation of the initial trajectory of each star in the cluster so that it is no longer in match with its initial conditions. In addition, it will provide the effect of energy exchange between stars with larger kinetic energy and those with smaller kinetic energy. This causes the stars with large mass to concentrate in the center of the cluster and those smaller stars in the edge. This process is known as mass segregation. Dynamical processes experienced by the cluster is also influenced by external factor such as tidal forces exerted by the host galaxy. By examining the structure of the star cluster within galactic environment, the cluster’s mass can be deduced. The total mass of the star cluster can be calculated from the tidal force. It means the relation between gravitational interaction on the star cluster and Galaxy potential. To find out the dynamic process, we will start from the spatial distribution form of the star cluster.

In this work, the spatial distribution will be reviewed in terms of the density function to the distance (radial) function from the center of the axis. The radial density profiles used are King, Generalised King (GKing), Optimised Generalised King (OGKing), Elson Fall Freeman (EFF), Generalised Density Profile (GDP), and Restricted Generalised Density Profile (RGDP) and the reasons as explained from [1]. In determining the parameters of each of these profiles, a Bayesian framework is used which is a posterior from model or likelihood with priors used. Furthermore, to determine the best radial density profile in the star cluster, a Bayes factor calculation was performed.
Bayes factor is the probability of the posterior ratio of evidence or marginal likelihood in a particular model with other models. Bayes factor interpretation used taken from [2]. The best model can be obtained based on this Bayes factor.

As the case study, we use Blanco 1 which is located in (RA, Dec) (J2000) = (0.85, -29.96). In the calculation, age (log year), distance (pc), and E(B-V) of this cluster are obtained as 7.75, 135.6, and 0.051, respectively from fitting isochrone [3] (see figure 1 on the left). The main reason for choosing Blanco 1 as the main object of this work is the shape that almost symmetrical circle for an open cluster depicted from the spatial distribution (see the right panel of figure 1). So, it is interesting to examine the structure using various density profile and calculate the value of its eccentricity. We can also predict whether there is a mass segregation phenomenon within the star cluster or not in the age of 56 million years. On that figure, we can see the Projected Spatial Distribution (PSD). The coordinates are far more easily measured than parallax (before GAIA), only a small fraction of the objects with the stellar positions have distance estimations. Furthermore, the relative uncertainties in the celestial coordinates yield far more precise measurements of distances perpendicular to the line of sight than those achieved by parallaxes along this line so far [1]. This explains why most of the previous works devoted to studying the spatial distribution of stars in clusters have been done using the PSD.

Figure 1. (left) Color-magnitude diagram of Blanco 1 with the best-fit isochrone. Red dots and blue line are photometric data from 381 member stars and the model from [3], respectively. (right) The PSD of the Blanco 1. The color scale defines the number density of stars per square arcsecond.

2. Data and Methods

2.1 Astrometric and Photometric Data

The data used are right ascension (α), declination (δ), photometry including the error value denoted by $d_i$ from Gaia DR 2 and the probability membership > 50% of star cluster denoted by $p_m$ from [4]. Based on these, the likelihood ($L$) function can be expressed as a probability in the data set $D = \{d_i, p_m\}$ given by the model ($M'$), parameter ($\theta$), and a limitation on the model ($I$)

$$L(D|\theta, M', I) = \prod_{i}^{N} [p_m \cdot p(d_i|\theta, M', I) + (1 - p_m) \cdot U(d_i|I)].$$  \hspace{1cm} (1)

The assumption used is that the stars are modeled by a uniform spatial distribution ($U$) within a maximum radius ($R_{max}$) of 11.5 pc. In addition, each star in the data set is not affected by the measurements of other stars or known as statistically independent.

2.2 Density Profiles

One of the earliest PSD function is King profile [5]
\[ n(r) = k \left[ \frac{1}{\left(1 + \left(\frac{r}{r_c}\right)^2\right)^{\frac{1}{2}}} - \frac{1}{\left(1 + \left(\frac{r}{r_t}\right)^2\right)^{\frac{1}{2}}} \right]^{2} \]  

(2)

with \( n(r), k, r_c, \) and \( r_t \) are the number density of star at the distance \( r \) from the center \((\text{star}/p\:c^2)\), the normalized constant of star density, core radius \((pc)\), and tidal radius \((pc)\), respectively.

[1] developed a more general form of equation (2) and known as GKing and OGKing. GKing equation as

\[ n(r) = k \left[ \frac{1}{\left(1 + \left(\frac{r}{r_c}\right)^{1/\alpha}\right)^{\frac{1}{\alpha}}} - \frac{1}{\left(1 + \left(\frac{r}{r_t}\right)^{1/\alpha}\right)^{\frac{1}{\alpha}}} \right]^{\beta} \]  

(3)

\( \alpha \) and \( \beta \) as a parameter that will be fit with the data. OGKing is GKing profile with the values of \( \alpha \) and \( \beta \) fixed at the maximum-a-posterior (MAP) values of the GKing parameters. This maximizes the Bayesian evidence and reduces the dimensionality of the parameter space.

Apart from the King profile, there are other profiles that can describe the PSD of the star cluster, namely the EFF, GDP, and RGDP profiles. The EFF profile is the profile of Elson, Fall, Freeman [6] which was originally created for the distribution of young cluster. The PSD function of the EFF profile is

\[ n(r) = n(0) \left(1 + \left(\frac{r}{r_c}\right)^2\right)^{-\gamma/2} \]  

(4)

with \( n(0) \) and \( \gamma \) are center of the core density and the tilt of the radius profile, respectively. GDP created by [7] as

\[ n(r) = \frac{k}{(r/c)^{\gamma} \left(1 + \left(r/c\right)^{2}\right)^{(\gamma-\beta)\alpha}} \]  

(5)

with \( \gamma, \alpha, \) and \( \beta \) are inner, outer, and width of the transition of regions constant, respectively. If \( \gamma \) parameter of equation [5] is equal to zero, that is RGDP.

The core and tidal radius eccentricity can be calculated with

\[ e_{rc} = 1 - \frac{r_{cb}}{r_{ca}}, e_{rt} = 1 - \frac{r_{tb}}{r_{ta}} \]  

(6)

\( r_{ca}, r_{cb}, r_{ta}, \) and \( r_{tb} \) are the core semi-major radius, core semi-minor radius, tidal semi-major radius, and tidal semi-minor radius, respectively.

### 2.3 Mass and Segregation

The total mass of the star cluster can be calculated from the tidal radius. It means that there is a relation between gravitational interaction on the star cluster and the Galactic potential. This can be stated using the Jacobi radius with Hill’s approximation, see [8]. The Jacobi radius can be expressed with

\[ r_j = \left(\frac{GM}{4\Omega_0 A_0}\right)^{1/3} \]  

(7)

\( G, M, \Omega_0, \) and \( A_0 \) are gravitational constant, total mass, circular frequency of cluster’s orbit around the center of the Galaxy, and the relation of Oort constant \( \Omega_0 = A_0 - B_0 \) with \( A_0=15.3 \pm 0.4 \: \text{km/kpc}, B_0=-11.9 \pm 0.4 \: \text{km/kpc}, \) respectively.
In comparison to those profiles, we also examine the biaxially symmetrical model. The biaxially symmetrical model needs new projection coordinates ($x, y$) that are rotated by the angle ($\Phi$). In the biaxially symmetrical model, there is a radius which is characterized by semi-major ($r_a$) and semi-minor ($r_b$) axis (see [1]). On the other hand, to evaluate the structure in the star cluster whether there is mass segregation or not, it can be seen using the luminosity segregated model. That is

$$r_c(\theta_a, G) = r(\theta_a) + \kappa(G - G_{mode})$$

(8)

with $r(\theta_a)$ is

$$r(\theta_a) = \frac{r_a r_b}{\sqrt{(r_a \sin(\theta_a))^2 + (r_b \cos(\theta_a))^2}}$$

(9)

the new radius after being projected by $\Phi$, $\kappa$ is a slope that does not depend on $\theta_a$, $G$ is magnitude, and $G_{mode}$ is a mode of magnitude. If $G > G_{mode}$ with positive $\kappa$, then the bright stars are in the core. This is a system of star cluster that illustrates many large-mass stars concentrated in the center of the cluster.

2.4 Bayesian Framework

In determining the parameters of each profile and posterior probability distributions model and evidence, we use the Nested Sampling algorithm from [9]. The Nested Sampling algorithm is implemented in the PyMultiNest module by [10]. To derive the parameters of the radial density profile, the Nested Sampling algorithm do as MCMC method, that is calculating a random probability of posterior as explained by [11]

$$P(\theta | D, M') \propto P(D | \theta, M') P(\theta | M')$$

(10)

where $P(D | \theta, M')$ is likelihood function and $P(\theta | M')$ is prior or range value of parameters. For each parameter, data and model are denoted as $\theta, D$ nd $M'$.

The prior used in this work is non-informative priors for parameters $\alpha, \beta$, and $\gamma$ used with the range of 0-15. The other parameters $\Phi$ used a uniform prior with the range of $[-\pi/2, \pi/2]$ and normal prior for $\kappa$ with the range of (0.0,0.5).

3. Results and Analysis

The parameters that have been determined using the Nested Sampling algorithm in this work are written in table 1 for radially symmetric, table 2 for biaxially symmetric, and table 3 for luminosity segregated models. Those three tables show the maximum-a-posterior obtained used in fitting the model to the data.

The distribution of each parameter and correlation obtained for the Blanco 1, is illustrated in figure 7 for GKing profile of the radially symmetrical. For the more information regarding the distribution of parameters obtained and summarized in the covariance matrix in table B.1-B.18, you can access http://github.com/yudarhd/Master-s-Thesis/blob/master/Blanco1.pdf

The correlation of each parameter obtained is marked in red. For the OGKing profile, the parameter values of $\alpha$ and $\beta$ use the values obtained from MAP on the GKing profile. In this case, the MAP GKing value with a radially symmetric, biaxial symmetric, and luminosity segregated models, i.e $0.14^{+0.59}_{-0.19}$ and $1.82^{+0.77}_{-0.57}$, $0.102^{+0.53}_{-0.13}$ and $1.99^{+0.63}_{-0.54}$, $0.12^{+0.58}_{-0.18}$ and $1.72^{+0.63}_{-0.53}$, respectively.

From the table 1-3, we see that some parameters have huge uncertainties. It could happen because the range of prior used is too large. We need to narrow down the prior range. The large range of prior is used due to the non-informative priors that only based on the distribution of the data.

| Table 1. Maximum-a-posterior estimation of the inferred parameters in each radially symmetric model |
|---------------------------------------------------------------|
| **Model** | $\alpha_0$ | $\delta_0$ | $r_c$ | $r_b$ | $\alpha$ | $\beta$ | $\gamma$ |
|-----------|-------------|-------------|--------|--------|----------|--------|--------|
| Radial    |             |             |        |        |          |        |        |
| GKing     |             |             |        |        |          |        |        |
| OGKing    |             |             |        |        |          |        |        |
Table 2. Maximum-a-posterior estimation of the inferred parameters in each biaxially symmetric model. Ellipticities are derived from a posterior using the inferred parameters

| Model  | $\alpha$ [°] | $\delta$ [°] | $\Phi$ [rad] | $r_{ca}$ [pc] | $r_{ta}$ [pc] | $r_{eb}$ [pc] |
|--------|--------------|--------------|-------------|--------------|--------------|--------------|
| EFF    | 0.95±0.06   | -29.93±0.07  | 3.44±0.90  | 3.36±1.03   | -            | -            |
| GDP    | 0.95±0.06   | -29.93±0.07  | 2.39±0.77  | 2.90±0.82   | 0.07±0.36   |              |
| GKing  | 0.95±0.06   | -29.93±0.07  | 1.94±0.50  | 14.36±23.94 | 0.14±0.58   | -            |
| King   | 0.94±0.06   | -29.93±0.07  | 1.96±0.41  | 15.08±16.78 | -            | -            |
| OGKing | 0.95±0.06   | -29.93±0.07  | 1.91±0.35  | 16.53±16.15 | -            | -            |
| RGDP   | 0.95±0.06   | -29.93±0.07  | 2.24±1.13  | 0.32±0.37   | 3.29±0.85   | -            |

Table 3. Maximum-a-posterior estimation of the inferred parameters in each luminosity segregated model. Ellipticities are derived from a posterior using the inferred parameters

| Model  | $\alpha$ [°] | $\delta$ [°] | $\Phi$ [rad] | $r_{ca}$ [pc] | $r_{ta}$ [pc] | $r_{eb}$ [pc] |
|--------|--------------|--------------|-------------|--------------|--------------|--------------|
| EFF    | 0.95±0.06   | -29.93±0.07  | 3.44±0.90  | 3.36±1.03   | -            | -            |
| GDP    | 0.95±0.06   | -29.93±0.07  | 2.39±0.77  | 2.90±0.82   | 0.07±0.36   |              |
| GKing  | 0.95±0.06   | -29.93±0.07  | 1.94±0.50  | 14.36±23.94 | 0.14±0.58   | -            |
| King   | 0.94±0.06   | -29.93±0.07  | 1.96±0.41  | 15.08±16.78 | -            | -            |
| OGKing | 0.95±0.06   | -29.93±0.07  | 1.91±0.35  | 16.53±16.15 | -            | -            |
| RGDP   | 0.95±0.06   | -29.93±0.07  | 2.24±1.13  | 0.32±0.37   | 3.29±0.85   | -            |
After fitting the distribution profile model to the data, then the distribution profile model is determined which can be considered as good by looking at the Bayes factor value obtained for each distribution profile model. A Bayes factor is the ratio of the likelihood of one particular model \( m_{oa} \) to the likelihood of another \( m_{ob} \). The higher value Bayes factor of a model, the better that model compared to another. The Bayes factor obtained from the posterior probability in each radial density distribution on radially symmetrical, biaxially symmetrical, and luminosity segregated models are written in tables 4, 5, and 6, respectively. The best results of the distribution profile fitting of the data by Bayes factor are illustrated in figure 2 for the radially symmetric, figure 3 for the biaxially symmetric, and figure 4 for the luminosity segregated models. We can see from the results obtained in these figures that from various kinds of probabilities of the model described by gray lines, the fittest one is marked by the red line. The red line is matched to the maximum-a-posterior.

**Figure 2. OGKing of the radially symmetric model.** Black dots, grey lines, and red lines are the data that have been binned with Poisson uncertainties, 100 samples from the posterior distribution, and inferred density of the radially symmetric profiles shown by means of the MAP, respectively.

**Figure 3. OGKing of the biaxially symmetric model.** Black dots, grey lines, and red lines are the data that have been binned with Poisson uncertainties, 100 samples from the posterior distribution, and inferred density of the biaxially symmetric profiles shown by means of the MAP, respectively.
Overall, the value of this evidence divided into two groups. First, the King profile family (including GKing and OGKing) and the second one is the non-King profile family (that are EFF, GDP, and RGDP). Bayes factor obtained is in the form of natural logarithms. In determining this model, tables 4-6 can be read in:

\[
\frac{p(D|M_{\text{raw}})}{p(D|M_{\text{column}})}
\]

Based on that, the evidence value obtained in the distribution profile with the radially symmetrical model has large evidence for the OGKing profile on GDP with an intermediate category. Furthermore, the smallest evidence value is obtained for the GDP profile in each distribution profile. These models can be considered as good for this star cluster, i.e OGKing, EFF, GKing, King, RGDP, and GDP. We can see that the profiles of the King profile family dominate over those of the non-King profile family. This means the King profile family can reflect the data on the profile fitting results in flexibility.

Figure 4. OGking of the luminosity segregated model. Black dots, grey lines, and red lines are the data that have been binned with Poisson uncertainties, 100 samples from the posterior distribution, and inferred density of the biaxially symmetric profiles shown by means of the MAP, respectively.

Figure 5. OGking of the luminosity segregated model. The data are binned in three bins of the $G$ band: $G < 12, 12 < G < 15$, and $G > 15$, with the colours of green.
cyan, and magenta, respectively. Three coloured solid lines are the MAP for the binned. Also the grey 100 samples from the posterior distribution.

Table 4. Bayes factors on the off-diagonal elements, with the evidence for the radially symmetric model

|     | EFF  | GDP  | GKing | King  | OGKing | RGDP  |
|-----|------|------|-------|-------|--------|-------|
| EFF | 0.00 | 44.51| 2.06  | 3.17  | 0.30   | 5.71  |
| GDP | 0.02 | 0.00 | 0.05  | 0.07  | <1e-2  | 0.13  |
| GKing | 0.48 | 21.58| 0.00  | 1.54  | 0.15   | 2.77  |
| King | 0.31 | 14.02| 0.65  | 0.00  | 0.10   | 1.80  |
| OGKing | 3.29 | 146.46| 6.79  | 10.45 | 0.00   | 18.80 |
| RGDP | 0.18 | 7.79 | 0.36  | 0.56  | 0.05   | 0.00  |

Table 5. Bayes factors on the off-diagonal elements, with the evidence for the biaxially symmetric model

|     | EFF  | GDP  | GKing | King  | OGKing | RGDP  |
|-----|------|------|-------|-------|--------|-------|
| EFF | 0.00 | 25.01| 4.06  | 4.99  | 0.41   | 2.32  |
| GDP | 0.04 | 0.00 | 0.16  | 0.20  | 0.02   | 0.09  |
| GKing | 0.25 | 6.17 | 0.00  | 1.23  | 0.10   | 0.57  |
| King | 0.20 | 5.02 | 0.81  | 0.00  | 0.08   | 0.46  |
| OGKing | 2.41 | 60.40| 9.80  | 12.04 | 0.00   | 5.60  |
| RGDP | 0.43 | 10.80| 1.75  | 2.15  | 0.18   | 0.00  |

Table 6. Bayes factors on the off-diagonal elements, with the evidence for the luminosity segregated model

|     | EFF  | GDP  | GKing | King  | OGKing | RGDP  |
|-----|------|------|-------|-------|--------|-------|
| EFF | 0.00 | 20.97| 2.82  | 1.83  | 0.28   | 3.24  |
| GDP | 0.05 | 0.00 | 0.13  | 0.09  | 0.01   | 0.15  |
| GKing | 0.36 | 7.45 | 0.00  | 0.65  | 0.10   | 1.15  |
| King | 0.55 | 11.48| 1.54  | 0.00  | 0.15   | 1.77  |
| OGKing | 3.63 | 76.11| 10.22 | 6.63  | 0.00   | 11.75 |
| RGDP | 0.31 | 6.48 | 0.87  | 0.56  | 0.09   | 0.00  |

Table 7. Mode of the distribution of total mass of the cluster with units in solar masses

|     | GKing | King | OGKing |
|-----|-------|------|--------|
| Ctr | 1054  | 1066 | 1259   |
| Ell | 1545  | 1389 | 1715   |
| Seg | 1454  | 1258 | 1896   |

For the distribution profile with the biaxially symmetrical model, the results of the Bayes factor following the pattern of the radially symmetrical. That is the highest evidence value obtained by the OGKing profile with the strong category and the smallest evidence value obtained by the GDP profile with no category for each profile in the distribution. The eccentricity values obtained for the King
profile family had an average of 0.08 ± 0.01 for the core radius eccentricity values and for the tidal radius eccentricity values the average value obtained is 0.10. The eccentricity value obtained means that the Galactic potential does not provide a strong interaction with this star cluster. Therefore, the two values obtained in the core radius and the tidal radius do not differ greatly.

The luminosity segregated obtained cannot be ignored. This is indicated by κ parameters of 0.06 and 0.12. The calculation obtained for the $G_{\text{mode}}$ is 14.35. The parameter of $\kappa$ means if the value is positive then the stars with large mass are concentrated in the center of the star cluster within the system (see Figure 5). In figure 5, from the MAP obtained it shows that the core radius has increased on these profiles. The evidence values obtained for each profile, following the same with the radially symmetrical and biaxially symmetrical models (see table 3). The criteria for OGKing is the intermediate category.

In determining the total mass of the star cluster, we use an approximation of the tidal radius of the gravitational interactions on the star cluster with the Galactic potential. The average distribution of the total mass is illustrated in figure 6 and the average distribution of the mode is written in table 7. Based on these result, the derived total mass of the Blanco 1 is $1623^{+273}_{-364} M_\odot$.

**Figure 6.** The total mass distribution for King’s family. These are derived from each radially symmetric (upper left), biaxially symmetric (upper right), and luminosity segregated (bottom).

**4. Conclusion**

The Bayesian framework method is highly effective in determining parameters and calculations to determine the better model in the radial density distribution profile. This is caused by the Bayesian framework that has a prior factor which makes the determination of the parameters faster. The largest evidence value from the Bayes factor is dominated by the King profile family (King, OGKing, and GKing). This is due to the profile of the King profile family specifically, the OGKing profile has $\alpha$
and $\beta$ parameters obtained from the maximum-a-posterior GKing profile. The smallest evidence value is dominated by the GDP and RGDP profiles in each distribution model. This is due to the high correlation of the $r_c$ and $\beta$ parameters which make the physical interpretation lose. The value of ellipticity obtained gives a different value on the core radius and tidal radius. This means that there are some interactions with the potential of the Galaxy. Based on the parameter of $\kappa$, we get the positive value that has meaning this open cluster undergo some mass segregation.

Figure 7. The distribution of each parameter and correlation obtained for GKing profile of the radially symmetrical.

Acknowledgment
This work is supported by *Program Penelitian Tesis Magister Kemenristekdikti* 2019 No. 2/E1/KP.PTNBH/2019.

References
[1] Olivares J, Moraux E, Sarro L M, Bouy H, Berihuete A, Barrado D, Huelamo N, Bertin E, and Bouvier J 2018 A&A 612 A70
[2] Jeffreys H 1961 *Theory of Probability, 3rd edition*. (Oxford: Oxford University Press)
[3] Maiz-Appellaniz J and Weiler M 2018 A&A 619 A180
[4] Cantat-Gaudin T et al. 2018 A&A 618 A93
[5] King I 1962 AJ 67 471-485
[6] Elson R A W, Fall S M, and Freeman K C 1987 AJ 323 54
[7] Lauer T R, Ajhar E A, Byun Y I, Dressler A, Faber S M, Grillmair C, Kormendy J, Richstone D, and Tremaine S 1995 AJ 110 2622-2666
[8] Binney J and Tremaine S 2008 *Galactic Dynamics: Second Edition* (Princeton: Princeton University Press, NJ USA)

[9] Skilling J 2006 *International Society for Bayesian Analysis* 1 833-860

[10] Buchner J, Georgakakis A, Nandra K, Hsu L, Rangel C, Brightman M, Merloni A, Salvato M, Donley J, and Kocevski D 2014 *A&A* 564 A125

[11] Hidayat Y A, Arifyanto M I, Aprilia, and Hakim M I 2019 *J. Phys.: Conf. Ser.* 1231 012029