



**HAL**  
open science

# BALRoGO: Bayesian Astrometric Likelihood Recovery of Galactic Objects – Global properties of over one hundred globular clusters with Gaia EDR3

Eduardo Vitral

► **To cite this version:**

Eduardo Vitral. BALRoGO: Bayesian Astrometric Likelihood Recovery of Galactic Objects – Global properties of over one hundred globular clusters with Gaia EDR3. *Monthly Notices of the Royal Astronomical Society*, 2021, 504 (1), pp.1355-1369. 10.1093/mnras/stab947 . hal-03373421

**HAL Id: hal-03373421**

<https://hal.sorbonne-universite.fr/hal-03373421v1>

Submitted on 11 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# BALRoGO: Bayesian Astrometric Likelihood Recovery of Galactic Objects – Global properties of over one hundred globular clusters with *Gaia* EDR3

Eduardo Vitral<sup>1</sup>\*

<sup>1</sup>*Sorbonne Université, CNRS, UMR 7095, Institut d’Astrophysique de Paris, 98 bis bd Arago, 75014 Paris, France*

Accepted XXX. Received YYY; in original form ZZZ

## ABSTRACT

We present BALRoGO: Bayesian Astrometric Likelihood Recovery of Galactic Objects, a public code to measure the centers, effective radii, and bulk proper motions of Milky Way globular clusters and Local Group dwarf spheroidals, whose data are mixed with Milky Way field stars. Our approach presents innovative methods such as surface density fits allowing for strong interloper contamination and proper motion fits using a Pearson VII distribution for interlopers, instead of classic Gaussian-mixture recipes. We also use non-parametric approaches to represent the color-magnitude diagram of such stellar systems based in their membership probabilities, previously derived from surface density and proper motion fits. The robustness of our method is verified by comparing its results with previous estimates from the literature as well as by testing it on mock data from  $N$ -body simulations. We applied BALRoGO to *Gaia* EDR3 data for over one hundred Milky Way globular clusters and nine Local Group dwarf spheroidals, and we provide positions, effective radii, and bulk proper motions. Finally, we make our algorithm available as an open source software.

**Key words:** methods: data analysis – astrometry – proper motions – (Galaxy:) globular clusters: general – galaxies: dwarf – stars: kinematics and dynamics

## 1 INTRODUCTION

The galactic neighborhood is a replete garden of important objects, such as globular clusters and dwarf spheroidal galaxies, that help to explain the key ingredients of galactic evolution and other unsolved mysteries in astrophysics. Globular clusters (GCs) are among the oldest relics of the Universe, sometimes reaching ages close to 13 Gyr (Marín-Franch et al. 2009), which are spherically shaped, compact collections of stars that orbit the Milky Way (MW), and whose origin is still uncertain (Peebles & Dicke 1968; Peebles 1984; Searle & Zinn 1978; Bullock & Johnston 2005; Abadi, Navarro & Steinmetz 2006; Peñarrubia, Walker & Gilmore 2009). They have recently been discovered to harbor multiple stellar populations (e.g. Carretta et al. 2009) and are frequently served as test laboratories for intermediate-mass black holes search (e.g. van der Marel & Anderson 2010; Noyola, Gebhardt & Bergmann 2008, and more recently Vitral & Mamon 2021). Dwarf spheroidal galaxies (dSphs), as well, are composed of old stellar systems and are located farther away than GCs, but still in the Local Group. They are much fainter group of stars and are often studied as tracers of dark matter (e.g., Battaglia, Helmi & Breddels 2013, Walker 2013 and Boldrini et al. 2020).

The study of both GCs and dSphs made a major leap as data from the *Gaia* astrometric mission became available, specially with its second and early third release (hereafter *Gaia* DR2 and *Gaia* EDR3, respectively). This mission provided an overall astrometric coverage of stellar velocities, positions and magnitudes of more than  $10^9$  stars

in the MW and beyond and therefore allowed many deeper studies of spherical systems such as GCs and dSphs, which can be well separated from MW field stars (i.e., interlopers) when combining both position, distance, and most importantly, proper motions. New bulk proper motions were derived in *Gaia* Collaboration et al. (2018), Baumgardt et al. (2019) and Vasiliev (2019a) with *Gaia* DR2 for more than 150 GCs, and different methods were used to do so, relying whether on Gaussian mixtures or iterative routines to differentiate tracers from interlopers. Similarly, McConnachie & Venn (2020b) derived bulk proper motions for dSphs using *Gaia* DR2, and later with *Gaia* EDR3 (McConnachie & Venn 2020a) by using a Bayesian routine also employing Gaussian mixtures, inspired by the work of Pace & Li (2019). Recently, by analyzing the GC NGC 6397, Vitral & Mamon (2021) showed that the separation between GC stars and interlopers was much more robust by assigning distribution functions to the proper motion space which were not based on a Gaussian mixture, as usual, but rather considering a GC Gaussian component plus a Pearson VII (Pearson 1916) distribution for the interlopers. An alternative to such problem is to consider not one single Gaussian for interlopers, but rather a multiple Gaussian component, such as in Vasiliev & Baumgardt (2021), who recently derived GC astrometric parameters with *Gaia* EDR3.

In fact, with the arrival of *Gaia* data, extracting GC members, as well as dSphs became a very important topic, for which a robust approach is required in order not to let interloper contamination bias one’s results. For example, if the selection of GC stars at its outskirts is not well made, one could misinterpret the presence of tidal tails, and thus derive wrong conclusions on the cluster evolution.

Given the importance of a separation algorithm between such

\* E-mail: vitral@iap.fr

galactic objects and interlopers, [Bustos Fierro & Calderón \(2019\)](#) proposed a method based on Gaussian mixtures and clustering algorithms, with only one step not performed by a machine and applied it to eight GCs. Their work joined previous attempts to separate cluster members from interlopers through parametric and non-parametric methods (e.g., [Vasilevskis et al. 1965](#) and [Galadi-Enriquez et al. 1998](#), respectively). In this paper, we also present an algorithm to extract GCs and dSphs members embedded in MW interlopers, which employs a combination of Bayesian fits, giving the membership probability of each star in the field, and non-parametric cuts on the color-magnitude diagram. Our method takes as an input a cone search delivered by GAIA and, from that data, estimates the object center in right ascension and declination (i.e.,  $\alpha$  and  $\delta$ ), the bulk proper motions,  $\mu_{\alpha,*}$  (i.e.,  $[d\alpha/dt] \cos \delta$ ) and  $\mu_{\delta}$ , and the effective radius (half projected number radius) by fitting discrete data. To validate our method, we compare bulk proper motions provided in the literature using GAIA EDR3, and we compare our effective radii estimates for over a hundred sources with other estimates. We use GAIA EDR3 data from the two GCs NGC 6752 and NGC 6205 (M 13) along with the Draco dSph galaxy as illustrative examples in our Figures, throughout the paper.

In addition, we derive effective radii, bulk proper motions and centers for over a hundred GCs from the NEW GALACTIC CATALOG (NGC) using GAIA EDR3 and make it available in Table format. Our algorithm developed with PYTHON is also provided as an open source code named BALRoGO: BAYESIAN ASTROMETRIC LIKELIHOOD RECOVER OF GALACTIC OBJECTS<sup>1</sup>, and allows the user to adjust the inputs of our algorithm to better suit any particular source.

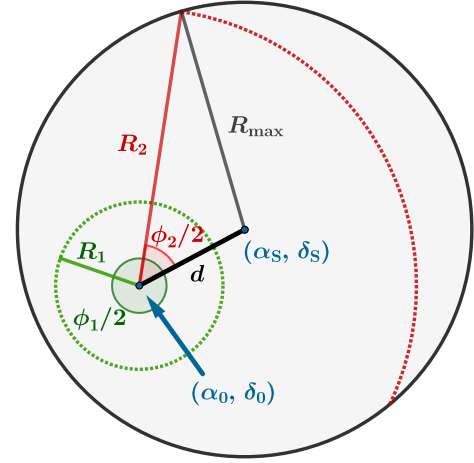
## 2 METHODOLOGY

Our algorithm is specially designed to deal with data from the GAIA astrometric mission, and therefore we begin by acquiring all the stars in a two degrees cone search for each of the GCs and dSphs in Table C1. The first two steps of our approach deal with all the stars in the two degrees field of view, and use only positional ( $\alpha$ ,  $\delta$ ) information. As it is going to be shown, they allow for interlopers in their fitting routines.

### 2.1 Center estimation

The first step of our analysis is the estimation of the GC or dSph (hereafter, galactic object for short) center. Although BALRoGO allows for an iterative frequentist approach, this can be dangerous when dealing with objects that suffer from crowding issues ([Arenou et al. 2018](#)), since the amount of stars decreases when one approaches the center too much. Therefore, as for many other steps, we opt for a Bayesian fit of the galactic object center.

This is done by fitting a Plummer profile ([Plummer 1911](#)) plus a constant contribution of interlopers to the surface density of the stars inside the two degrees cone (in some cases, we selected smaller radii due to high MW contamination, or due to parasite satellites). Such approach aims for nothing more than finding a center of mass of the stellar distribution, by fitting the center that best suits an spherical shaped collection of stars. Even if the galactic object does not follow precisely a Plummer model (e.g., core-collapsed GCs), other profiles, such as Sérsic ([Sérsic 1963](#); [Sersic 1968](#)), Hernquist ([Hernquist 1990](#)) or ([Kazantzidis et al. 2004](#)) for example also follow a circular



**Figure 1.** *Geometry of the center estimation:* The original cone search, centered in  $(\alpha_S, \delta_S)$  is shown as the gray circle of radius  $R_{\max}$ , while the new fitted center is  $(\alpha_0, \delta_0)$ . The angles  $\phi_1$  and  $\phi_2$  represent the effective circular sections (of respective radii  $R_1$  and  $R_2$ ) where we consider circular symmetry. The distance between the two centers is labeled as  $d$ , and we consider this value to approach zero for the cone searches around the centers from the SIMBAD data base.

symmetry, and therefore should have their centers well determined by a Plummer fit. In addition, allowing for the contribution of interlopers in the fit enables the user not to lose many galactic object stars and thus have a better knowledge of the radial extent of the source, as well as the radial membership probability of a star with respect to the galactic object.

Following the normalization of [Vital & Mamon \(2020, equations 15–17\)](#):

$$\Sigma(R) = \frac{N_{\infty}}{\pi R_{\text{scale}}^2} \tilde{\Sigma} \left( \frac{R}{R_{\text{scale}}} \right), \quad (1a)$$

$$N(R) = N_{\infty} \tilde{N} \left( \frac{R}{R_{\text{scale}}} \right), \quad (1b)$$

where  $N_{\infty}$  is the projected number of tracers at infinity, one can write the global surface density as:

$$\tilde{\Sigma}(X) = \tilde{\Sigma}_{\text{sys}}(X) + \mathcal{R} \frac{\tilde{N}_{\text{sys,tot}}}{X_{\text{max}}^2 - X_{\text{min}}^2}, \quad (2)$$

where  $\mathcal{R} = N_{\text{ilop,tot}}/N_{\text{sys,tot}}$  is the ratio between the number of interlopers and system (galactic object) stars,  $\tilde{N}_{\text{sys,tot}} = \tilde{N}_{\text{sys}}(X_{\text{max}}) - \tilde{N}_{\text{sys}}(X_{\text{min}})$  is the total number of system stars, and  $X = R/a$  is the normalized projected radius, with  $a$  being the Plummer scale radius, or the Plummer projected, two-dimensional, half number radius. The normalized surface density of the analyzed system, for a Plummer profile reads:

$$\tilde{\Sigma}_{\text{sys}}(X) = \frac{1}{(1 + X^2)^2}, \quad (3)$$

while the normalized projected number of stars for the same model is:

$$\tilde{N}_{\text{sys}}(X) = \frac{1}{\pi} \int_{X_{\text{min}}}^{X_{\text{max}}} \phi(X) X \tilde{\Sigma}_{\text{sys}}(X) dX, \quad (4)$$

<sup>1</sup> <https://gitlab.com/eduardo-vital/balrogo>

where  $\phi(X)$  is the angle corresponding to the effective circular section where we analyze the data (see Fig 1), and whose analytical expression for small cone apertures (i.e.  $R_{\max} \ll 1$  radian) is:

$$\phi(X = R/a) = \begin{cases} 2\pi & , \text{ if } R \leq R_{\max} - d, \\ 2 \arccos \left[ \frac{R^2 + d^2 - R_{\max}^2}{2 R d} \right] & , \text{ if } R > R_{\max} - d, \end{cases} \quad (5)$$

where  $R_{\max}$  is the maximum radius of the original cone search (usually two degrees) and  $d$  is the distance between the fitted center and the center from the original cone search ( $\alpha_S, \delta_S$ ), set as the source center on SIMBAD<sup>2</sup> by the automatic GAIA advanced query. The projected radius  $R$  is defined, in spherical trigonometry, as:

$$R = \arccos[\sin \delta \sin \delta_0 + \cos \delta \cos \delta_0 \cos(\alpha - \alpha_0)], \quad (6)$$

where  $(\alpha_0, \delta_0)$  is the center of the galactic object. The likelihood function is therefore written as:

$$\mathcal{L} = \prod_i \frac{\phi(X) X}{\pi a} \frac{\tilde{\Sigma}(X)}{\tilde{N}_{\text{sys,tot}}(1 + \mathcal{R})}. \quad (7)$$

We minimize  $-\log \mathcal{L}$  with the DIFFERENTIAL\_EVOLUTION routine from the SCIPY.OPTIMIZE method and find the centers that best fit the data, later displaying them in Table C1. For simplicity, since we expect the quantity  $d$  to approach zero, given the reliable previous measurements of the centers in SIMBAD, one can assume the case where  $\phi(X) = 2\pi$ , and thus derive:

$$\tilde{N}_{\text{sys}}(X) = \frac{X^2}{1 + X^2}. \quad (8)$$

Notably, for more general cases, the BALRoGO routine allows to use the more general representation of  $\phi(X)$ , and thus to account for a more complicated expression for  $\tilde{N}_{\text{sys}}(X)$ , presented in appendix A.

## 2.2 Surface density

The next step of our algorithm is to fit the surface density of the galactic object plus interlopers once again, using a maximum likelihood estimation (MLE), but this time with fixed centers from the previous step. It may seem as an unnecessary step given the previous one, but it actually allows the fitting routine to better explore the other parameters ranges like the scale radius and the interloper fraction, and thus derive more robust results. In this step, we also reduced the maximum projected radius whenever there was a parasite galactic object on the field or when the MW contamination was too intense, disturbing the fits.

In this step, we not only tested Plummer profiles, but also Sérsic models (Sérsic 1963; Sérsic 1968), which can be useful for core-collapsed clusters such as NGC 6397, and the Kazantzidis model, which is motivated from dynamical simulations of repeated tidal encounters (Kazantzidis et al. 2004). Equations 2, 6 and 7 remain valid for these two other models, while the normalized surface density and projected number of stars become:

$$\tilde{\Sigma}_{\text{sys}}(X) = \frac{b^{2n}(n)}{2n\Gamma(2n)} \exp\left[-b(n)X^{1/n}\right], \quad (9a)$$

<sup>2</sup> <http://simbad.u-strasbg.fr/simbad/>

$$\tilde{N}_{\text{sys}}(X) = \frac{\gamma(2n, b(n)X^{1/n})}{\Gamma(2n)}. \quad (9b)$$

for Sérsic, where  $\gamma(a, x) = \int_0^x t^{a-1} e^{-t} dt$  is the lower incomplete gamma function and  $X = R/R_e$ ,  $R_e$  being the effective projected radius. And for Kazantzidis:

$$\tilde{\Sigma}_{\text{sys}}(X) = K_0(X)/2, \quad (10a)$$

$$\tilde{N}_{\text{sys}}(X) = 1 - X K_1(X). \quad (10b)$$

Where  $K_0$  and  $K_1$  are modified Bessel functions of the second kind and  $X = R/a_K$ ,  $a_K$  being the radius of density slope  $-2$ .

We compared the likelihood of these three models by means of Bayesian inference, using the corrected Akaike Information Criterion (AICc, Sugiura 1978):

$$\text{AICc} = \text{AIC} + 2 \frac{N_{\text{free}}(1 + N_{\text{free}})}{N_{\text{data}} - N_{\text{free}} - 1}, \quad (11)$$

where AIC is the original Akaike Information Criterion (Akaike 1973):

$$\text{AIC} = -2 \ln \mathcal{L}_{\text{MLE}} + 2 N_{\text{free}}, \quad (12)$$

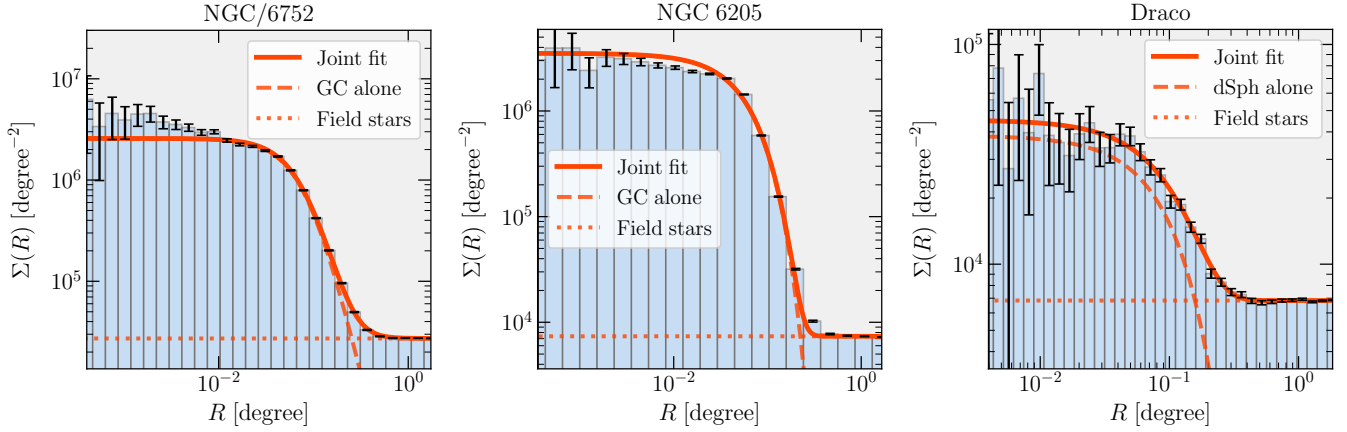
Our motivation for this criterion is the same one portrayed in Vitral & Mamon (2021), section 8.2. Whenever the difference between two models (i.e.,  $\Delta\text{AICc} = \text{AICc}_i - \text{AICc}_j$ ) had an absolute value smaller than two (i.e., no particular preference), priority was given to Plummer, Kazantzidis and Sérsic profiles, respectively. The results of our fits of surface density model plus constant interloper contribution are displayed in Figure 2 for NGC 6752, NGC 6205 (M 13) and the Draco dSph, respectively.

## 2.3 Data cleaning

The two previous steps required no data cleaning, since the filtering could bias the results, for example, by forcing radial incompleteness towards the center of the galactic object (mostly because of crowding issues) and cutting of fainter stars at the tracer outskirts, thus damaging the fitting of the center and other surface density parameters. However, the next step, which is aimed to look into the proper motion space is, in its turn, much more sensible to poor astrometry. This is a consequence of the much more subtle measurements of proper motions than sky coordinates, given that the GAIA astrometric mission has not yet a long timeline data base to fully trust its main velocity values without any filtering. Another reason why a conservative data cleaning of proper motions is necessary is that our proper motion fits themselves do not take into account the convolution of the interloper component with GAIA errors, which will be latter justified. Therefore, our data cleaning follows a similar approach as done in Vitral & Mamon (2021) for the GC NGC 6397, with some few changes, detailed in the following.

### 2.3.1 Maximum projected radius

We first filtered the GAIA stars inside a maximum projected radius: we selected stars inside a cone of radius  $r_{\text{cut}}$ , where  $r_{\text{cut}} = 10 R_{\text{scale}}$ .  $R_{\text{scale}}$  was  $R_e$ ,  $a$ , and  $a_K$  for Sérsic, Plummer and Kazantzidis, respectively. For some sources, this limit was changed in order to account for a more representative subset, but never out-passed the range between 0.03 and two degrees.



**Figure 2.** *Surface density fits:* Fits for the surface density of the galactic object (NGC 6752, NGC 6205 or M 13 and the Draco dSph, respectively) plus a constant contribution of interlopers, according to section 2.2. The *histogram* shows the empirical profile, using logarithmic radial bins extending from the innermost bin point to 2 degrees. The *curves* show different models: our MLE fit (*red*) of a Plummer, Sérsic and Sérsic models respectively (*dashed*) plus constant field stars surface density (*dotted*), as well as the total (*solid*) to compare with the data. The error bars were calculated considering only Poisson noise.

### 2.3.2 Low error stars

The next step was to retain only stars with  $\epsilon_\mu < \sigma_{\text{limit}}$ , where  $\sigma_{\text{limit}}$  is an error threshold and  $\epsilon_\mu$  is the proper motion error (semi-major axis of the error ellipse), defined in [Lindgren et al. \(2018, eq. B.2\)](#):

$$\epsilon_\mu = \sqrt{\frac{1}{2}(C_{33} + C_{44}) + \frac{1}{2}\sqrt{(C_{44} - C_{33})^2 + 4C_{34}^2}}, \quad (13)$$

$$C_{33} = \epsilon_{\mu_{\alpha,*}}^2, \quad (14)$$

$$C_{34} = \epsilon_{\mu_{\alpha,*}} \epsilon_{\mu_\delta} \rho, \quad (15)$$

$$C_{44} = \epsilon_{\mu_\delta}^2, \quad (16)$$

where  $\epsilon$  denotes the error or uncertainty and where  $\rho$  is the correlation coefficient between  $\mu_{\alpha,*}$  and  $\mu_\delta$ .<sup>3</sup> The error threshold was derived by default, as the following routine:

(i) First, we selected the stars inside the maximum radius mentioned above.

(ii) With this subset, we generated a two-dimensional histogram in proper motion space and assigned the two highest local peaks of the distribution with PYTHON's routine `PEAK_LOCAL_MAX` from the `SKIMAGE.FEATURE` method, which would correspond to the tracer and interloper clumps of stars. The binning of the histogram was chosen such that each dimension  $x_i$  had a number of bins computed as:

$$N_{\text{bin}} = \left\lceil \frac{4 \times (\max[x_i] - \min[x_i])}{\min[\eta_{.84}(x_i) - \eta_{.50}(x_i), \eta_{.50}(x_i) - \eta_{.16}(x_i)]} \right\rceil, \quad (17)$$

where  $\eta_n(x_i)$  is the  $n$ -th percentile of the  $x_i$  data. The typical values of  $N_{\text{bin}}$  were around 300.

(iii) We naively estimated the Gaussian dispersion of each clump by taking the distance between the peaks found above and the closest point where the histogram reached  $e^{-1/2}$  times the value at the peak.

(iv) We assigned the interloper peak as the one with great dispersion, while the narrow dispersion clump (calculated as above) was considered as the tracer velocity dispersion  $\sigma_{\text{tracer}}$ . The few cases where this relation was not satisfied (e.g., for 47 Tuc) could be corrected by selecting a narrow field of view around the source center, allowing for very few interlopers.

<sup>3</sup> We use the standard notation  $\mu_{\alpha,*} = \cos \delta [d\alpha/dt]$ ,  $\mu_\delta = d\delta/dt$ .

(v) The error threshold  $\sigma_{\text{limit}}$  was taken as  $\sigma_{\text{limit}} = k \sigma_{\text{tracer}}$ , where the default  $k$  was set as 0.5. For many clusters this value was changed in order to select a more representative subset. The GCs had all  $k$  values between 0.2 and two, while the dSphs had  $k$  up to five, given their greater distances, and consequently, much higher errors (otherwise, we would have too few tracers for dSphs in order to fit them properly).

### 2.3.3 Quality flags

We only kept stars whose astrometric solution presented a sufficiently low uncertainty:

$$\sqrt{\frac{\chi^2}{\nu' - 5}} < 1.2 \max \{1, \exp[-0.2(G - 19.5)]\}, \quad (18)$$

where  $\nu'$  is the number of points (epochs) in the astrometric fit of a given star and 5 is the number of free parameters of the astrometric fit (2 for the position, 1 for the parallax and two for the proper motions). Eq. (18) gives a sharper HR diagram, removing artifacts such as double stars, calibration problems, and astrometric effects from binaries. It is more optimized than the (`astrometric_excess_noise` < 1) criterion, used in [Baumgardt et al. \(2019\)](#) and [Vasiliev \(2019a\)](#), especially for brighter stars ( $G \lesssim 15$ ), according to [Lindgren et al. \(2018\)](#). We consider this step to substitute possible filters based on the GAIA `RUWE` parameter.

Second, we only kept stars with good photometry, by selecting the ones that satisfied:

$$C^*(r) < N \sigma_C, \quad (19)$$

where we select  $N = 3$ , such as in [Vasiliev & Baumgardt \(2021\)](#) and [McConnachie & Venn \(2020a\)](#), while  $\sigma_C$  follows equation (18) from [Riello et al. \(2020\)](#).  $C^*(r)$  is defined such as:

$$C^*(r) = C(r) - f(G_{\text{BP}} - G_{\text{RP}}), \quad (20)$$

with  $C(r)$  being the corrected excess factor and  $f(x) = \sum_i a_i x^i$ ,  $a_i$  reported in Table 2 from [Riello et al. \(2020\)](#). Eq. (19) performs an additional filter for unreliable astrometric solutions (mainly in the

cases of blended stars), affecting mainly faint sources in crowded areas. Those variables correspond to the following quantities in the GAIA DR2 and EDR3 archive:

- $\chi^2$ : `astrometric_chi2_al`
- $\nu'$ : `astrometric_n_good_obs_al`
- $C(r)$ : `phot_bp_rp_excess_factor`
- $G_{BP} - G_{RP}$ : `bp_rp`

## 2.4 Proper motions

After having cleaned the data set according to the previous section, we proceed to fit the proper motion joint distribution of galactic objects plus interlopers. These two components present two clumps of data in proper motion space which allow one to assign membership probabilities to each clump.

Vasiliev (2019a) and McConnachie & Venn (2020a) used Bayesian approaches based on Gaussian mixture models (a Gaussian distribution for both cluster and interloper clumps), while the original work from Gaia Collaboration et al. (2018) used an iterative method based on clustering approaches and Baumgardt et al. (2019) performed an iterative cleaning routine based on an  $n - \sigma$  outlier rejection. Even though Bayesian approaches tend to be more reliable, the use of Gaussian mixtures must, however, be taken with caution, since the distribution of field stars has often much wider wings than a Gaussian distribution (it can be cleverly bypassed by assigning multiple Gaussians to the interloper component, such as in Vasiliev & Baumgardt 2021). This trend on the interloper proper motion distribution was first noticed in Vitral & Mamon (2021, see appendix B), where the authors opted for a symmetric Pearson VII distribution for the interlopers and the results were much more robust. The trend is again confirmed in this work, where we employ a refined version of their Bayesian method, by assigning a symmetric Gaussian distribution for the galactic object plus a non-symmetric Pearson VII distribution for interlopers. The choice of a non-symmetric distribution allows for a much better adjustment of the interlopers, which improves the membership probability of the stars in the subset, while the impact on the bulk proper motions of the galactic object varies from cluster to cluster<sup>4</sup>.

The main drawback of such an approach is that in order to introduce a non-symmetric Pearson VII distribution, which is analytically more complicated than a symmetric distribution, we abandon the convolution of the interloper distribution with Gaussian errors. This was not an issue for Vitral & Mamon (2021), who used a polynomial approximation to this convolution for the symmetric case, which was simpler, in order to avoid extra integrals. This becomes much more computationally costly whenever accepting the non symmetricity of the data. We therefore try to counter this issue by trusting in our data cleaning described in the previous section, which applied conservative cuts on proper motion errors. Moreover, ignoring the error convolution for the non-Gaussian field population should not impact significantly the fits, since the interlopers show a much broader distribution, which is less affected than the narrow galactic object component (Gaussian). We further test this assumption in section 3.4.

<sup>4</sup> For example, there is no strong difference when using symmetric and non-symmetric interloper distributions for some GCs such as NGC 6397 and NGC 6121 (M4)

### 2.4.1 Probability distribution function

To construct the probability distribution function (PDF) of the analyzed subset, we consider the PDF from these tracers and from the MW contaminants:

$$\text{PDF} = f_{\text{GO}} \text{PDF}_{\text{GO}} + (1 - f_{\text{GO}}) \text{PDF}_{\text{MW}}, \quad (21)$$

where  $f_{\text{GO}}$  is the fraction of galactic object stars. The PDF of galactic objects is a straightforward Gaussian:

$$\text{PDF}_{\text{GO}}(\mu_{\alpha,*}, \mu_{\delta}) = \frac{\exp\left(-\zeta^2\right)}{2\pi\sigma_{\text{GO}}^2}, \quad (22a)$$

$$\zeta^2 = \frac{(\mu_{\alpha,*} - \mu_{\alpha,*\text{GO}})^2 + (\mu_{\delta} - \mu_{\delta,\text{GO}})^2}{2\sigma_{\text{GO}}^2} \quad (22b)$$

where  $\sigma_{\text{GO}}$  is the convolved proper motion dispersion of the galactic object stars, and  $\mu_{\alpha,*\text{GO}}$  and  $\mu_{\delta,\text{GO}}$  are their bulk proper motions in  $(\alpha, \delta)$ . The convolution of the Gaussian component is done by considering  $\sigma_{\text{GO}}^2 = \sigma_{\text{GO,int}}^2 + \epsilon^2$ , where  $\epsilon$  is the proper motion error (see equation [20] from Vitral & Mamon 2021) and  $\sigma_{\text{GO,int}}$  is the intrinsic dispersion of the source. For the PDF of interlopers, we first shift the origin to the bulk proper motions of the interlopers clump, and then rotate the reference frame into the main axis of the interlopers proper motion ellipsoidal distribution:

$$\mu_x = (\mu_{\alpha,*} - \mu_{\alpha,*\text{MW}}) \cos\theta + (\mu_{\delta} - \mu_{\delta,\text{MW}}) \sin\theta, \quad (23a)$$

$$\mu_y = -(\mu_{\alpha,*} - \mu_{\alpha,*\text{MW}}) \sin\theta + (\mu_{\delta} - \mu_{\delta,\text{MW}}) \cos\theta, \quad (23b)$$

where  $\mu_{\alpha,*\text{MW}}$  and  $\mu_{\delta,\text{MW}}$  are the contaminants bulk proper motions in  $(\alpha, \delta)$  and  $\theta$  is the angle between the original  $(\mu_{\alpha,*}, \mu_{\delta})$  frame and the new one. Then, if we call the semi-major and semi-minor axis of the Pearson VII ellipsoidal distribution  $a_x$  and  $a_y$ , we can write the interlopers PDF as:

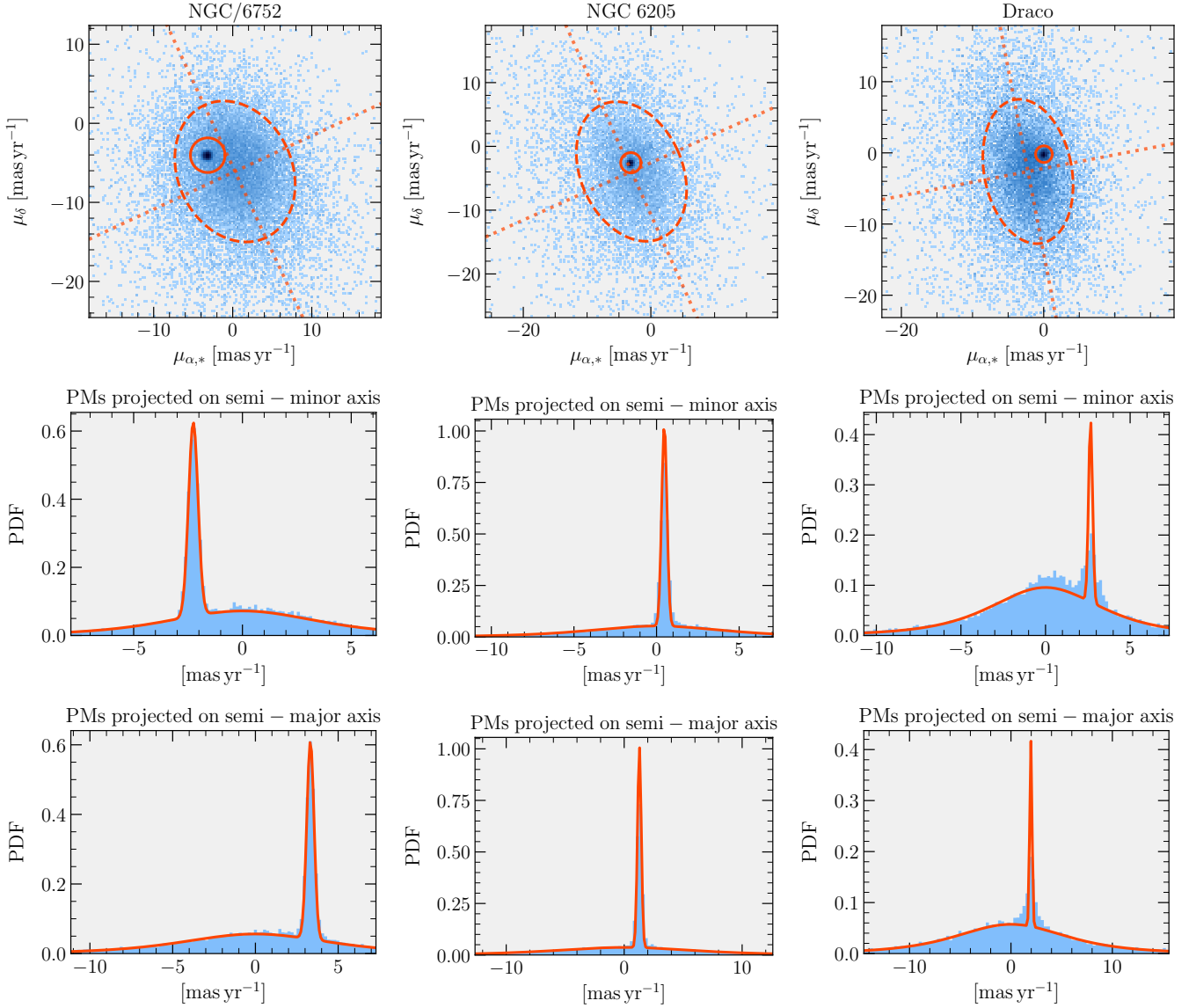
$$\text{PDF}_{\text{MW}} = \frac{\left[\frac{\Gamma\left(-\frac{1}{2} - \frac{\tau}{2}\right)}{\Gamma\left(-1 - \frac{\tau}{2}\right)}\right]^2 \left\{ \left[1 + \left(\frac{\mu_x}{a_x}\right)^2\right] \left[1 + \left(\frac{\mu_y}{a_y}\right)^2\right] \right\}^{(1+\tau)/2}}{\pi a_x a_y}, \quad (24)$$

where  $\Gamma(x)$  is the gamma function of  $x$  and  $\tau$  is an intrinsic slope of the distribution (with  $\tau < -2$ ). Thus, we had ten free parameters, which were fitted by an MLE routine using the DIFFERENTIAL\_EVOLUTION method in PYTHON.

In Figure 3, we present the outcome of those fits for NGC 6752, NGC 6205 (M 13) and the Draco dSph, respectively. The first row shows the fit over the entire proper motion space, where one can verify the asymmetry of the interlopers distribution, while the second and third rows display the fits projected on the semi-minor and semi-major axis respectively.

### 2.4.2 Handling errors

In order to derive statistical errors of our Bayesian estimates, such as the bulk proper motion uncertainties shown in Table C1, we used PYTHON'S NUMDIFFTOOLS.HESSIAN method to compute the Hessian



**Figure 3.** *Proper motion fits:* We display here the results of the proper motion fits of NGC 6752, NGC 6205 (M 13) and the Draco dSph in the first, second and third columns, respectively. The first row displays the entire proper motion subset color-coded by stellar counts, from light blue to dark blue. The dashed ellipse displays the Pearson VII asymmetric distribution, with its semi-minor and semi-major directions as two perpendicular dotted lines, while the continuous circle represents the galactic object (globular cluster or dwarf spheroidal) proper motion mean with a radius equals ten times its intrinsic dispersion, for better visualization. The second and third rows display the fits (solid red) projected on the semi-minor and semi-major axis respectively, with the data in blue.

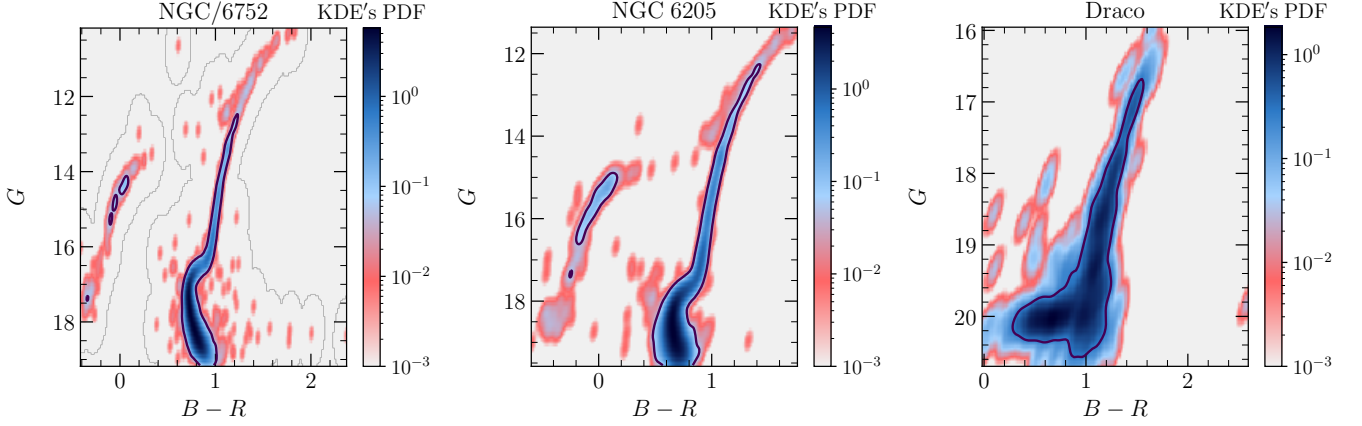
matrix of the proper motion PDF (i.e., eq [21]). After, we assigned the uncertainties of each parameter as the square root of the respective diagonal position of the inverted Hessian matrix. To these statistical uncertainties, one should expect to incorporate a systematic error at the level of  $\sim 0.025$  mas yr<sup>-1</sup> for GAIA EDR3 (as estimated by Lindegren et al. 2020 and Vasiliev & Baumgardt 2021) and of  $\sim 0.06$  mas yr<sup>-1</sup> for GAIA DR2 (Vasiliev 2019b).

## 2.5 Color magnitude diagram

Once one has a precise analytical description of both the surface density and the proper motion distribution of the ensemble of field stars plus the analyzed galactic object, it is much easier to extract tracer members by means of membership probabilities. However,

the incredible amount of astrophysical information from the GAIA releases allows to go even further, by analyzing how likely it is for a star to be part of the color-magnitude diagram of the tracer. It also allows to spot particular groups of stars such as binaries and blue stragglers (e.g., Leonard 1989) that lie away from the CMD.

The last step of our default filtering routine aims at constraining the region covered by the galactic object on the color-magnitude diagram (CMD). Since we do not have an analytical form to correctly describe the CMD, we opt, for the first time until now, to not use a Bayesian method, but rather a non-parametric Kernel Density Estimation (KDE) approach. We do so with the PYTHON method SCIPY.STATS.GAUSSIAN\_KDE, similarly to Vitral & Mamon (2021). Nevertheless, our approach has some few differences:



**Figure 4.** *Color-magnitude diagram fits:* We display the color-magnitude diagram (CMD) of NGC 6752, NGC 6205 (M 13) and the Draco dSph color-coded by the Kernel Density Estimation (KDE) non-parametric PDF. The black line contours indicate  $3\text{-}\sigma$  ( $2.5\text{-}\sigma$ , for Draco) confidence regions, which we used to filter out interlopers and binaries that lie away from the CMD.

- We first select stars with a membership probability greater than 0.8 (this limit is a modifiable parameter in BALRoGO).
- We used a KDE bandwidth of half the one derived by the Silverman’s rule (Silverman 1986).
- We used the membership probabilities of each star as weights to the KDE routine.
- We selected stars inside a  $3\text{-}\sigma$  density contour in the CMD (this  $n\text{-}\sigma$  limit is a modifiable parameter in BALRoGO, and set as 2.5 for dSphs in this article).

The CMD of NGC 6752, NGC 6205 (M 13) and the Draco dSph are presented in Figure 4, with the  $3\text{-}\sigma$  ( $2.5\text{-}\sigma$ , for the dSph) contour region displayed as a black thick line.

## 2.6 Mock data

For astronomers interested in generating mock data sets, BALRoGO has a `MOCK` method equipped with many capabilities such as (1) Projecting cartesian data into sky coordinates<sup>5</sup>, (2) Adding field stars uniformly distributed in an spherical cap and following a Pearson VII distribution in proper motions space, and (3) Adding realistic GAIA EDR3 errors to velocities. We describe below these three functionalities.

### 2.6.1 Coordinate transformation

BALRoGO’s method `CART6D_TO_GAIA` is able to convert cartesian coordinates into plane of sky coordinates by making use of `ASTROPY` (Astropy Collaboration et al. 2013) `COORDINATES` method, allowing the users to decide whether they want a realistic data set with proper motion uncertainties or not. Moreover, if the users want to shift a certain source from one  $(\alpha, \delta)$  position in the sky to another one, this can be promptly done by calling the method `ANGLE.TRANSROT_SOURCE`, which takes into account the spherical symmetry of the sky projection.

<sup>5</sup>  $\alpha, \delta, \mu_{\alpha,*}$  and  $\mu_{\delta}$

### 2.6.2 Field stars

When observing regions of the sky narrow enough (such as the two degrees cone searches we made in this work), one can expect to find uniformly distributed field stars in the field of view. By taking into account spherical trigonometry, one can generate such random distribution by inverting the probabilities  $\Pr\{r < R\}$  and  $\Pr\{\theta < \Theta\}$ , where  $R$  and  $\Theta$  are the angular distance of a tracer from the source’s center and the angle between this tracer and the source’s center with respect to the increasing declination axis. We have, thus:

$$R = \arccos [(1 - U) + U \cos (R_{\text{lim}})] , \quad (25a)$$

$$\Theta = 2\pi U , \quad (25b)$$

where  $R_{\text{lim}}$  is the maximum distance from the source’s center and  $U$  is a random variable uniformly distributed between zero and one. Similarly, one is able to generate random proper motions from a symmetric Pearson VII distribution ( $a_y = a_x$  in equation [24]) by inverting the probability  $\Pr\{|\mu| < M\}$ :

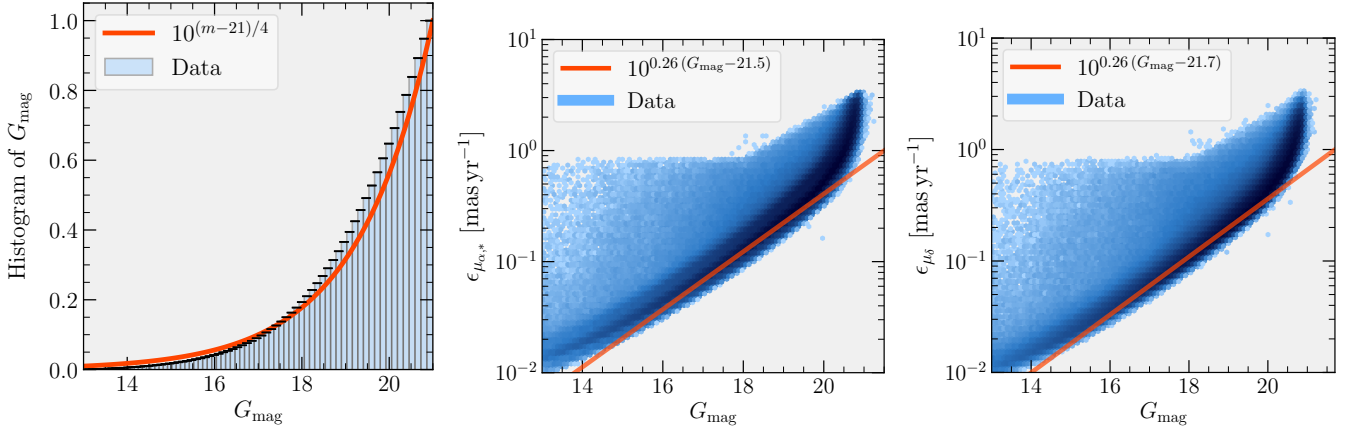
$$M = a \sqrt{U^{1/(1+\tau/2)} - 1} , \quad (26)$$

where  $a$  and  $\tau$  are the scale radius and slope from the Pearson VII distribution, and once again  $U$  is a random variable uniformly distributed between zero and one. We distribute those proper motion moduli azimuthally by choosing angles from a distribution such as the one from equation (25b). The precise derivation of the probabilities above is presented in appendix B.

### 2.6.3 Realistic GAIA EDR3 uncertainties

The proper motion uncertainties in GAIA EDR3 present a clear dependence on the apparent magnitude: In Figure 5, we stacked all GAIA EDR3 stars in a two degrees cone search around the nearby GCs NGC 6121 (M 4), NGC 5139 ( $\omega$  Cen), NGC 6397 and NGC 6752, in order to show that dependence. For that reason, to generate realistic GAIA EDR3 proper motion uncertainties, BALRoGO first generates random magnitudes, again, by inverting the probability





**Figure 5.** *GAIA EDR3 uncertainties:* On the **left**, we show the cumulative histogram of *GAIA EDR3*  $G$  magnitudes in blue, along with the function  $10^{(m-21)/4}$  in red. In the **middle**, we display the *GAIA EDR3*  $\mu_{\alpha,*}$  uncertainties blue color-coded by stellar counts, with the function  $10^{0.26(m-21.5)}$  in red. In the **right**, we display the *GAIA EDR3*  $\mu_{\delta}$  uncertainties blue color-coded by stellar counts, with the function  $10^{0.26(m-21.7)}$  in red. The *GAIA EDR3* data used for those plots is the stack of all stars in a two degrees cone search around the nearby globular clusters NGC 6121 (M 4), NGC 5139 ( $\omega$  Cen), NGC 6397 and NGC 6752.

$\Pr\{G_{\text{mag}} < m\}$ , which can be easily derived from the cumulative distribution of  $G_{\text{mag}}$  magnitudes from *GAIA EDR3* (i.e., the `phot_g_mean_mag` parameter), by imposing a threshold magnitude  $m_{\text{lim}}$ . The respective equations for a distribution of magnitude  $m$  are:

$$\Pr\{G_{\text{mag}} < m\} = 10^{(m-m_{\text{lim}})/f}, \quad (27a)$$

$$m = f \log_{10} U + m_{\text{lim}}, \quad (27b)$$

where  $U$  is a random variable uniformly distributed between zero and one, and  $m_{\text{lim}} = 21$  and  $f = 4$ , according to the fits displayed in Figure 5. Once one has a random set of magnitudes from equation (27b), we can again make use of the fits displayed in Figure 5 to assign:

$$\epsilon_{\mu_{\alpha,*}} = 10^{0.26(m-21.5)}, \quad (28a)$$

$$\epsilon_{\mu_{\delta}} = 10^{0.26(m-21.7)}. \quad (28b)$$

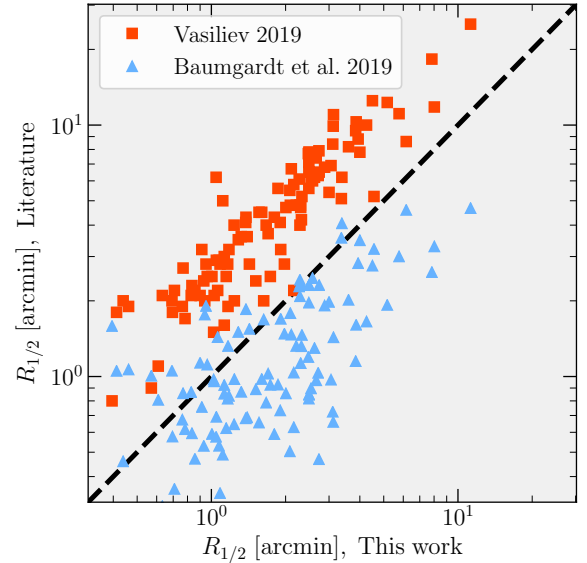
Those uncertainties are provided to the user, after adding Gaussian errors to  $\mu_{\alpha,*}$  and  $\mu_{\delta}$ , with respective standard deviation of  $\epsilon_{\mu_{\alpha,*}}$  and  $\epsilon_{\mu_{\delta}}$ .

### 3 RESULTS AND DISCUSSION

In Table C1, we display the main results of our fits, including our Bayesian estimates of the center, bulk proper motion and half number radii for over a hundred GCs and a few dSph galaxies measured by *GAIA EDR3*. This section aims to compare our results with previous estimates and discuss the implications of our new methods.

#### 3.1 Centers

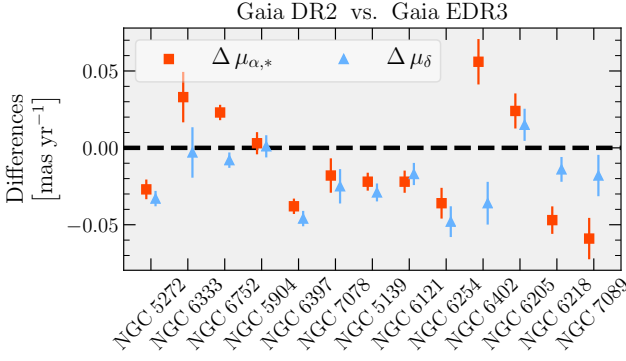
Our GC centers, derived according to section 2.1, were compared to the estimates from Goldsbury et al. (2010) for a robustness check: The median separation between their centers and ours is of 0.12



**Figure 6.** *Effective radii:* Comparison between the half number radii in arcmin derived by BALRoGO from *GAIA EDR3* data in the x-axis, and the same quantity derived by Baumgardt et al. (2019, blue triangles) and Vasiliev (2019a, red squares) from *GAIA DR2* data in the y-axis. We display the  $x = y$  line in dashed black.

arcsec, while the median of their reported uncertainties is of 0.2 arcsec. Similarly, we compared our estimates with the sources from the Harris (2010) catalog, and obtained a median separation of 0.67 arcsec.

This strengthens our initial assumption of small separations (i.e.,  $d \approx 0$ , in Figure 1), and gives us confidence in our Bayesian center estimation. The maximum separation between our measurements and those from Goldsbury et al. (2010) and Harris (2010) was of 26.5 arcsec, for the GC NGC 4147, followed by NGC 6553, with a separation of 26.1 arcsec and then by NGC 6558, with a separation of 9.5 arcsec.



**Figure 7.** *Improvement of the Gaia catalog:* Comparison of proper motion means derived by BALRoGO from GAIA DR2 and GAIA EDR3 data for the first ten globular clusters from the Messier catalogue plus NGC 6397, NGC 6752 and NGC 5139 ( $\omega$  Cen). We display, in the y-axis, the differences  $\Delta\mu_{\alpha,*} = \mu_{\alpha,*\text{DR2}} - \mu_{\alpha,*\text{EDR3}}$  and  $\Delta\mu_{\delta} = \mu_{\delta\text{DR2}} - \mu_{\delta\text{EDR3}}$  as red squares and blue triangles, respectively, for the 13 globular clusters mentioned above, distributed along the x-axis. The errors bars were calculated as explained in section 3.3, and we plot a dashed black horizontal line at the value of zero as a reference.

### 3.2 Effective radii

Figure 6 displays the effective (two-dimensional) radii derived from GAIA EDR3 by BALRoGO in the x-axis along with the same quantity derived from GAIA DR2 by Vasiliev (2019a) and Baumgardt et al. (2019), in red and blue respectively, in the y-axis. The conversion to projected half number radii was straightforward for both Sérsic and Plummer fits, since the scale radius used in both models was already the half number radius. For the Kazantzidis model, we multiplied the Kazantzidis scale radius  $a_K$  from equations 10a and 10b by 1.257 in order to retrieve the equivalent half number radius.

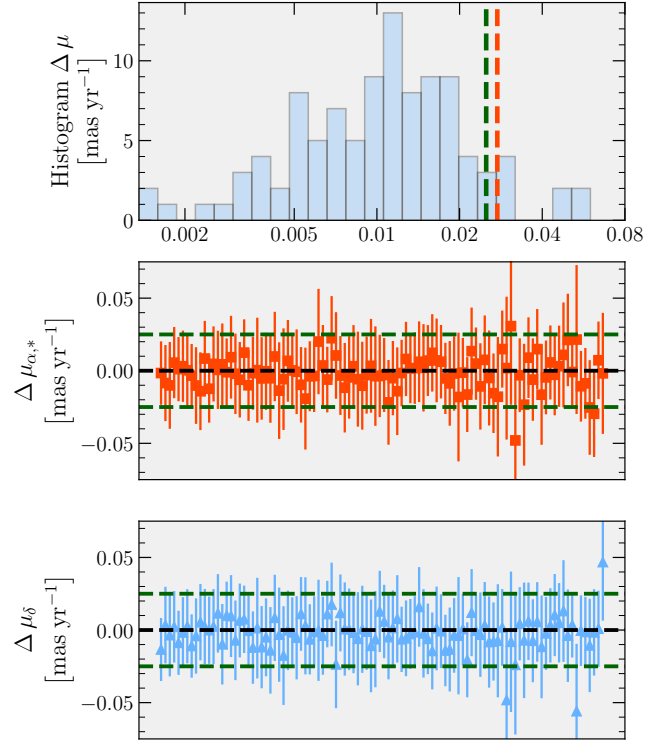
The values from Vasiliev (2019a) are in general higher, which he mentions in his Figure C.1 to be a consequence of the incompleteness in the central regions of his filtered GAIA catalog, which may lead his derived scale radius to be much larger than the actual half number radius of the cluster computed from all stars. In contrast, the measurements from Baumgardt et al. (2019) seem slightly smaller than ours, which in turn may indicate that our measurements are also slightly overestimated due to the intrinsic incompleteness of GAIA.

In any case, it is impressive that BALRoGO stands right in the middle of both estimates, which can be considered as a reliable indicator of an adequate goodness of fit. This highlights one of the strengths of our surface density fit method, which is taking into account a constant distribution of MW field stars, and neglecting any data filtering in this first step: It avoids a forced incompleteness towards both the cluster center and outskirts, due to crowding and fainter stars respectively, which are associated with worse astrometric solutions that would be filtered out in most filtering routines.

### 3.3 Proper motions

#### 3.3.1 Gaia DR2 vs. Gaia EDR3

In this subsection we stress the important difference between the bulk proper motions derived with GAIA DR2 and GAIA EDR3. In order to make such a comparison possible, we decided to perform the same analysis, but this time using GAIA DR2, for 13 GCs: the first ten GCs of the Messier catalogue, plus NGC 6397, NGC 6752 and



**Figure 8.** *Bulk proper motions:* Comparison of proper motion means derived from GAIA EDR3 data by BALRoGO and Vasiliev & Baumgardt (2021) for all the globular clusters in Table C1. The two *bottom plots* display, in the y-axis, the differences  $\Delta\mu_{\alpha,*} = \mu_{\alpha,*\text{this work}} - \mu_{\alpha,*\text{other work}}$  and  $\Delta\mu_{\delta} = \mu_{\delta\text{this work}} - \mu_{\delta\text{other work}}$  as red squares and blue triangles, respectively, for the globular clusters in Table C1, distributed along the x-axis in a similar fashion than Figure 7. In the *top plot*, we display the histogram of  $\Delta\mu = \sqrt{(\mu_{\alpha,*\text{other work}} - \mu_{\alpha,*\text{this work}})^2 + (\mu_{\delta\text{other work}} - \mu_{\delta\text{this work}})^2}$  in  $\text{mas yr}^{-1}$ , with the uncertainty floor of the GAIA EDR3 data reported in Vasiliev & Baumgardt (2021), of  $0.025 \text{ mas yr}^{-1}$  as a dashed green line and the median uncertainty in these differences, calculated with uncertainty propagation, as a dashed red line.

NGC 5139 ( $\omega$  Cen). We show, in Figure 7, the differences between the bulk proper motions ( $\mu_{\alpha,*}$ ,  $\mu_{\delta}$ ) estimated by BALRoGO from GAIA DR2 and from GAIA EDR3.

The uncertainty of those mean values was calculated as  $\epsilon = \sqrt{\epsilon_{\text{DR2}}^2 + \epsilon_{\text{EDR3}}^2}$ , where  $\epsilon_i$  stands for the uncertainties on the estimated values from the catalog  $i$ . One can notice that the disagreements, dominated by DR2 errors, lie in-between the uncertainty floor of the GAIA DR2 mission reported in Vasiliev (2019b), of  $\sim 0.06 \text{ mas yr}^{-1}$ . This is an important indicator of the improvement of GAIA EDR3, with more reliable measurements and a longer baseline, which in turns leads to smaller systematic uncertainties of the order of  $\sim 0.025 \text{ mas yr}^{-1}$  (e.g., Lindegren et al. 2020 and Vasiliev & Baumgardt 2021).

#### 3.3.2 Comparison with the literature

Gaia Collaboration et al. (2018), Baumgardt et al. (2019) and Vasiliev (2019a) measured bulk proper motions for over a hundred GCs with GAIA DR2 by using different methods previously mentioned. In this work, we update those information with the new GAIA EDR3, which has more precise and robust measurements of proper motions given

its longer baseline. Due to this fact, the values of proper motions means for nearly all GCs changed by more than their respective errors published in the works mentioned above. Therefore, it would be unfair to compare our results using GAIA EDR3 with their results obtained from GAIA DR2 modelling.

Fortunately, [Vasiliev & Baumgardt \(2021\)](#) recently provided bulk proper motion fits for 170 GCs, which allow for such comparison. We show, in Figure 8, the differences between the bulk proper motions ( $\mu_{\alpha,*}$ ,  $\mu_{\delta}$ ) estimated from GAIA EDR3 by BALRoGO and by their work in the two bottom panels, in a similar fashion than displayed in Figure 7, but for all GCs in Table C1. The upper plot displays the histogram of  $\Delta\mu = \sqrt{(\mu_{\alpha,*\text{VB21}} - \mu_{\alpha,*\text{V21}})^2 + (\mu_{\delta\text{VB21}} - \mu_{\delta\text{V21}})^2}$  in  $\text{mas yr}^{-1}$  (VB21 stands for [Vasiliev & Baumgardt 2021](#) and V21 for this work) with the uncertainty floor of the GAIA EDR3 data reported in [Lindegren et al. \(2020\)](#) and [Vasiliev & Baumgardt \(2021\)](#), of  $0.025 \text{ mas yr}^{-1}$  as a dashed green line and the median uncertainty in these differences, calculated with uncertainty propagation, as a dashed red line.

We observe a very good agreement between the measurements using BALRoGO and the measurements from [Vasiliev & Baumgardt \(2021\)](#): Most of the sources lie below the median uncertainty of  $\approx 0.03 \text{ mas yr}^{-1}$ , with the exception of NGC 6440, NGC 6453, NGC 6522, NGC 6528 and NGC 6540, which present a  $\Delta\mu \approx 0.05 \text{ mas yr}^{-1}$ , with high statistical uncertainties, on the order of  $0.05 \text{ mas yr}^{-1}$ . The reason for such disagreement is likely the small amount of tracers of NGC 6453, NGC 6522, NGC 6528, NGC 6540, which all have a small extension (i.e.,  $R_{1/2} \lesssim 0.6 \text{ arcmin}$ ), and are therefore more affected by data cleaning and also the increased amount of interlopers in the proper motion space of NGC 6440. In this plot, it is important to mention that the differences are certainly smaller than what could be expected from formal error bars, since both studies use the same EDR3 data (different from Figure 7), in which case the systematic uncertainty cancels out (it could be viewed as the calibration error on the proper motion zero-point, which varies across the sky at this level).

In addition to the comparison between GCs made above, our proper motion fits of dSphs provide a very good agreement with the estimates from [McConnachie & Venn \(2020a\)](#), with differences  $\lesssim 0.05 \text{ mas yr}^{-1}$  for all of our fits in Table C1. In the very late stages of this work, [Li et al. \(2021\)](#) also provided bulk proper motion fits<sup>6</sup> for 46 dSphs, presenting a very good overall agreement with our measurements, a part from the Bootes I dSph, for which our measurements are closer to those from [McConnachie & Venn \(2020a\)](#). This gives us confidence on the use of a non-Gaussian mixture in our Bayesian fit, along with the choice of a Pearson VII non-symmetric distribution of proper motions for interlopers, even though we neglected the convolution of the field stars component with Gaussian errors, after cleaning the data.

### 3.4 Convolution with Gaussian errors

As previously pointed out, our conservative data cleaning from section 2.3 should be enough to counter the fact that we do not convolve the Pearson VII distribution with Gaussian errors. In order to test this assumption, Pierre Boldrini kindly provided a GC mock using the initial condition  $N$ -body generator MAGI ([Miki & Umemura 2018](#)). Adopting a distribution-function-based method, it ensures that the

<sup>6</sup> They apply a similar method than presented in [Vasiliev \(2019a\)](#), therefore also relying on Gaussian mixtures.

**Table 1.** Comparison of estimates on the bulk  $\mu_{\alpha,*}$  and  $\mu_{\delta}$  from the clean and inaccurate mock data sets (see section 3.4).

Data set	$\mu_{\alpha,*}$	$\mu_{\delta}$
	[mas yr <sup>-1</sup> ]	[mas yr <sup>-1</sup> ]
Clean	$4.141 \pm 0.011$	$4.040 \pm 0.011$
Inaccurate	$4.146 \pm 0.012$	$4.039 \pm 0.012$

final realization of the cluster is in dynamical equilibrium ([Miki & Umemura 2018](#)). This GC mock was inspired by the real cluster NGC 6397, and therefore followed a Sérsic profile with stellar mass of  $1.17 \times 10^5 M_{\odot}$ , Sérsic radius of 3.14 pc, Sérsic index of 3.3 ([Vital & Mamon 2021](#)), along with orbital and tidal radius of 5.91 kpc ([Vasiliev 2019a](#)) and 59.9 pc (using the relation from [Bertin & Varri 2008](#)), respectively.

To this data, composed of 57500 stars, we added five times more stars, following spatial and velocity distributions of GAIA EDR3 field stars, as described in section 2.6. From this new data set, we created an extra one with realistic GAIA EDR3 proper motion uncertainties, constructed accordingly to section 2.6, in order to test if the lack of convolution with Gaussian errors could significantly impact our results. We randomly selected  $10^4$  stars in both subsets (hereafter clean and inaccurate subsets, for simplicity) and ran BALRoGO's routine on them. Figure 9 displays the proper motion fits of the inaccurate subset, similar to Figure 3.

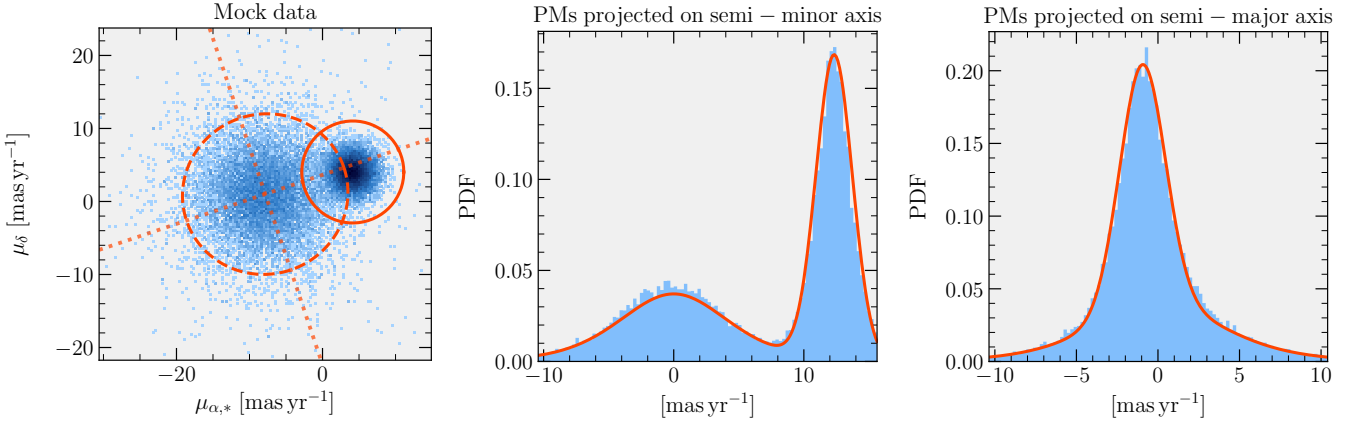
The fitted bulk  $\mu_{\alpha,*}$  and  $\mu_{\delta}$  from the clean and inaccurate subsets, along with their respective uncertainties are displayed in Table 1. We can verify that even without convolving the global proper motion PDF with Gaussian errors, the fits on the bulk proper motion from the data sets with and without uncertainties agree within the  $1-\sigma$  error bars, and their disagreement lies below the GAIA EDR3 uncertainty floor of  $0.025 \text{ mas yr}^{-1}$ . This strengthens our confidence in the proper motion cleaning routine from BALRoGO, as well as in its fits.

### 3.5 Kurtosis of the interlopers proper motions

This work confirms the tendency of field stars proper motions to follow a Pearson VII distribution, instead of a Gaussian, and we address now the interpretation of this result. Such an effect could be explained by the fact that the proper motions delivered by GAIA are not a direct measurement of the velocity, i.e. a measure of space variation per time, but rather a measurement of angular velocity, which neglects the distance of the stars.

According to the central limit theorem ([Laplace 1810](#)), when one considers the ensemble of independent and identically distributed random variables sharing the same dispersion and mean, their properly normalized sum tends toward a normal distribution regardless of the variables original distribution. In the case of GCs and dSph, this can be considered as an adequate approximation, since these sources contain generally tens or hundreds of thousands of stars ([Binney & Tremaine 2008](#)) which are located at distances that can be considered barely the same for a distant observer. This means that when converting the spatial velocities of its stars into angular velocities (i.e., dividing by their distance and turning them into proper motions), their originally quasi-independent and identically distributed velocities remain as such, as well as their similar dispersion, and therefore the variation around their mean is close to a Gaussian.

In the case of MW field stars however, since they have completely different distances, the distribution of proper motions is drawn away



**Figure 9.** *Mock data:* We display here the results of the proper motion fits of our mock data set with realistic GAIA EDR3 errors. The image on the first column displays the entire proper motion subset color-coded by stellar counts, from light blue to dark blue. The dashed ellipse displays the fitted Pearson VII symmetric distribution, with its main axis directions as two perpendicular dotted lines, while the continuous circle represents the galactic object (mock globular cluster) proper motion mean with a radius equals five times its intrinsic dispersion, for better visualization. The second and third columns display the fits (solid red) projected on the semi-minor and semi-major axis respectively, with the data in blue.

from an independent and identically distributed assumption, with each random measurement having a particular dispersion. In fact, their variation around their mean depends on the distance they lie from the observer, some of them much closer and others much farther than the galactic object analyzed. As a consequence, we expect to find more outliers, i.e., stars with proper motions that deviate more strongly from their mean, and thus a higher kurtosis (wider tails) in the distribution. That is why the Pearson VII distribution, with its wider tails, is better adapted to fit the proper motions of interlopers than a Gaussian distribution.

## 4 CONCLUSIONS

We present a new algorithm aimed at measuring some of the main structural parameters of galactic objects such as globular clusters and dwarf spheroidal galaxies. This algorithm, named BALRoGO: BAYESIAN ASTROMETRIC LIKELIHOOD RECOVERY OF GALACTIC OBJECTS, performs Bayesian and non-parametric fits in order to extract stars drowned in Milky Way interlopers, by accessing their membership probabilities (with respect to the galactic object analyzed).

Our approach presents innovative points which have been previously used in Vitral & Mamon (2021), but with some new improvements. Among the new approaches used in this algorithm, we highlight the following:

- Bayesian surface density fits considering a constant contribution from Milky Way interlopers.
- We allow the surface density modelling of globular clusters with profiles such as Sérsic and Kazantzidis, instead of the generally used King models (King 1966). Cored density profiles are well handled by a Plummer profile or low Sérsic indexes.
- The proper motion Bayesian fit is not based on a Gaussian mixture, but rather a Pearson VII distribution for interlopers, given their distribution wider tails.
- We use the membership probabilities of the surface density and proper motion fits in order to derive a non-parametric representation of the color-magnitude diagram, and then select cluster members inside confidence regions from this representation.

Comparisons between our method and previous works such as Baumgardt et al. (2019), Vasiliev (2019a), McConnachie & Venn (2020a) and Vasiliev & Baumgardt (2021) indicate strong agreement of bulk proper motions for globular clusters and dwarf spheroidal galaxies, and of scale radii for globular clusters. In addition, we make available our measurements of center, bulk proper motions and scale radii for over a hundred globular clusters from the NGC catalog, along with a few dwarf spheroidal galaxies in Table C1 (see data availability section).

The dynamics of such stellar systems is a fundamental key to understand some of the main aspects of galaxy evolution, as well as the astrophysical impacts of dark matter. With new releases of the GAIA astrometric mission, along with other astrometric data sets, the future of dynamical modeling is very enticing. For that reason, we believe it is important to make algorithms such as ours available, which can be easily used in order to derive important parameters of many stellar systems.

## ACKNOWLEDGEMENTS

First, I want to thank Alexandre Macedo for all his efforts towards making BALRoGO public, readable and accessible by anyone. I acknowledge the important comments of Gary Mamon, in particular for proposing equation (4) and providing equation (5) in a more complicated form than presented here. I also thank Pierre Boldrini, who provided the globular cluster mock data presented in section 3.4 and gave insightful comments on this work.

I acknowledge the referee, Eugene Vasiliev, for his comprehensive analysis of this work and important suggestions that improved the data cleaning and the proper motion fitting procedure. I finally thank Zhao Su for noticing a bad center fit for NGC 1904 in the submitted version.

Eduardo Vitral was funded by an AMX doctoral grant from École Polytechnique.

This work greatly benefited from public software packages such as NUMPY (van der Walt et al. 2011), ASTROPY (Astropy Collaboration et al. 2013), EMCEE (Goodman & Weare 2010), MATPLOTLIB (Hunter

2007), as well as the SPYDER Integrated Development Environment (Copyright (c) 2009- Spyder Project Contributors and others).

## DATA AVAILABILITY

The data that support the plots within this paper and other findings of this study are available from the corresponding author upon reasonable request. Table C1 can be accessed in text format at [https://gitlab.com/eduardo-vitral/balrogo/-/raw/master/table\\_gc\\_dsph.dat](https://gitlab.com/eduardo-vitral/balrogo/-/raw/master/table_gc_dsph.dat). The code repository link of BALRoGO, on GitLab, is <https://gitlab.com/eduardo-vitral/balrogo>.

## REFERENCES

- Abadi M. G., Navarro J. F., Steinmetz M., 2006, *MNRAS*, **365**, 747
- Akaike H., 1973, *Information Theory and an Extension of the Maximum Likelihood Principle*. Springer New York, New York, NY, pp 199–213
- Arenou F., et al., 2018, *A&A*, **616**, A17
- Astropy Collaboration et al., 2013, *A&A*, **558**, A33
- Battaglia G., Helmi A., Breddels M., 2013, *New Astron. Rev.*, **57**, 52
- Baumgardt H., Hilker M., Sollima A., Bellini A., 2019, *MNRAS*, **482**, 5138
- Bertin G., Varri A. L., 2008, *ApJ*, **689**, 1005
- Binney J., Tremaine S., 2008, *Galactic Dynamics: Second Edition*. Princeton University Press, Princeton, NJ, USA
- Boldrini P., Miki Y., Wagner A. Y., Mohayaee R., Silk J., Arbey A., 2020, *MNRAS*, **492**, 5218
- Bullock J. S., Johnston K. V., 2005, *ApJ*, **635**, 931
- Bustos Fierro I. H., Calderón J. H., 2019, *MNRAS*, **488**, 3024
- Carretta E., et al., 2009, *A&A*, **505**, 117
- Gaia Collaboration et al., 2018, *A&A*, **616**, A12
- Galadí-Enriquez D., Jordi C., Trullols E., 1998, *VizieR Online Data Catalog*, pp J/A+A/337/125
- Goldsbury R., Richer H. B., Anderson J., Dotter A., Sarajedini A., Woodley K., 2010, *AJ*, **140**, 1830
- Goodman J., Weare J., 2010, *Communications in Applied Mathematics and Computational Science*, **5**, 65
- Harris W. E., 2010, arXiv e-prints, p. arXiv:1012.3224
- Hernquist L., 1990, *ApJ*, **356**, 359
- Hunter J. D., 2007, *Computing in Science & Engineering*, **9**, 90
- Kazantzidis S., Mayer L., Mastropietro C., Diemand J., Stadel J., Moore B., 2004, *ApJ*, **608**, 663
- King I. R., 1966, *AJ*, **71**, 64
- Laplace P., 1810, *Mémoires de l'Académie Royale des Sciences de Paris*, **10**, 301
- Leonard P. J. T., 1989, *AJ*, **98**, 217
- Li H., Hammer F., Babusiaux C., Pawlowski M. S., Yang Y., Arenou F., Du C., Wang J., 2021, arXiv e-prints, p. arXiv:2104.03974
- Lindgren L., et al., 2018, *A&A*, **616**, A2
- Lindgren L., et al., 2020, arXiv e-prints, p. arXiv:2012.03380
- Marín-Franch A., et al., 2009, *ApJ*, **694**, 1498
- McConnachie A. W., Venn K. A., 2020a, *Research Notes of the American Astronomical Society*, **4**, 229
- McConnachie A. W., Venn K. A., 2020b, *AJ*, **160**, 124
- Miki Y., Umemura M., 2018, *MNRAS*, **475**, 2269
- Noyola E., Gebhardt K., Bergmann M., 2008, *ApJ*, **676**, 1008
- Pace A. B., Li T. S., 2019, *ApJ*, **875**, 77
- Peñarrubia J., Walker M. G., Gilmore G., 2009, *MNRAS*, **399**, 1275
- Pearson K., 1916, *Philosophical Transactions of the Royal Society of London Series A*, **216**, 429
- Peebles P. J. E., 1984, *ApJ*, **277**, 470
- Peebles P. J. E., Dicke R. H., 1968, *ApJ*, **154**, 891
- Plummer H. C., 1911, *MNRAS*, **71**, 460
- Riello M., et al., 2020, arXiv e-prints, p. arXiv:2012.01916
- Searle L., Zinn R., 1978, *ApJ*, **225**, 357

- Sérsic J. L., 1963, *Bull. Assoc. Argentina de Astron.*, **6**, 41
- Sérsic J. L., 1968, *Atlas de galaxias australes*. Cordoba, Argentina: Observatorio Astronómico
- Silverman B. W., 1986, *Density estimation for statistics and data analysis*
- Sugiura N., 1978, *Communications in Statistics - Theory and Methods*, **7**, 13
- Vasilevskis S., Sanders W. L., van Altena W. F., 1965, *AJ*, **70**, 806
- Vasiliev E., 2019a, *MNRAS*, **484**, 2832
- Vasiliev E., 2019b, *MNRAS*, **489**, 623
- Vasiliev E., Baumgardt H., 2021, arXiv e-prints, p. arXiv:2102.09568
- Vitral E., Mamon G. A., 2020, *A&A*, **635**, A20
- Vitral E., Mamon G. A., 2021, *A&A*, **646**, A63
- Walker M., 2013, *Dark Matter in the Galactic Dwarf Spheroidal Satellites*. p. 1039, doi:10.1007/978-94-007-5612-0\_20
- van der Marel R. P., Anderson J., 2010, *ApJ*, **710**, 1063
- van der Walt S., Colbert S. C., Varoquaux G., 2011, *Computing in Science Engineering*, **13**, 22

## APPENDIX A: PROJECTED NUMBER

We present here the solution of equation 4, for a distribution following the Plummer (Plummer 1911) profile, in the approximation of small cone apertures in the sky (i.e.  $R_{\max} \ll 1$  radian, where  $R_{\max}$  and  $d$  can be seen in Figure 1). For the case where  $R \leq R_{\max} - d$  ( $d$  also defined in Figure 1), one has equation 8.

However, whenever  $R > R_{\max} - d$ , the indefinite integral of equation 4 yields:

$$\begin{aligned} \tilde{N}(x) = & -\frac{a^2 \arccos \left[ \frac{R^2 + d^2 - R_{\max}^2}{2Rd} \right]}{\pi(a^2 + R^2)} + \\ & \left( \sqrt{\Xi_1} \left( -\sqrt{\Xi_2} \times \right. \right. \\ & \left. \left. \arctan \left[ \frac{(d^2 - R_{\max}^2)^2 - R^2 (d^2 + R_{\max}^2)}{(d^2 - R_{\max}^2) \sqrt{\Xi_1}} \right] + \right. \right. \\ & \left. \left. (a^2 + d^2 - R_{\max}^2) \arctan \left[ \frac{\Xi_3}{\sqrt{\Xi_2} \sqrt{\Xi_1}} \right] \right) \right) \div \\ & (2\pi \sqrt{\Xi_2} \sqrt{\Xi_1}) + C. \end{aligned} \quad (\text{A1})$$

where  $C = 1$  is an integration constant,  $a$  is the Plummer effective radius, and  $\Xi_1$ ,  $\Xi_2$  and  $\Xi_3$  are defined as:

$$\Xi_1 = -d^4 - (R^2 - R_{\max}^2)^2 + 2d^2 (R^2 + R_{\max}^2) \quad (\text{A2a})$$

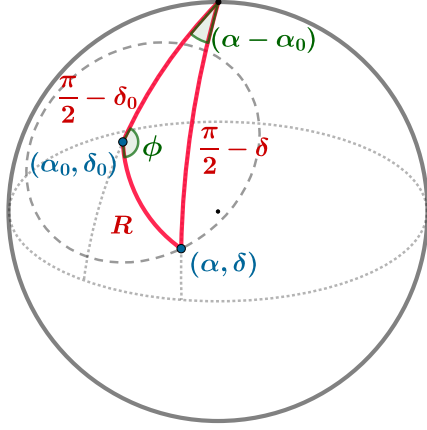
$$\Xi_2 = a^4 + (d^2 - R_{\max}^2)^2 + 2a^2 (d^2 + R_{\max}^2) \quad (\text{A2b})$$

$$\Xi_3 = (d^2 - R_{\max}^2)^2 + a^2 (d^2 + R_{\max}^2) - R^2 (a^2 + d^2 + R_{\max}^2) \quad (\text{A2c})$$

This treatment can be chosen in the BALRoGO method `POSITION.FIND_CENTER()`, by providing the argument `method="ml_robust"`.

## APPENDIX B: FIELD STARS MOCK DATA

In this section, we derive equations (25) and (26).



**Figure B1.** *Spherical geometry:* Representation of the physical situation of a source projected in the plane of sky.

## B1 Random positions

In order to generate  $n$  points uniformly distributed in a spherical cap, we shall first considerate Figure B1, with the geometry of the problem. From classical spherical trigonometry relations, it is straightforward to write:

$$\cos R = \sin \delta \sin \delta_0 + \cos \delta \cos \delta_0 \cos (\alpha - \alpha_0), \quad (\text{B1a})$$

$$\sin \phi = \frac{\cos \delta \sin (\alpha - \alpha_0)}{\sin R}, \quad (\text{B1b})$$

$$\sin \delta = \cos R \sin \delta_0 + \sin R \cos \delta_0 \cos \phi. \quad (\text{B1c})$$

We wish to generate an uniform distribution of points in a spherical cap of radius  $R_{\text{lim}}$ , so the probability of having a radius smaller than  $R$  can be written as:

$$\Pr\{r < R\} = \frac{\text{Surface}(R)}{\text{Surface}(R_{\text{lim}})}, \quad (\text{B2})$$

or, more precisely:

$$\Pr\{r < R\} = \frac{\int_0^R \int_0^{2\pi} \rho^2 \sin \theta \, d\theta \, d\varphi}{\int_0^{R_{\text{lim}}} \int_0^{2\pi} \rho^2 \sin \theta \, d\theta \, d\varphi} = \frac{1 - \cos R}{1 - \cos R_{\text{lim}}}, \quad (\text{B3})$$

where  $\rho$  is the radius of the sphere,  $\theta$  and  $\varphi$  are the longitudinal and latitudinal angles of the sphere, that in our description have their origin at  $(\alpha_0, \delta_0)$ . Similarly, the probability that  $\phi$  is smaller than an angle  $\Theta$  is:

$$\Pr\{\phi < \Theta\} = \Theta/2\pi. \quad (\text{B4})$$

Therefore, in order to derive equations (25), one just needs to invert the relations (B3) and (B4), with respect to  $R$  and  $\Theta$ , respectively. The following step, which is to convert the set of known  $\alpha_0, \delta_0, R$  and

$\phi$  into pairs of  $(\alpha, \delta)$  is done by first deriving  $\delta$  with equation (B1c) and then  $\alpha$  with:

$$\alpha = \begin{cases} \alpha_0 + \arccos(\Lambda_1) & , \text{ for } \Lambda_2 > 0 \text{ and } 0 \leq \phi \leq \pi \\ \alpha_0 + \arccos(-\Lambda_1) & , \text{ for } \Lambda_2 < 0 \text{ and } 0 \leq \phi \leq \pi \\ \alpha_0 - \arccos(\Lambda_1) & , \text{ for } \Lambda_2 > 0 \text{ and } \pi < \phi < 2\pi \\ \alpha_0 - \arccos(-\Lambda_1) & , \text{ for } \Lambda_2 < 0 \text{ and } \pi < \phi < 2\pi \end{cases} \quad (\text{B5})$$

where we have the following correspondences:

$$\Lambda_1 = \sqrt{1 - \frac{\sin^2 \phi \sin^2 R}{1 - \sin^2 \delta}} \quad (\text{B6a})$$

$$\Lambda_2 = \frac{\cos R - \sin \delta \sin \delta_0}{\cos \delta \cos \delta_0} \quad (\text{B6b})$$

## B2 Random proper motions

For the field stars, we generated random variables that followed a symmetric Pearson VII distribution. The statistical approach to do so was to invert the probability that a field star proper motion modulus is smaller than  $M$ , i.e. the cumulative distribution function of  $M$  for a symmetric Pearson VII distribution of scale radius  $a$  and characteristic slope  $\tau$ :

$$\begin{aligned} \text{CDF}(M) &\equiv \int_0^M f_{\text{PM}}(\mu) \, d\mu \\ &= - \int_0^M \frac{\tau + 2}{a} \frac{\mu}{a} \left[ 1 + \left( \frac{\mu}{a} \right)^2 \right]^{\tau/2} \, d\mu. \end{aligned} \quad (\text{B7})$$

where  $f_{\text{PM}}$  is the distribution function of proper motions moduli for the Pearson VII symmetric distribution. The result of the integral above is:

$$\text{CDF}(M) = 1 - \left[ 1 + (M/a)^2 \right]^{1+\tau/2}. \quad (\text{B8})$$

Thus, if  $U$  is a uniform random variable with boundaries  $[0, 1]$  (and thus,  $1 - U \equiv U$ , random PM variables can be generated with equation (26).

## APPENDIX C: STRUCTURAL PARAMETERS OF GLOBULAR CLUSTERS AND DWARF SPHEROIDAL GALAXIES

We present the main results of our paper in Table C1, such as effective radii and bulk proper motions for over a hundred globular clusters and some of the main Local Group dwarf spheroidal galaxies. This table can be accessed in text format at the footnote link<sup>7</sup>.

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.

<sup>7</sup> [https://gitlab.com/eduardo-vitral/balrogo/-/raw/master/table\\_gc\\_dsph.dat](https://gitlab.com/eduardo-vitral/balrogo/-/raw/master/table_gc_dsph.dat)

**Table C1.** Catalog of proper motions and other dynamical parameters of NGC globular clusters and Local Group dwarf spheroidal galaxies.

Name	Other ID	$\alpha$ [deg]	$\delta$ [deg]	$D$ [kpc]	$\mu_{\alpha,*}$ [mas yr <sup>-1</sup> ]	$\mu_{\delta}$ [mas yr <sup>-1</sup> ]	$\epsilon_{\mu_{\alpha,*}}$ [mas yr <sup>-1</sup> ]	$\epsilon_{\mu_{\delta}}$ [mas yr <sup>-1</sup> ]	$\Sigma$	$R_{1/2}$ [arcmin]	$\epsilon_{R_{1/2}}$ [arcmin]
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
NGC 104	47 Tuc	6.02242	-72.08147	4.43	5.250	-2.565	0.003	0.003	S	7.84	0.02
NGC 1261	-	48.06694	-55.21592	17.20	1.602	-2.065	0.005	0.005	P	1.38	0.01
NGC 1851	-	78.52821	-40.04675	11.33	2.140	-0.664	0.004	0.005	P	2.08	0.01
NGC 1904	M 79	81.04414	-24.52423	13.27	2.475	-1.591	0.005	0.005	P	1.56	0.01
NGC 2298	-	102.24756	-36.00532	10.80	3.311	-2.172	0.006	0.006	S	1.18	0.02
NGC 2419	-	114.53541	38.88193	83.18	-0.004	-0.526	0.011	0.011	S	0.91	0.02
NGC 2808	-	138.01293	-64.86349	10.21	1.000	0.276	0.005	0.005	S	3.11	0.01
NGC 288	-	13.18819	-26.58238	9.00	4.156	-5.706	0.004	0.004	S	2.57	0.02
NGC 3201	-	154.40393	-46.41249	4.60	8.348	-1.966	0.002	0.002	S	4.50	0.03
NGC 362	-	15.80943	-70.84885	9.17	6.683	-2.543	0.010	0.010	S	2.49	0.01
NGC 4147	-	182.51853	18.54192	18.20	-1.716	-2.103	0.010	0.010	P	0.86	0.02
NGC 4372	-	186.43919	-72.65907	5.76	-6.410	3.298	0.003	0.003	S	4.00	0.03
NGC 4590	M 68	189.86660	-26.74405	10.13	-2.736	1.777	0.004	0.004	P	2.11	0.02
NGC 4833	-	194.89122	-70.87650	6.56	-8.377	-0.962	0.006	0.006	S	2.48	0.02
NGC 5024	M 53	198.23025	18.16819	17.90	-0.142	-1.332	0.004	0.004	S	2.47	0.02
NGC 5053	-	199.11287	17.69872	17.20	-0.331	-1.217	0.004	0.005	S	2.29	0.03
NGC 5139	$\omega$ Cen	201.69698	-47.47950	5.24	-3.253	-6.757	0.003	0.003	S	11.25	0.03
NGC 5272	M 3	205.54873	28.37727	9.59	-0.153	-2.666	0.004	0.003	P	3.85	0.02
NGC 5286	-	206.61171	-51.37425	11.45	0.188	-0.157	0.006	0.006	S	1.90	0.01
NGC 5466	-	211.36366	28.53444	16.00	-5.359	-0.843	0.008	0.009	P	2.30	0.03
NGC 5634	-	217.40528	-5.97638	27.20	-1.689	-1.473	0.008	0.007	P	0.83	0.02
NGC 5694	-	219.90217	-26.53835	37.33	-0.481	-1.107	0.033	0.028	P	0.63	0.02
NGC 5824	-	225.99421	-33.06854	30.90	-1.204	-2.228	0.006	0.006	S	1.11	0.04
NGC 5897	-	229.35163	-21.01013	12.60	-5.413	-3.390	0.006	0.006	S	2.29	0.02
NGC 5904	M 5	229.63841	2.08097	7.57	4.073	-9.869	0.004	0.004	S	4.26	0.02
NGC 5927	-	232.00284	-50.67305	9.08	-5.051	-3.214	0.007	0.007	S	3.13	0.06
NGC 5946	-	233.86904	-50.65973	10.60	-5.317	-1.646	0.012	0.012	S	1.05	0.03
NGC 5986	-	236.51248	-37.78644	10.56	-4.193	-4.555	0.008	0.008	S	1.86	0.01
NGC 6093	M 80	244.26003	-22.97611	8.86	-2.930	-5.588	0.009	0.009	P	1.80	0.02
NGC 6101	-	246.45171	-72.20161	12.80	1.761	-0.258	0.003	0.003	P	2.27	0.03
NGC 6121	M 4	245.89669	-26.52584	1.93	-12.515	-19.011	0.004	0.004	S	6.17	0.09
NGC 6139	-	246.91658	-38.84919	9.80	-6.073	-2.701	0.010	0.010	S	1.40	0.05
NGC 6144	-	246.80774	-26.02352	8.90	-1.746	-2.614	0.009	0.009	P	1.63	0.02
NGC 6171	M 107	248.13274	-13.05381	5.70	-1.945	-5.973	0.006	0.005	S	2.09	0.02
NGC 6205	M 13	250.42348	36.46129	6.77	-3.137	-2.566	0.003	0.003	S	3.87	0.01
NGC 6218	M 12	251.80908	-1.94856	4.67	-0.201	-6.803	0.004	0.004	S	2.88	0.01
NGC 6229	-	251.74419	47.52642	30.62	-1.173	-0.456	0.009	0.009	P	0.71	0.01
NGC 6235	-	253.35565	-22.17748	13.52	-3.942	-7.590	0.012	0.012	S	1.03	0.02
NGC 6254	M 10	254.28768	-4.10033	4.96	-4.760	-6.609	0.008	0.008	P	3.60	0.02
NGC 6256	-	254.88617	-37.12135	6.40	-3.714	-1.635	0.018	0.017	K	1.17	0.04
NGC 6266	M 62	255.30247	-30.11237	6.41	-4.982	-2.962	0.011	0.011	S	3.06	0.04
NGC 6273	M 19	255.65702	-26.26793	8.27	-3.249	1.656	0.008	0.008	S	2.62	0.02
NGC 6284	-	256.11976	-24.76424	15.14	-3.206	-2.016	0.011	0.011	S	1.00	0.03
NGC 6287	-	256.28936	-22.70776	9.40	-5.002	-1.875	0.010	0.009	P	1.16	0.02
NGC 6293	-	257.54342	-26.58174	8.70	0.878	-4.322	0.012	0.011	P	1.51	0.03

*Notes:* Columns are (1): Source name; (2): Alternative ID; (3): right ascension derived according to section 2.1; (4): declination derived according to section 2.1; (5): Distances in kpc from Baumgardt et al. (2019); (6): Bulk proper motion in right ascension (i.e.,  $[d\alpha/dt] \cos \delta$ , in mas yr<sup>-1</sup>); (7): Bulk proper motion in declination (i.e.,  $d\delta/dt$ , in mas yr<sup>-1</sup>); (8): uncertainty of  $\mu_{\alpha,*}$ , in mas yr<sup>-1</sup>; (9): uncertainty of  $\mu_{\delta}$ , in mas yr<sup>-1</sup>; (10): Surface density model preferred by AICc (see section 2.2). 'P' stands for Plummer, 'S' for Sérsic and 'K' for Kazantzidis; (11): Effective (two-dimensional) radius, in arcmin; (12): uncertainty on the effective radius, in arcmin.

Table C1 – continued

Name	Other ID	$\alpha$	$\delta$	$D$	$\mu_{\alpha,*}$	$\mu_{\delta}$	$\epsilon_{\mu_{\alpha,*}}$	$\epsilon_{\mu_{\delta}}$	$\Sigma$	$R_{1/2}$	$\epsilon_{R_{1/2}}$
(1)	(2)	[deg]	[deg]	[kpc]	[mas yr <sup>-1</sup> ]	[mas yr <sup>-1</sup> ]	[mas yr <sup>-1</sup> ]	[mas yr <sup>-1</sup> ]	(10)	[arcmin]	[arcmin]
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
NGC 6304	–	258.63435	–29.46203	5.77	–4.083	–1.092	0.017	0.016	S	0.95	0.07
NGC 6316	–	259.15589	–28.14002	11.60	–4.975	–4.611	0.016	0.016	P	0.96	0.08
NGC 6325	–	259.49692	–23.76607	7.80	–8.295	–9.008	0.013	0.012	S	0.78	0.03
NGC 6333	M 9	259.79907	–18.51627	8.40	–2.174	–3.223	0.010	0.010	S	2.30	0.03
NGC 6341	M 92	259.28081	43.13600	8.44	–4.935	–0.627	0.004	0.004	S	2.66	0.01
NGC 6342	–	260.29224	–19.58745	8.43	–2.904	–7.122	0.010	0.010	S	2.16	0.16
NGC 6352	–	261.37069	–48.42210	5.30	–2.167	–4.436	0.006	0.005	S	3.00	0.08
NGC 6355	–	260.99435	–26.35342	8.70	–4.758	–0.570	0.017	0.015	S	1.13	0.16
NGC 6356	–	260.89577	–17.81304	15.10	–3.765	–3.400	0.008	0.008	S	2.01	0.07
NGC 6362	–	262.97906	–67.04835	7.36	–5.509	–4.769	0.002	0.003	S	2.74	0.02
NGC 6366	–	261.93433	–5.07988	3.79	–0.335	–5.161	0.004	0.004	S	3.38	0.04
NGC 6380	–	263.61666	–39.06918	9.80	–2.162	–3.233	0.018	0.018	S	0.90	0.05
NGC 6388	–	264.07275	–44.73566	10.74	–1.313	–2.712	0.008	0.008	S	2.48	0.02
NGC 6397	–	265.17540	–53.67441	2.44	3.254	–17.653	0.003	0.003	S	5.77	0.05
NGC 6401	–	264.65386	–23.90874	7.70	–2.765	1.438	0.021	0.019	S	0.69	0.05
NGC 6402	M 14	264.40062	–3.24594	9.31	–3.575	–5.057	0.007	0.007	S	2.21	0.01
NGC 6426	–	266.22795	3.17013	19.80	–1.805	–2.981	0.012	0.012	P	0.83	0.03
NGC 6440	–	267.21945	–20.35960	8.24	–1.188	–3.973	0.021	0.020	S	1.07	0.06
NGC 6441	–	267.55440	–37.05147	11.83	–2.541	–5.372	0.010	0.010	P	1.05	0.03
NGC 6453	–	267.71571	–34.59866	11.60	0.205	–5.982	0.033	0.032	K	0.41	0.04
NGC 6496	–	269.76531	–44.26599	9.12	–3.064	–9.259	0.005	0.005	P	1.43	0.03
NGC 6517	–	270.45989	–8.95957	10.60	–1.564	–4.470	0.011	0.015	S	1.15	0.06
NGC 6522	–	270.89201	–30.03400	8.00	2.583	–6.491	0.034	0.030	S	0.46	0.03
NGC 6528	–	271.20670	–30.05580	7.45	–2.128	–5.659	0.036	0.029	K	0.57	0.04
NGC 6535	–	270.96044	–0.29765	6.50	–4.219	–2.938	0.009	0.009	K	1.06	0.03
NGC 6539	–	271.20722	–7.58588	7.85	–6.893	–3.539	0.010	0.010	S	1.91	0.08
NGC 6540	–	271.53581	–27.76529	5.20	–3.749	–2.819	0.035	0.034	P	0.40	0.06
NGC 6541	–	272.00984	–43.71493	7.95	0.285	–8.843	0.005	0.005	S	2.71	0.02
NGC 6544	–	271.83574	–24.99742	2.60	–2.309	–18.610	0.011	0.010	K	4.57	0.10
NGC 6553	–	272.31533	–25.90775	6.75	0.358	–0.441	0.021	0.019	P	1.29	0.09
NGC 6558	–	272.57642	–31.76354	7.20	–1.744	–4.152	0.021	0.020	K	0.61	0.05
NGC 6569	–	273.41198	–31.82648	10.59	–4.135	–7.353	0.011	0.011	P	0.77	0.04
NGC 6584	–	274.65667	–52.21581	13.18	–0.093	–7.204	0.007	0.006	S	1.32	0.02
NGC 6624	–	275.91879	–30.36106	7.40	0.127	–6.947	0.014	0.014	P	1.24	0.04
NGC 6626	M 28	276.13704	–24.86987	5.43	–0.302	–8.930	0.016	0.016	S	1.70	0.12
NGC 6637	M 69	277.84622	–32.34811	8.80	–5.063	–5.834	0.011	0.011	P	1.73	0.03
NGC 6638	–	277.73436	–25.49643	10.32	–2.497	–4.079	0.014	0.014	S	0.76	0.04
NGC 6642	–	277.97596	–23.47616	8.05	–0.176	–3.898	0.016	0.014	P	0.70	0.04
NGC 6652	–	278.94010	–32.99074	10.00	–5.488	–4.261	0.008	0.008	S	0.93	0.02
NGC 6656	M 22	279.09980	–23.90477	3.23	9.840	–5.618	0.007	0.007	S	8.02	0.15
NGC 6681	M 70	280.80317	–32.29214	9.31	1.440	–4.738	0.008	0.006	P	1.67	0.03
NGC 6712	–	283.26804	–8.70596	6.95	3.356	–4.443	0.009	0.009	P	1.98	0.05
NGC 6715	M 54	283.76386	–30.47986	24.13	–2.682	–1.381	0.003	0.003	S	2.73	0.13
NGC 6717	Pal 9	283.77515	–22.70150	7.10	–3.147	–5.013	0.010	0.010	S	0.94	0.04
NGC 6723	–	284.88813	–36.63226	8.30	1.024	–2.417	0.005	0.005	S	2.33	0.02
NGC 6749	–	286.31399	1.89998	7.80	–2.844	–5.997	0.012	0.012	P	1.38	0.05
NGC 6752	–	287.71710	–59.98458	4.25	–3.163	–4.034	0.003	0.003	P	5.17	0.02
NGC 6760	–	287.80024	1.03046	7.95	–1.091	–3.610	0.008	0.008	S	2.32	0.10
NGC 6779	M 56	289.14820	30.18348	9.68	–2.010	1.612	0.005	0.005	S	1.60	0.01



Table C1 – *continued*

Name	Other ID	$\alpha$	$\delta$	$D$	$\mu_{\alpha,*}$	$\mu_{\delta}$	$\epsilon_{\mu_{\alpha,*}}$	$\epsilon_{\mu_{\delta}}$	$\Sigma$	$R_{1/2}$	$\epsilon_{R_{1/2}}$
(1)	(2)	[deg]	[deg]	[kpc]	[mas yr <sup>-1</sup> ]	[mas yr <sup>-1</sup> ]	[mas yr <sup>-1</sup> ]	[mas yr <sup>-1</sup> ]	(10)	[arcmin]	[arcmin]
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
NGC 6809	M 55	294.99878	-30.96479	5.30	-3.430	-9.314	0.003	0.003	S	3.94	0.02
NGC 6838	M 71	298.44369	18.77918	3.99	-3.414	-2.656	0.004	0.004	S	3.37	0.12
NGC 6864	M 75	301.52017	-21.92228	21.61	-0.591	-2.794	0.009	0.009	P	1.08	0.01
NGC 6934	–	308.54736	7.40445	15.40	-2.653	-4.691	0.015	0.016	P	1.24	0.02
NGC 6981	M 72	313.36541	-12.53734	17.00	-1.266	-3.360	0.006	0.005	P	1.13	0.01
NGC 7006	–	315.37278	16.18790	40.10	-0.144	-0.645	0.021	0.023	P	0.44	0.01
NGC 7078	M 15	322.49304	12.16699	10.22	-0.652	-3.808	0.005	0.005	P	3.12	0.01
NGC 7089	M 2	323.36257	-0.82332	10.51	3.447	-2.174	0.010	0.010	S	2.54	0.01
NGC 7099	M 30	325.09221	-23.17988	8.00	-0.742	-7.301	0.006	0.005	S	2.16	0.01
NGC 7492	–	347.11116	-15.61142	26.55	0.763	-2.319	0.008	0.009	S	1.02	0.03
Bootes dSph	Boo dSph	210.00004	14.49996	–	-0.399	-1.065	0.018	0.016	P	12.45	2.86
Carina dSph	PGC 19441	100.40293	-50.96613	–	0.528	0.118	0.008	0.008	P	8.71	0.30
Draco dSph	UGC 10822	260.05981	57.92121	–	0.032	-0.183	0.008	0.009	S	6.98	0.24
Fornax dSph	ESO 356-4	39.99709	-34.44920	–	0.379	-0.358	0.002	0.002	S	14.80	0.07
Leo I dSph	PGC 29488	152.11721	12.30651	–	-0.078	-0.090	0.014	0.014	S	2.93	0.05
Leo II dSph	PGC 34176	168.36717	22.15282	–	-0.092	-0.143	0.027	0.026	S	2.41	0.09
Sculptor dSph	PGC 3589	15.03901	-33.70893	–	0.103	-0.149	0.003	0.003	P	9.71	0.11
Sextans dSph	LEDA 88608	153.26204	-1.61463	–	-0.360	0.031	0.015	0.013	P	21.39	1.28
Ursa Minor dSph	PGC 54074	227.29745	67.21439	–	-0.133	0.045	0.010	0.010	S	13.97	0.74