

---

# Testing the limits of the ELT

*Release private beta*

**Alarich Herzner**

Dec 13, 2021



# CONTENTS:

<b>1 Abstract</b>	<b>1</b>
<b>2 Acknowledgement</b>	<b>3</b>
<b>3 Introduction</b>	<b>5</b>
<b>4 Galactic Potential</b>	<b>7</b>
4.1 Circular Velocity . . . . .	8
4.2 Angular Velocity . . . . .	10
4.3 Mass Distribution . . . . .	10
4.4 Surface Mass Density . . . . .	10
<b>5 Initialization</b>	<b>13</b>
5.1 Sampling distribution functions . . . . .	13
5.2 Cluster . . . . .	14
5.3 Initial Mass Function (IMF) . . . . .	14
5.4 Density Profile . . . . .	16
5.5 Field Stars (FS) . . . . .	16
5.6 Positions . . . . .	21
5.7 Velocities . . . . .	21
<b>6 Integration</b>	<b>31</b>
6.1 Quadrature, Adaptive, General-purpose (QAG) . . . . .	31
6.2 Quadrature, Adaptive, General-purpose, Infinite interval (QAGI) . . . . .	32
6.3 Velocity Verlet Algorithm . . . . .	32
6.4 Barnes-Hut Algorithm (BH) . . . . .	32
<b>7 Coordinate Systems</b>	<b>35</b>
7.1 Galactocentric Cartesian (GCA) . . . . .	35
7.2 Galactocentric Polar (GCP) . . . . .	35
7.3 Local Standard of Rest (LSR) . . . . .	36
7.4 Heliocentric Cartesian (HCA) . . . . .	36
7.5 Heliocentric Galactic Polar (HGP) . . . . .	37
7.6 Heliocentric Equatorial Polar (HEQ) . . . . .	37
7.7 Heliocentric Telescope Polar (HTP) . . . . .	38
<b>8 Database</b>	<b>39</b>
8.1 Entity Relationship Diagram . . . . .	40
<b>9 Mock Observations</b>	<b>41</b>
9.1 Creating Source Objects . . . . .	41

<b>10 Data Reduction</b>	<b>43</b>
10.1 Detecting Stars . . . . .	43
<b>11 Cluster analysis</b>	<b>45</b>
11.1 Velocity approximation . . . . .	45
11.2 DBSCAN . . . . .	46
11.3 Performance . . . . .	46
<b>12 Experiments</b>	<b>49</b>
12.1 Parameter optimization . . . . .	49
12.2 Setup . . . . .	50
12.3 Results and Interpretation . . . . .	50
<b>13 Conclusions</b>	<b>55</b>
<b>14 Appendix</b>	<b>57</b>
<b>15 Indices and tables</b>	<b>65</b>
<b>Bibliography</b>	<b>67</b>

---

**CHAPTER  
ONE**

---

**ABSTRACT**

Todo?: write



---

**CHAPTER  
TWO**

---

**ACKNOWLEDGEMENT**

Todo!: write



---

**CHAPTER  
THREE**

---

## **INTRODUCTION**

The Extremely Large Telescope (ELT) is currently under construction. This first next generation 40m class telescope will have the resolution and sensitivity needed to study the motions of individual stars in the galactic centre (GC). These position and motions can be used to detect star clusters using some type of clustering algorithm and consequently estimate their Initial Mass Function (IMF). Whether the IMF of star clusters is universal is subject of discussion [2] at this time. The study of Young Massive Clusters (YMCs) [29] near the GC will hopefully give more insight into this hotly debated topic. YMCs are tightly grouped clusters of stars, generally containing  $>10000$  members. All original members are still present, the IMF is well sampled in all mass regimes. These facts combine to give a good picture of the end product of a star formation event.

Given detected star clusters the question remains, how reliable these results are. If the true classification of the studied set of stars is known, the performance of the clustering algorithm and hence the reliability of the results can be calculated. However, this is generally not the case.

In this master thesis an N-body simulation containing cluster stars and field stars under the influence of the milky way potential is performed. Snapshots are taken at different timesteps and fed to ScopeSim [22] to create mock observations. With the help of Photutils [8] positions are extracted and the DBSCAN algorithm used to detect cluster and field stars. Finally, the reliability of the results can be determined by comparing its results with the initially simulated stars.



---

CHAPTER  
**FOUR**

---

## GALACTIC POTENTIAL

Modelling the gravitational potential of the Milky Way has been a ever evolving topic in recent decades, [39] provides a brief history. Here the potential is used to approximate the force on the simulated stars and to initialize field stars.

The model of choice consists of four parts: black hole, bulge, disc and the dark matter halo.

The **black hole** is represented by a Keplerian potential:

$$\Phi_{bh}(r) = -\frac{G * M_{bh}}{r}$$

with  $r$  being the spherical radius.

The **disk** can be modelled via a Miyamoto Nagai potential [49]

$$\begin{aligned} \Phi_{disk}(R, z) &= -\frac{G * M_{disk}}{\sqrt{R^2 + (a_{disk} + \sqrt{z^2 + b_{disk}^2})^2}} \\ \rho_{disk}(R, z) &= \frac{b_{disk}^2 M_{disk}}{4\pi} \frac{a_{disk} R^2 + \left[ a_{disk} + 3(z^2 + b_{disk}^2)^{\frac{1}{2}} \right] \left[ a_{disk} + (z^2 + b_{disk}^2)^{\frac{1}{2}} \right]^2}{\left\{ R^2 + \left[ a_{disk} + (z^2 + b_{disk}^2)^{\frac{1}{2}} \right]^2 \right\}^{\frac{5}{2}} (z^2 + b_{disk}^2)^{\frac{3}{2}}} \end{aligned}$$

with  $R$  the cylindrical radius and with  $z$  the distance to the reference plane.

For the **bulge** the Hernquist potential [44] is used

$$\begin{aligned} \Phi_{bulge}(r) &= -\frac{G * M_{bulge}}{(r + a_{bulge})} \\ \rho_{bulge}(r) &= \frac{M_{bulge}}{2\pi r} \frac{a_{bulge}}{(r + a_{bulge})^3} \end{aligned}$$

$a$  is the scale-length of the spheroid potential

and NFW potential [26] is used for the **dark matter halo** potential

$$\Phi_{halo}(r) = \frac{-4\pi G \rho_s r_s^3 \ln\left(1 + \frac{r}{r_s}\right)}{r}$$

where  $\rho_s$  is the characteristic density and  $r_s$  the scale length.

Parameters for bulge and disk taken from [30] and the radius  $r_s$  for the halo from [7]:

Table 4.1: MW model parameters

Parameter	Value	Unit
$M_{bh}$	$4 * 10^6$	$M_\odot$
$M_{disk}$	$10^{11}$	$M_\odot$
$a_{disk}$	6.5	$kpc$
$b_{disk}$	0.26	$kpc$
$M_{bulge}$	$3.4 * 10^{10}$	$M_\odot$
$a_{bulge}$	0.70	$kpc$
$r_s$	16	$kpc$

$\rho_s$  can be determined by imposing

$$v_c(R_0, z = 0) = 220 \left[ \frac{km}{s} \right]$$

$$R_0 = 8 [kpc]$$

## 4.1 Circular Velocity

The circular velocity  $v_c$  is defined via

$$v_c(R, z) = \sqrt{R \frac{\partial \Phi(R, z)}{\partial R}} \quad (4.1)$$

with in the presented case total potential being

$$\Phi(R, z) = \Phi_{bh}(R, z) + \Phi_{disk}(R, z) + \Phi_{bulge}(R, z) + \Phi_{halo}(R, z) \quad (4.2)$$

therefore

$$v_c^2 = v_{c,bh}^2 + v_{c,disk}^2 + v_{c,bulge}^2 + v_{c,halo}^2$$

$$v_{c,bh}^2 = \frac{GM_{bh}R^2}{(R^2 + z^2)^{3/2}}$$

$$v_{c,disk}^2 = \frac{GM_{disk}R^2}{\left( (a_{disk} + \sqrt{b_{disk}^2 + z^2})^2 + R^2 \right)^{3/2}}$$

$$v_{c,bulge}^2 = \frac{GM_{bulge}R^2}{\sqrt{R^2 + z^2} (a_{bulge} + \sqrt{R^2 + z^2})^2}$$

$$v_{c,halo}^2 = \frac{4\pi G \rho_s R^2 r_s^3 \log \left( \frac{\sqrt{R^2 + z^2}}{r_s} + 1 \right)}{(R^2 + z^2)^{3/2}} - \frac{4\pi G \rho_s R^2 r_s^2}{(R^2 + z^2) \left( \frac{\sqrt{R^2 + z^2}}{r_s} + 1 \right)}$$

Inserting the parameters given in Table 4.1 and setting  $G \approx 4.302 * 10^{-6} \left[ \frac{kpc \ km^2}{M_\odot s^2} \right]$  results in  $\rho_s \approx 4.5 * 10^6 \left[ \frac{M_\odot}{kpc^3} \right]$

The circular velocity can be used for fitting the model parameters [1]. Fig. 4.1 displays the velocity for the chosen model and its components.

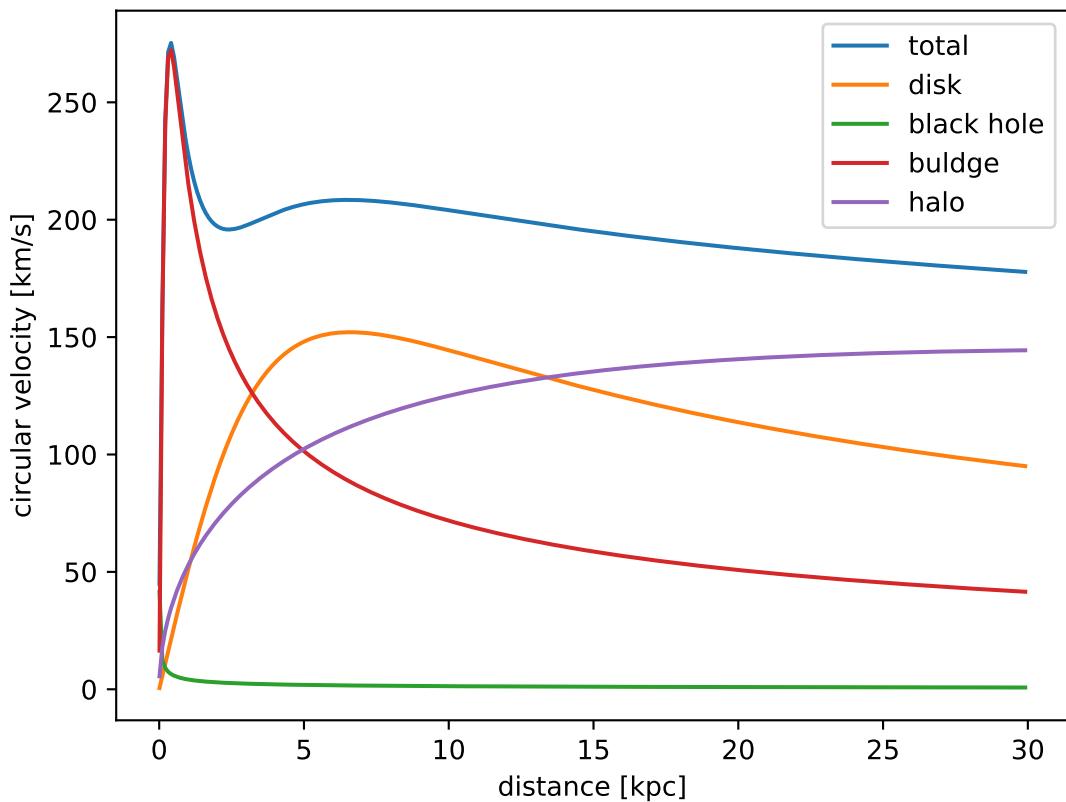


Fig. 4.1: Galactic rotation curve for the chosen MW model

## 4.2 Angular Velocity

The angular velocity is required for initializing disc star velocities and can be obtained from the potential:

$$\Omega^2(R) = \frac{1}{R} \frac{\partial \Phi(R, 0)}{\partial r}$$

For the total potential (4.2) this derivation yields:

$$\Omega^2(R) = \frac{G}{R} \left\{ -\frac{M_{bulge}}{(a_{bulge} + R)^2} + \frac{2M_{disk}R^3}{[(a_{disk} + b_{disk})^2 + R^4]^{1.5}} + \frac{M_{bh}}{R^2} - \frac{4\pi p_s r_s^3}{R^2 + Rr_s} + \frac{4\pi p_s r_s^3 \ln\left(\frac{R+r_s}{r_s}\right)}{R^2} \right\}$$

## 4.3 Mass Distribution

The mass inside a volume is calculated by numerical integration of the density. GSL implementation of Monte Carlo Integration is used. For further details refer to the [GSL documentation](#).

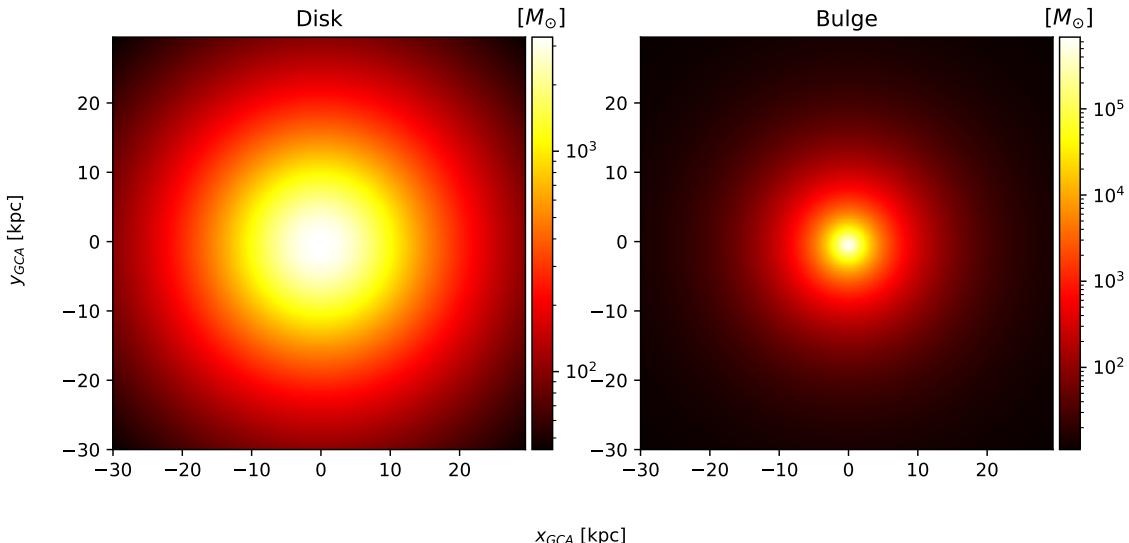


Fig. 4.2: Mass distribution at  $z_{GCA} = 1.0$  [kpc]

## 4.4 Surface Mass Density

The surface mass density (SMD) is used to estimate the vertical velocity dispersion (5.26) and is defined by

$$\Sigma(R) = 2 \int_0^\infty \rho(R, z) dz \quad (4.3)$$

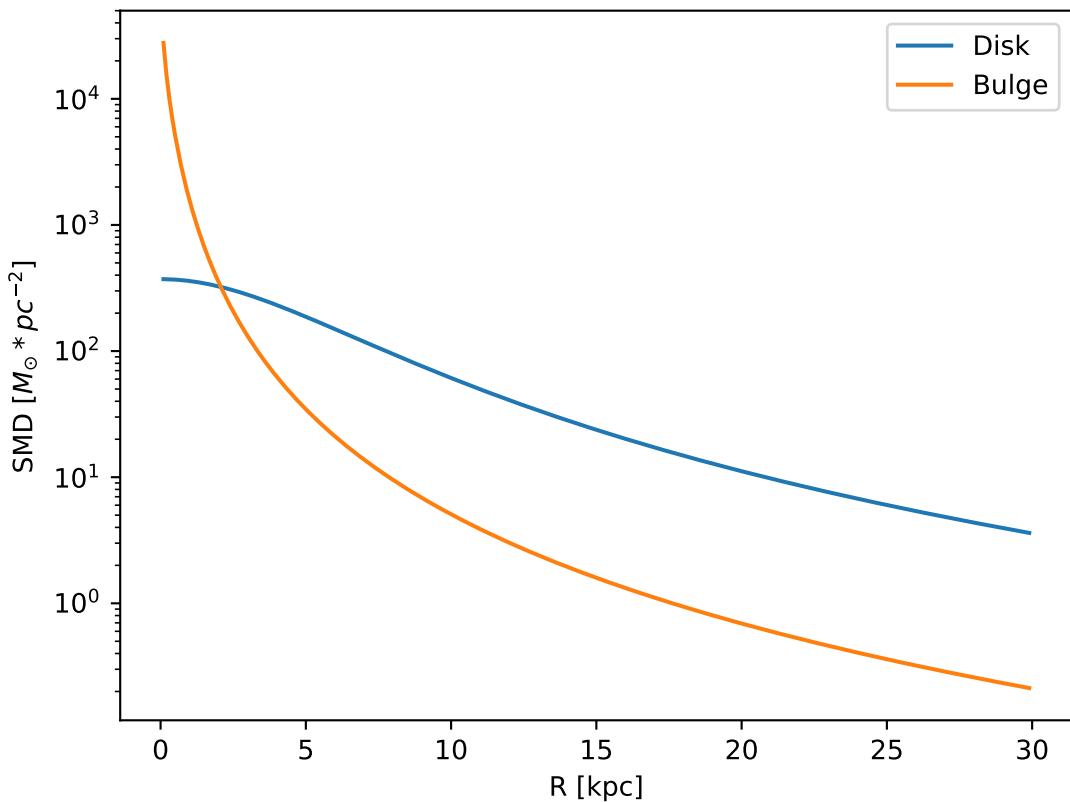


Fig. 4.3: The surface mass density of the disc and bulge



## INITIALIZATION

### 5.1 Sampling distribution functions

Space, velocity and mass distributions are sampled during various steps of the initialization using the following two methods. For further details and alternative methods see [11].

#### 5.1.1 Rejection sampling

A proposed distribution function  $g(x)$  is sampled and each resulting sample accepted with a probability of  $\frac{f(x)}{Cg(x)}$ .  $C$  is a constant ensuring  $f(x) \leq Cg(x) \forall x$  and  $f(x)$  the target distribution.

If  $g(x)$  is the pdf of a uniform distribution with  $x \in [a, b]$

$$g(x) = \frac{1}{b-a}$$

and  $m$  the maximum of  $f(x)$  within  $[a, b]$  then  $C = \frac{m}{b-a}$  and the acceptance criterion becomes

$$u \leq f(x)$$

here  $u$  is a uniform distribution  $U(0, Cg(x)) = U(0, m)$ . While  $Cg(x)$  has to be larger than  $f(x)$  the area between them should be as small as possible to ensure a decent acceptance fraction.

#### 5.1.2 Inverse transform sampling

If the cumulative probability function (cpf) of the probability distribution  $f(y)$  is continuous and its inverse obtainable, this method can be utilized and is more efficient than rejection sampling. The cpf describes the probability that the random variable is smaller or equal to  $x$

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(y)dy$$

The inverse  $F^{-1}(y)$  may be calculated via  $F(F^{-1}(y)) = y$ . With  $y$  drawn from a uniform distribution the results of  $F^{-1}(y)$  have the desired probability distribution  $f(y)$ .

## 5.2 Cluster

A cluster can be characterised via a number of parameters, namely its density profile, metallicity, binary fraction and initial mass function (IMF). The following sections cover methods implemented by the author. Alternatively cluster stars can be generated using the open source code McLuster [47] which supports more density profiles and all other parameters mentioned above. The generated file can then be imported by setting the appropriate parameters.

## 5.3 Initial Mass Function (IMF)

The IMF describes the initial mass distribution of stars, that is, how many stars have which mass in some set of stars at the start. It is subject of debate whether the IMF of star clusters is universal [2] and the present day mass function (PDMF) usually differs from the IMF. Figure 13 in [42] depicts the Salpeter IMF [55] and the PDMF of Arches. Hence, the IMF is deemed an adequate estimate for the PDMF of a generic YMSC. Sampling routines for the following two IMFs have been implemented by the author.

### 5.3.1 Salpeter IMF

$$p(m) = \frac{dN}{dm} = A * m^{-\alpha} \quad (5.1)$$

This comparatively simple and widely used ([32], [54], [43]) IMF can be sampled using inverse transform sampling. To do so one must integrate the IMF over the mass and normalize it, yielding the normalized cumulative distribution function (NCDF).

$$P(m) = \int_{m_{min}}^m A * m^{-\alpha} dm = \frac{A}{-\alpha + 1} (m^{-\alpha+1} - m_{min}^{-\alpha+1}) \quad (5.2)$$

$A$  is defined by normalization:

$$P(m_{max}) \equiv 1 \rightarrow A = \frac{-\alpha + 1}{m_{max}^{-\alpha+1} - m_{min}^{-\alpha+1}}$$

With this (5.2) becomes:

$$P(m) = \frac{m^{-\alpha+1} - m_{min}^{-\alpha+1}}{m_{max}^{-\alpha+1} - m_{min}^{-\alpha+1}}$$

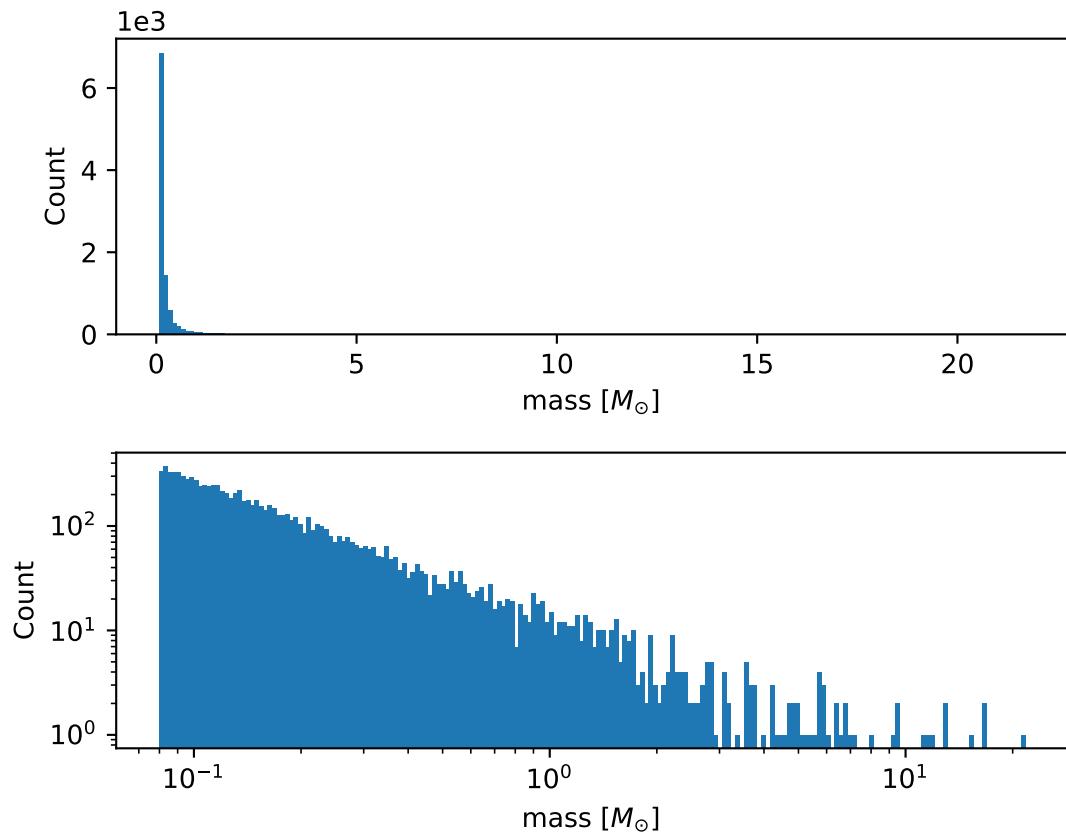
Since Salpeter is a power-law distribution function, the inverse of the NCPF can calculated analytically. After some simplifications the final result reads

$$m_{rand} = m_{min} * \left[ 1 + x * \left( \left( \frac{m_{max}}{m_{min}} - 1 \right)^{-\alpha+1} \right) \right]^{\frac{1}{-\alpha+1}}$$

where  $x$  is a random number in range [0,1]

### 5.3.2 Broken Power Law (BPL)

The BPL IMF has the benefit of being highly adaptable. It can be utilized to approximate any other IMF. It has been used to describe both globular cluster IMFs [3] as well as young star cluster IMFs ([28], [19]).

Fig. 5.1:  $10^3$  stars sampled from the Salpeter IMF

The following is a generalization of the equation given by [18] for  $n - 1$  intervals.

$$\xi(m) = A \begin{cases} k_1 m^{-\alpha_1} & \text{if } m_1 \leq m < m_2 \\ k_2 m^{-\alpha_2} & \text{if } m_2 \leq m < m_3 \\ \dots & \\ k_{n-1} m^{-\alpha_{n-1}} & \text{if } m_{n-1} \leq m < m_n \end{cases}$$

where  $A$  is a normalization constant and  $k_i$  is defined as

$$\begin{aligned} k_1 &= m_2^{\alpha_1} \\ k_2 &= m_2^{\alpha_2} \\ k_i &= k_{i-1} m_i^{\alpha_i - \alpha_{i-1}} \end{aligned}$$

As in the case of Salpeter, random samples are drawn with inverse transform sampling.

The NCDF can be obtained by integrating  $\xi(m)$  over the mass interval.

$$F(m) = \int_{m_1}^m \xi(m) dm$$

Where  $A$  is defined by the normalization constraint:

$$A * \sum_{i=1}^{n-1} \left( k_i \int_{m_i}^{m_{i+1}} m^{-\alpha_i} dm \right) = 1$$

Inverting the NCDF leads to

$$F^{-1}(y) = \begin{cases} \left[ \frac{(1-\alpha_1)y}{A*k_1} + m_1^{1-\alpha_1} \right]^{\frac{1}{1-\alpha_1}} & \text{if } 0 \leq y < \frac{A*k_1}{1-\alpha_1} (m_2^{1-\alpha_1} - m_1^{1-\alpha_1}) \\ \left\{ \left[ y - \frac{A*k_1}{1-\alpha_1} (m_2^{1-\alpha_1} - m_1^{1-\alpha_1}) \right] \frac{1-\alpha_2}{A*k_2} + m_2^{1-\alpha_2} \right\}^{\frac{1}{1-\alpha_2}} & \text{if } \frac{A*k_1}{1-\alpha_1} (m_2^{1-\alpha_1} - m_1^{1-\alpha_1}) \leq y < \sum_{i=1}^2 \frac{A*k_i}{1-\alpha_i} (m_{i+1}^{1-\alpha_i} - m_i^{1-\alpha_i}) \\ \dots & \\ \left\{ \left[ \sum_{i=1}^{n-2} y - \frac{A*k_i}{1-\alpha_i} (m_{i+1}^{1-\alpha_i} - m_i^{1-\alpha_i}) \right] \frac{1-\alpha_{n-1}}{A*k_{n-1}} + m_{n-1}^{1-\alpha_{n-1}} \right\}^{\frac{1}{1-\alpha_{n-1}}} & \text{if } \sum_{i=1}^{n-2} \frac{A*k_i}{1-\alpha_i} (m_{i+1}^{1-\alpha_i} - m_i^{1-\alpha_i}) \leq y < \sum_i^n \frac{A*k_i}{1-\alpha_i} (m_{i+1}^{1-\alpha_i} - m_i^{1-\alpha_i}) \end{cases}$$

where  $y$  is a random number in range  $[0,1]$

## 5.4 Density Profile

Different density profiles (King [46], Plummer [53] and Elson, Fall & Freeman [41]) have been used ([17], [35], [16]) to simulate YMCs. The Plummer profile has been implemented here as described in [38].

## 5.5 Field Stars (FS)

Any stars which does not belong to the simulated cluster is a FS and sampled from the disc and bulge potential described in *Galactic Potential*.

The cone of vision (COV) is defined by the angle of view  $\alpha$ , the view distance  $h$  (height of the cone), the view point  $vP$  (location of the observer) and the focus  $F$  (a point along the line of sight).

The COV is constructed by transforming a right circular cone, where the vertex is at the origin and the circular base normal to the z axis.

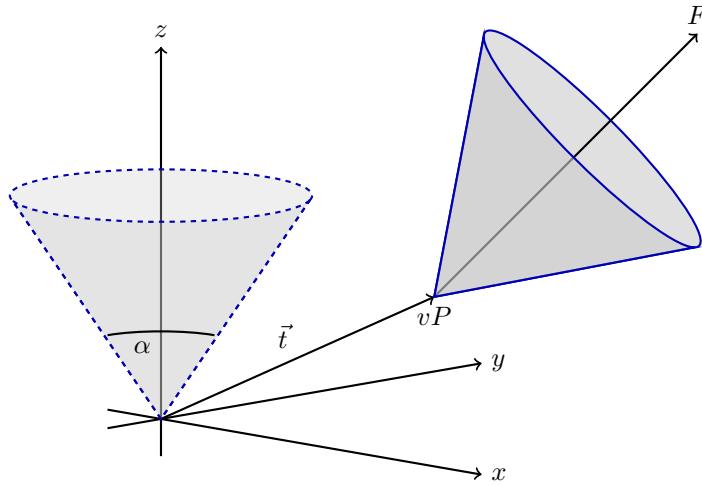


Fig. 5.2: Transformation of the cone of vision

Per transformation the tip of the cone is displaced from the origin to the view point  $vP$  and its axis is rotated to align with the line of sight  $l$ . Consequently, the transformation consists of both translation and rotation illustrated in the following figure.

Rotation and translation are both isometric transformations meaning area and volume of the cone are preserved [15] (p.175).

A unit quaternion  $\mathbf{q}$  is used in order to construct the rotation matrix. With rotation axis  $\vec{b}$  and angle  $\beta$  the quaternion is given by

$$\mathbf{q} = \left( \cos\left(\frac{\beta}{2}\right), \vec{b} \sin\left(\frac{\beta}{2}\right) \right)$$

The rotation axis  $\vec{b}$  is the normalized cross product of the original ( $\vec{z}$ ) and target ( $l$ ) cone axis

$$\vec{b} = \frac{\vec{z} \times \vec{l}}{\|\vec{z} \times \vec{l}\|}$$

The angle  $\beta$  between the vectors of interest can be calculated as follows

$$\beta = \text{atan2}(\tan(\beta)) = \text{atan2}\left(\frac{\sin(\beta)}{\cos(\beta)}\right) = \text{atan2}\left(\frac{\|\vec{z} \times \vec{l}\|}{\vec{z} \cdot \vec{l}}\right)$$

Next, quaternion is converted to the rotation matrix [21]. Using the homogeneous notation [34] (p. 57) the matrix becomes:

$$\mathbf{R} = \begin{bmatrix} q_1^2 + q_2^2 - q_3^2 - q_4^2 & -2q_1q_4 + 2q_2q_3 & 2q_1q_3 + 2q_2q_4 & 0 \\ 2q_1q_4 + 2q_2q_3 & q_1^2 - q_2^2 + q_3^2 - q_4^2 & -2q_1q_2 + 2q_3q_4 & 0 \\ -2q_1q_3 + 2q_2q_4 & 2q_1q_2 + 2q_3q_4 & q_1^2 - q_2^2 - q_3^2 + q_4^2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The translation matrix for the translation vector  $\vec{t}$  reads [34] (p. 66):

$$\mathbf{T}_{\text{translation}} = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The transformation matrix  $\mathbf{T}$  is the product of  $\mathbf{R}$  and  $\mathbf{T}_{\text{translation}}$

$$\mathbf{T} = \begin{bmatrix} q_1^2 + q_2^2 - q_3^2 - q_4^2 & -2q_1q_4 + 2q_2q_3 & 2q_1q_3 + 2q_2q_4 & t_x \\ 2q_1q_4 + 2q_2q_3 & q_1^2 - q_2^2 + q_3^2 - q_4^2 & -2q_1q_2 + 2q_3q_4 & t_y \\ -2q_1q_3 + 2q_2q_4 & 2q_1q_2 + 2q_3q_4 & q_1^2 - q_2^2 - q_3^2 + q_4^2 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5.3)$$

The total mass  $M$  of disc and bulge stars is obtained by integrating the respective density over the COV.

For a right circular cone

$$\tan\left(\frac{\alpha}{2}\right) = \frac{R}{h}$$

and on its surface

$$\begin{aligned} \tan\left(\frac{\alpha}{2}\right) &= \frac{r}{z} \\ z &= \frac{h}{R}r \end{aligned}$$

where  $R$  is the base radius of the cone and  $r = \sqrt{x^2 + y^2}$ . Hence  $\frac{h}{R}r \leq z \leq h$ . Both  $x$  and  $y$  are bound by the base radius. Choosing  $|x| \leq R$  implies  $|y| \leq \sqrt{R^2 - x^2}$ .

$$M = \int_{-R}^R \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \int_{\frac{h}{R}r}^h \rho \left( \mathbf{T} * \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right) dz dy dx$$

### 5.5.1 Bulge Stellar Mass Function

By rejection sampling the following function, given by [40], the mass of bulge stars is determined.

For  $m < 0.7$  the log-normal distribution equation (5.4) is used. Parameters are  $A = 3.6 * 10^{-4}$ ,  $m_c = 0.22$  and  $\sigma = 0.33$ . For  $m > 0.7$  a Salpeter slope (5.1) with parameters  $A = 7.1 * 10^{-5}$  and  $x = 1.3$  is chosen.

Samples are drawn until the sum of all samples  $M_s = \sum_{i=1}^n m_i$  is larger than the total mass  $M$ . If  $\sum_{i=1}^{n-1} m_i + \frac{m_n}{2} > M$  the last sample is removed.

### 5.5.2 Disk Stellar Mass Function

Stars belonging to the disk are given a mass by rejection sampling the PDMF as given by [40].

For  $m < 1$  the PDMF reads

$$\xi(\log(m)) = \frac{dN}{d\log(m)} = A * \exp\left[\frac{-(\log(m) - \log(m_c))^2}{2\sigma^2}\right] \quad (5.4)$$

or equivalently (this version is sampled)

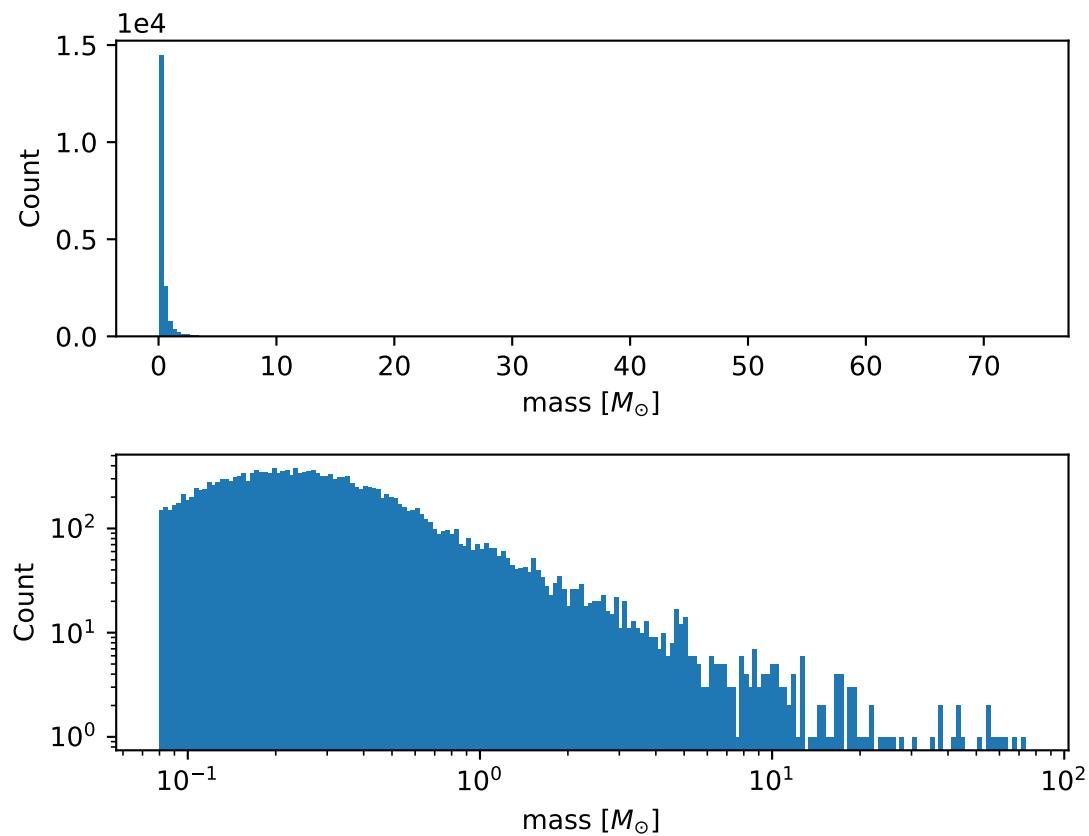
$$\xi(m) = \frac{dN}{dm} = \frac{A}{m \ln(10)} * \exp\left[\frac{-(\log(m) - \log(m_c))^2}{2\sigma^2}\right]$$

For  $m > 1$  the PDMF has the form

$$\xi(\log(m)) = \frac{dN}{d\log(m)} = Am^{-x}$$

or depending on  $m$  rather than  $\log(m)$

$$\xi(m) = \frac{dN}{dm} = \frac{A}{m \ln(10)} m^{-x}$$

Fig. 5.3:  $10^3$  stars sampled from the bulge MF

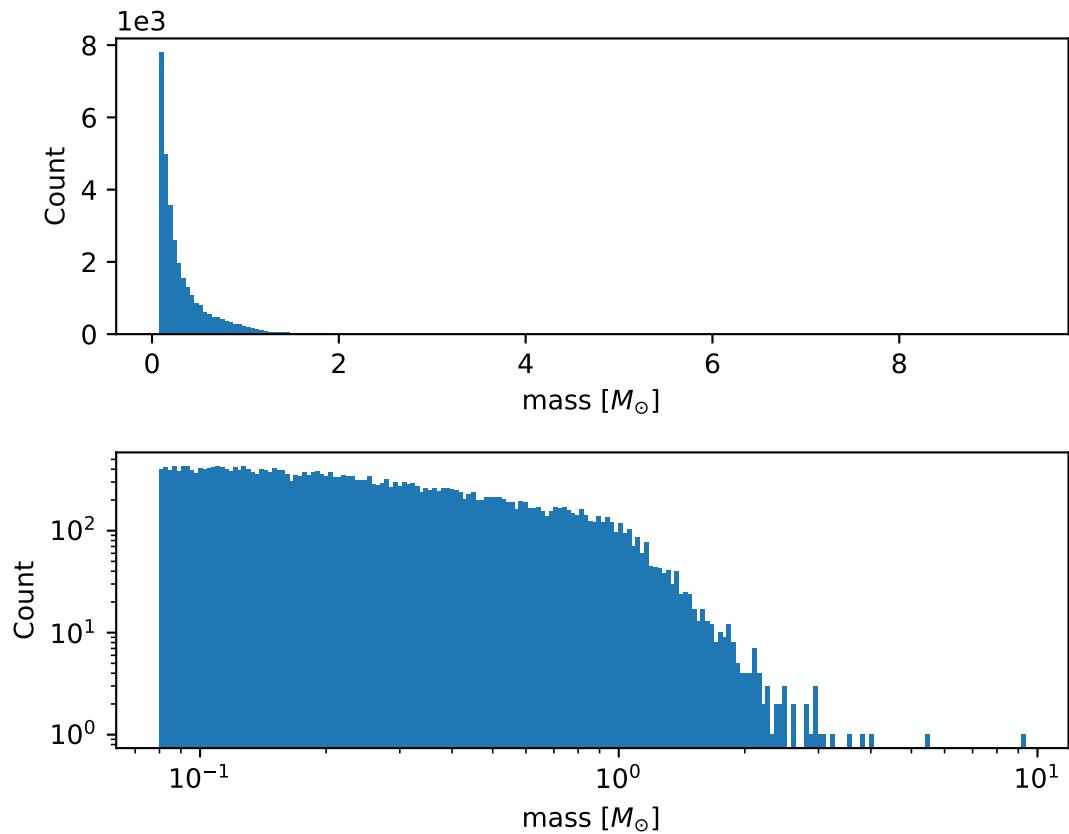


Fig. 5.4: Stars with total mass of  $10^3 M_{\odot}$  sampled from disk PDMF

## 5.6 Positions

The positions of the field stars within the cone of vision are generated in two steps of rejection sampling followed by the transformation (5.3).

In the first step trial positions are drawn from a uniform distribution within a cuboid containing the cone. The boundaries of the cuboid are given by

$$\begin{aligned} |x| &\leq R \\ |y| &\leq R \\ 0 &\leq z \leq h \end{aligned}$$

where  $R = h * \tan\left(\frac{\alpha}{2}\right)$  is the base radius of the cone.

These trial positions are rejected in case they lie outside the boundaries of the cone. The conditions for acceptance are:

$$\begin{aligned} \sqrt{x^2 + y^2} &\leq R \\ z &\geq h * \frac{\sqrt{x^2 + y^2}}{R} \end{aligned}$$

This method ensures that the positions are indeed homogeneously distributed which is essential for the second step.

The second step consists of rejection sampling the density distribution. The test variable is drawn from a uniform distribution ranging from the smallest to the largest possible density within the cone volume. If this test variable is smaller than the density at the trial position generated in step two, the trial position is accepted and rejected otherwise.

Finally the accepted position is transformed via the transformation matrix (5.3).

## 5.7 Velocities

### 5.7.1 Particle Kinematics in Cylindrical Coordinates

In terms of cartesian unit vectors, cylindrical unit vectors are given by

$$\begin{aligned} \hat{e}_R &= \hat{e}_x \cos(\phi) + \hat{e}_y \sin(\phi) \\ \hat{e}_\phi &= -\hat{e}_x \sin(\phi) + \hat{e}_y \cos(\phi) \\ \hat{e}_z &= \hat{e}_z \end{aligned}$$

Since  $\hat{e}_R$  depends on  $\phi$ , position vectors have the form

$$\vec{q} = R\hat{e}_R + z\hat{e}_z$$

derivation by time gives the velocity

$$\vec{v} = \dot{R}\hat{e}_R + R\dot{\phi}\hat{e}_\phi + \dot{z}\hat{e}_z \quad (5.5)$$

#### Lagrangian with axisymmetric potential

Using (5.5) the Lagrangian per unit mass reads

$$L = \frac{1}{2} \left( \dot{R}^2 + R^2 \dot{\phi}^2 + \dot{z}^2 \right) + \Phi(R, z)$$

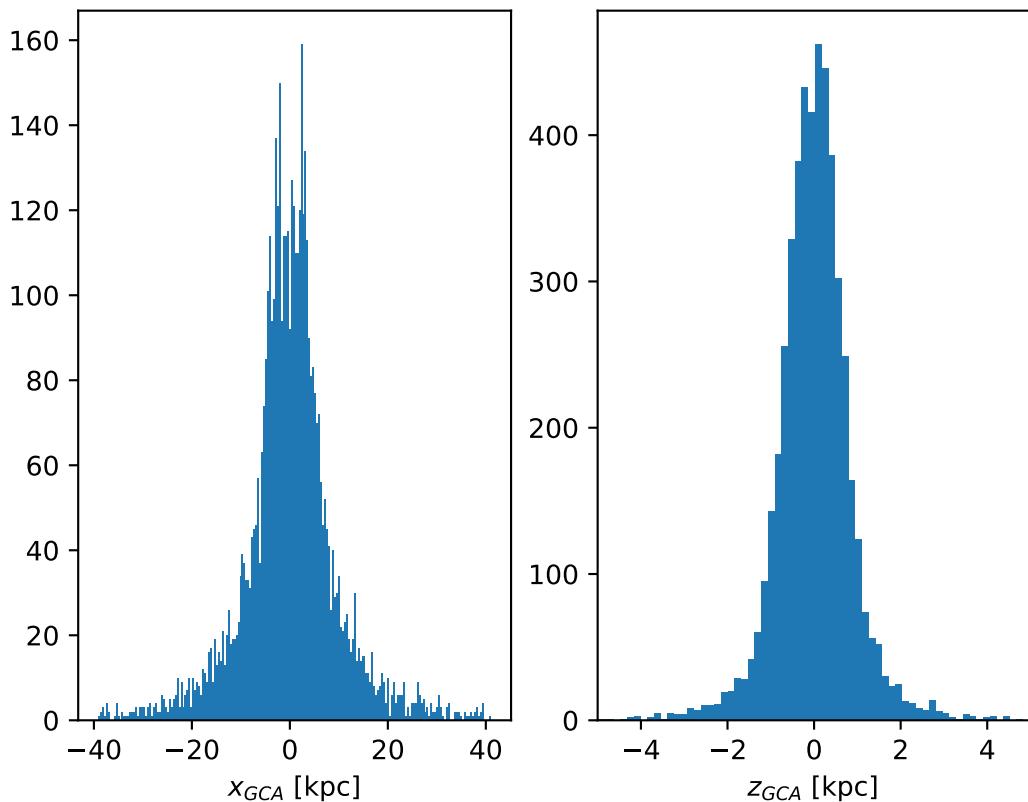


Fig. 5.5: Spatial distribution of a sample of disc and bulge stars

and using the Euler–Lagrange equation gives the conjugate momenta

$$\begin{aligned} p_R &= \frac{\partial L}{\partial \dot{R}} = \dot{R} \\ p_\phi &= \frac{\partial L}{\partial \dot{\phi}} = R^2 \dot{\phi} \\ p_z &= \frac{\partial L}{\partial \dot{z}} = \dot{z} \end{aligned} \quad (5.6)$$

### Hamiltonian with axisymmetric potential

Using the momenta in cylindrical coordinates (5.6) the Hamiltonian per unit mass with an axisymmetric potential reads ([6] p. 278)

$$H = \frac{1}{2} \left( p_R^2 + \frac{p_\phi^2}{R^2} + p_z^2 \right) + \Phi(R, z) \quad (5.7)$$

using Hamiltons equations gives

$$\dot{p}_R = -\frac{\partial H}{\partial R} = \frac{p_\phi^2}{R^3} - \frac{\partial \Phi}{\partial R} \quad (5.8)$$

$$\dot{p}_\phi = -\frac{\partial H}{\partial \phi} = -\frac{\partial \Phi}{\partial \phi} = 0 \quad (5.9)$$

$$\dot{p}_z = -\frac{\partial H}{\partial z} = -\frac{\partial \Phi}{\partial z} \quad (5.10)$$

Since  $\vec{L} = \vec{p} \times \vec{p}$  and thus

$$L_z = R^2 \dot{\phi} \quad (5.11)$$

, the second equation above signifies that, in the case of an axisymmetric potential, the z component of the angular momentum is conserved.

### 5.7.2 The phase space distribution function (DF)

A galaxy can be viewed as a number of stars in phase space. The phase space is a six dimensional space spanned by position  $\vec{q}$  and momentum  $\vec{p}$ . One point in phase space can therefor be described by a six dimensional vector  $\vec{w} = (\vec{q}, \vec{p})$ , the state of a whole galaxy as a DF  $f(\vec{q}, \vec{p}, t)$ .  $f(\vec{q}, \vec{p}, t) d^3 \vec{q} d^3 \vec{p}$  is the probability that some specific star is inside the 6D cuboid defined by  $d^3 \vec{q} d^3 \vec{p}$  at time  $t$ .

The conservation of probability in phase space is, similarly to the continuity equation in fluid dynamics, given by

$$\frac{\partial f}{\partial t} + \sum_{i=1}^6 \frac{\partial}{\partial w_i} (f \dot{w}_i) = 0$$

with  $\dot{\vec{w}} = (\dot{\vec{q}}, \dot{\vec{p}})$ .

The right summand can be changed as follows, using Hamilton equations

$$\begin{aligned}
 \sum_{i=1}^6 \frac{\partial}{\partial w_i} (f \dot{w}_i) &= \sum_{i=1}^3 \left( \frac{\partial}{\partial q_i} (f \dot{q}_i) + \frac{\partial}{\partial p_i} (f \dot{p}_i) \right) \\
 &= \sum_{i=1}^3 \left( \frac{\partial}{\partial q_i} \left( f \frac{\partial H}{\partial p_i} \right) - \frac{\partial}{\partial p_i} \left( f \frac{\partial H}{\partial q_i} \right) \right) \\
 &= \sum_{i=1}^3 \left( \frac{\partial f}{\partial q_i} \frac{\partial H}{\partial p_i} + f \frac{\partial^2 H}{\partial q_i \partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial H}{\partial q_i} - f \frac{\partial^2 H}{\partial p_i \partial q_i} \right) \\
 &= \sum_{i=1}^3 \left( \frac{\partial f}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial H}{\partial q_i} \right)
 \end{aligned}$$

And therefor

$$\frac{\partial f}{\partial t} + \sum_{i=1}^3 \left( \frac{\partial f}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial H}{\partial q_i} \right) = 0 \quad (5.12)$$

which is the collisionless Boltzmann Equation (CBE).

The zeroth moment of the DF is the number density:

$$\nu(\vec{x}) \equiv \int f(\vec{x}, \vec{v}) d^3 \vec{v}$$

Mean velocities are given by the first moment:

$$\bar{v}_i(\vec{x}) \equiv \frac{1}{\nu(\vec{x})} \int v_i f(\vec{x}, \vec{v}) d^3 \vec{v}$$

Todo(?): Continue here (showing = 0 for 2/3 directions)

### 5.7.3 Jeans equations

Using Einstein notation for  $i = 1, 2, 3$  the CBE (5.12) is given by ([6] p. 277):

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial H}{\partial q_i} = 0$$

Since the galactic potential (4.2) is axisymmetric, it is convenient to express this equation in cylindrical coordinates.

$$\frac{\partial f}{\partial t} + p_R \frac{\partial f}{\partial R} + \frac{p_\phi}{R^2} \frac{\partial f}{\partial \phi} + p_z \frac{\partial f}{\partial z} - \left( \frac{\partial \Phi}{\partial R} - \frac{p_\phi^2}{R^3} \right) \frac{\partial f}{\partial p_R} - \frac{\partial \Phi}{\partial \phi} \frac{\partial f}{\partial p_\phi} - \frac{\partial \Phi}{\partial z} \frac{\partial f}{\partial p_z} = 0 \quad (5.13)$$

It is assumed that the galaxy is statistically in a steady state ([4]) i.e.  $\frac{\partial f}{\partial t} = 0$ . Due to this assumption and taking (5.9) into account (5.13) simplifies to

$$p_R \frac{\partial f}{\partial R} + \frac{p_\phi}{R^2} \frac{\partial f}{\partial \phi} + p_z \frac{\partial f}{\partial z} - \left( \frac{\partial \Phi}{\partial R} - \frac{p_\phi^2}{R^3} \right) \frac{\partial f}{\partial p_R} - \frac{\partial \Phi}{\partial z} \frac{\partial f}{\partial p_z} = 0 \quad (5.14)$$

Multiplying equation (5.14) by  $p_R$  and integrating over all momenta leads to (todo: derive?)

$$\frac{\partial \nu \bar{v}_R^2}{\partial R} + \frac{\partial \nu \bar{v}_R \bar{v}_z}{\partial z} + \nu \left( \frac{\bar{v}_R^2 - \bar{v}_\phi^2}{R} + \frac{\partial \Phi}{\partial R} \right) = 0 \quad (5.15)$$

### 5.7.4 The Epicyclic Approximation

Individual stars in the disk are on nearly circular orbits. Such orbits can be approximated by circular orbits with additional retrograde elliptical orbits around the guiding center.

The derivation of this approximation starts with Hamilton's equations for an axisymmetric potential.

Rearranging (5.7) and using (5.11) gives

$$H = \frac{1}{2} (p_R^2 + p_z^2) + \frac{L_z^2}{2R^2} + \Phi(R, z)$$

With the effective potential given by

$$\Phi_{\text{eff}}(R, z) = \frac{L_z^2}{2R^2} + \Phi(R, z)$$

leads to

$$H_{\text{eff}} = \frac{1}{2} (p_R^2 + p_z^2) + \Phi_{\text{eff}}(R, z)$$

Here  $\frac{1}{2m} (p_R^2 + p_z^2)$  is the kinetic energy in the  $(R, z)$  plane or meridional plane. The angular momentum term in the effective potential is not a real potential energy even though sometimes called centrifugal potential. It really is the angular kinetic energy. The given definition of  $\Phi_{\text{eff}}$  is only valid because  $L_z$  is conserved.

Using  $\Phi_{\text{eff}}$  (5.8) and (5.10) can be written as

$$\dot{p}_R = -\frac{\partial \Phi_{\text{eff}}}{\partial R} \quad (5.16)$$

$$\dot{p}_z = -\frac{\partial \Phi_{\text{eff}}}{\partial z} \quad (5.17)$$

These equations describe harmonic oscillations in the effective potential. The minimum of effective potential is the minimum of the real potential energy, together with a contribution from the angular kinetic energy.

$$\begin{aligned} \frac{\partial \Phi_{\text{eff}}}{\partial R} &= \frac{\partial \Phi}{\partial R} - \frac{L_z^2}{R^3} = 0 \\ \frac{\partial \Phi_{\text{eff}}}{\partial z} &= 0 \end{aligned}$$

The first condition states that the attractive force ( $-\frac{\partial \Phi_{\text{eff}}}{\partial R}$ ) has to balance the centrifugal force. This is the condition for circular orbits with angular momentum  $L_z$ . The second condition is clearly satisfied in the equatorial plane ( $z = 0$ ). The coordinates of this guiding center are defined as  $(R_g, \phi_g, z_g)$ .

In preparation for a Taylor series expansion about the guiding center,  $x$  is defined as

$$x \equiv R - R_g$$

If  $R = R_g$  then  $x = 0$  and therefore the guiding center is at  $(x, z) = (0, 0)$ .

$$\Phi_{\text{eff}} = \Phi_{\text{eff}}(R_g, 0) + \left. \frac{\partial \Phi_{\text{eff}}}{\partial R} \right|_{(R_g, 0)} x + \left. \frac{\partial \Phi_{\text{eff}}}{\partial z} \right|_{(R_g, 0)} z + \frac{1}{2} \left. \frac{\partial^2 \Phi_{\text{eff}}}{\partial R^2} \right|_{(R_g, 0)} x^2 + \frac{1}{2} \left. \frac{\partial^2 \Phi_{\text{eff}}}{\partial z^2} \right|_{(R_g, 0)} z^2 + \frac{1}{2} \left. \frac{\partial^2 \Phi_{\text{eff}}}{\partial x \partial z} \right|_{(R_g, 0)} xz + \mathcal{O}(xz^2)$$

The first order terms are zero (since  $\Phi_{\text{eff}}$  is minimized at the guiding center) and so is the  $xz$ , the latter due to symmetry about  $z = 0$ .

In the epicyclic approximation  $\mathcal{O}(xz^2)$  and higher order terms are neglected.

With this approximation (5.16) and (5.17) become

$$\begin{aligned}\dot{p}_x &= -\frac{\partial \Phi_{\text{eff}}}{\partial x} \approx -\frac{\partial^2 \Phi_{\text{eff}}}{\partial R^2} \Big|_{(R_g, 0)} x \equiv -\kappa^2 x \\ \dot{p}_z &= -\frac{\partial \Phi_{\text{eff}}}{\partial z} \approx -\frac{\partial^2 \Phi_{\text{eff}}}{\partial z^2} \Big|_{(R_g, 0)} z \equiv -\nu^2 z\end{aligned}\quad (5.18)$$

where the epicyclic frequency  $\kappa$  is the frequency of small radial and the vertical frequency  $\nu$  the frequency of small vertical oscillations.

With potential  $\Phi(R, z)$  (5.18) can be written as

$$\kappa^2 = \frac{\partial^2 \Phi}{\partial R^2} \Big|_{(R_g, 0)} + \frac{3L_z^3}{R_g^4} \quad (5.19)$$

The combination of (5.6) with (5.8), given  $\dot{p}_R = 0$  and by definition  $\Omega = \dot{\phi}$ , results in the circular angular frequency

$$\Omega^2 = \frac{1}{R} \frac{\partial \Phi}{\partial R} \Big|_{(R_g, 0)} = \frac{L_z^2}{R^4} \quad (5.20)$$

The derivative of (5.20) is

$$\frac{\partial^2 \Phi}{\partial R^2} = \Omega^2 + R \frac{d\Omega^2}{dR} \quad (5.21)$$

Inserting (5.20) and (5.21) into (5.19) yields

$$\kappa^2(R_g) = \left( R \frac{d\Omega^2}{dR} + 4\Omega^2 \right) \Big|_{R=R_g} \quad (5.22)$$

The variance is generally given as the mean of squares minus the square of means. In the epicyclic approximation the mean velocity in  $R$  and  $z$  is zero and therefor

$$\begin{aligned}\sigma_{v_\phi}^2 &= \overline{v_\phi^2} - \overline{v_\phi}^2 \\ \sigma_{v_R}^2 &= \overline{v_R^2} \\ \sigma_{v_z}^2 &= \overline{v_z^2}\end{aligned}\quad (5.23)$$

## 5.7.5 Disk

The velocity distribution of stars in the milky way disk is approximated with the help of Jeans equations as well as relations and constraints based on observations.

For a flat rotation curve the radial velocity dispersion exponentially decreases with increasing radius [59] (p. 114)

$$\sigma_{v_R} \propto e^{-\frac{R}{h}} \quad (5.24)$$

where  $h$  in the case of the Miyamoto Nagai potential is the radial scale length  $a$ .

The relation (5.24) still requires a constant factor  $k$ , which can be determined by means of the Toomre parameter  $Q$  at some distance  $R_{ref}$ .

$Q$  is the ratio between the actual and minimum velocity dispersion  $\sigma_{v_{R,min}}$  [57] (p. 1234)

$$\begin{aligned}\sigma_{v_{R,min}} &= \frac{3.36G\Sigma}{\kappa} \\ Q \equiv \frac{\sigma_{v_R}}{\sigma_{v_{R,min}}} &= \frac{\kappa\sigma_{v_R}}{3.36G\Sigma}\end{aligned}$$

where  $\kappa$  denotes the epicyclic frequency (5.22).

In the solar neighborhood  $Q_* = 2.7 \pm 0.4$  and  $\sigma_{v_R} = (38 \pm 2) \left[ \frac{km}{s} \right]$  [6] (p. 497)

The constant  $k$  can therefor be approximated via

$$k \cong Q \sigma_{v_{R,min}} e^{\frac{R}{\kappa}} \quad (5.25)$$

Under the approximation of isothermal sheets (introduced in [59]), the vertical velocity dispersion only depends on the surface density [58]

$$\sigma_{v_z} = \pi G \Sigma(R) z_0 \quad (5.26)$$

with  $z_0$  being the vertical scale length  $b$  when using the Miyamoto Nagai potential.

Combining the equation for the circular velocity (4.1) and Jeans equation (5.15) and multiplication by  $\frac{R}{\nu}$  results in

$$\frac{R}{\nu} \frac{\partial \nu \overline{v_R^2}}{\partial R} + \frac{R}{\nu} \frac{\partial \nu \overline{v_R v_z}}{\partial z} + \overline{v_R^2} - \overline{v_\phi^2} + v_c^2 = 0 \quad (5.27)$$

For a razor thin disc the spatial density  $\nu$  can be replaced with the surface density  $\Sigma$  [45], which does not depend on  $z$ . With this and by using relations (5.23), (5.27) simplifies to

$$\frac{R}{\Sigma} \frac{\partial \nu \sigma_{v_R}^2}{\partial R} + R \frac{\partial \overline{v_R v_z}}{\partial z} + \sigma_{v_R}^2 - \sigma_{v_\phi}^2 - \overline{v_\phi^2} + v_c^2 = 0 \quad (5.28)$$

In the epicyclic approximation, azimuthal and radial dispersion have the following relation ([6] p. 170)

$$\frac{\sigma_{v_\phi}^2}{\sigma_{v_R}^2} = \frac{\kappa^2}{4\Omega^2} \quad (5.29)$$

With this relation (5.28) can be written as

$$\frac{R}{\Sigma} \frac{\partial (\Sigma \sigma_{v_R}^2)}{\partial R} + R \frac{\partial \overline{v_R v_z}}{\partial z} + \sigma_{v_R}^2 - \sigma_{v_R}^2 \frac{\kappa^2}{4\Omega^2} - \overline{v_\phi^2} + v_c^2 = 0$$

rearrangements lead to

$$\overline{v_\phi^2} - v_c^2 = \sigma_{v_R}^2 \left( 1 - \frac{\kappa^2}{4\Omega^2} + \frac{R}{\Sigma \sigma_{v_R}^2} \frac{\partial (\Sigma \sigma_{v_R}^2)}{\partial R} + \frac{R}{\sigma_{v_R}^2} \frac{\partial \overline{v_R v_z}}{\partial z} \right) \quad (5.30)$$

Due to the exponential dependency of the surface mass density ([6] p. 325) and the radial dispersion  $\sigma_{v_R}$  (5.24) on  $R$ , the derivation after  $R$  summand can be solved as follows

$$\frac{R}{\Sigma \sigma_{v_R}^2} \frac{\partial (\Sigma \sigma_{v_R}^2)}{\partial R} = \frac{1}{\Sigma \sigma_{v_R}^2} \frac{\partial R}{\partial \ln(R)} \frac{\partial (\Sigma \sigma_{v_R}^2)}{\partial R} = \frac{1}{\Sigma \sigma_{v_R}^2} \frac{\partial (\Sigma \sigma_{v_R}^2)}{\partial \ln(R)} = \frac{1}{\Sigma \sigma_{v_R}^2} \left( -\frac{R \Sigma \sigma_{v_R}^2}{a} - \frac{R \sigma_{v_R}^2 \Sigma}{a} \right) = -2 \frac{R}{a}$$

With this and under the assumption that the retrograde elliptical orbit is aligned with the coordinate axes (5.30) becomes

$$\overline{v_\phi^2} - v_c^2 = \sigma_{v_R}^2 \left( 1 - \frac{\kappa^2}{4\Omega^2} - \frac{R}{a} \right) \quad (5.31)$$

The velocity components can be sampled from Gaussian distributions. Dispersions are obtained by first evaluating the epicyclic frequency (5.22) and the surface mass density (4.3), followed by the radial velocity dispersion (5.24) with constant factor (5.25), the vertical (5.26) and the azimuthal velocity dispersion (5.29). As noted before, the mean velocity in  $R$  and  $z$  is zero. For  $\phi$  the mean velocity is calculated from (5.31), where the circular velocity  $v_c$  is given by (4.1).

## 5.7.6 Bulge

For spherically symmetric distribution functions, such as, with the present approximation, the bulge, the spherically symmetric Jeans equation can be used [37]

$$\frac{\partial(\rho\sigma_r^2)}{\partial r} + 2\frac{\beta\sigma_r^2}{r} + \rho\frac{\partial\Phi}{\partial r} = 0 \quad (5.32)$$

where  $\beta$  is an indicator for anisotropy in radial and tangential direction.

$$\beta = 1 - \frac{\sigma_\phi^2}{\sigma_r^2}$$

If the distribution is isotropic, as assumed here,  $\beta = 0$  and (5.32) simplifies to

$$\frac{\partial(\rho\sigma_r^2)}{\partial r} + \rho\frac{\partial\Phi}{\partial r} = 0$$

and the radial dispersion can be calculated by integration

$$\sigma_r^2 = \frac{1}{\rho} \int_r^\infty \rho \frac{\partial\Phi}{\partial r} dr$$

However,  $\Phi$  is the composite potential consisting of bulge, disc and halo and therefore not spherically symmetric. As suggested in [45], contribution by the disc to the bulge dispersion may be accounted for by spherically averaging the disc potential. This is achieved by expressing cylinder in spherical coordinates and integrating over the relevant angle

$$\sigma_{r,disk}^2 = \frac{1}{\rho} \int_r^\infty \frac{\rho}{2\pi} \int_0^{2\pi} \frac{\partial\Phi(r \sin(\theta), r \cos(\theta))}{\partial r} d\theta dr$$

As neither of these integrals have a analytical solution, this calculation is computationally intensive. Therefore a lookup table has been implemented in which discrete  $r$  and corresponding bulge dispersion are stored.

In the figure above results are compared with observations from [20]. These results were generated using a view distance of 7.5kpc. The other parameters are given in Table 4.1. The radial velocity dispersion heavily depends on the scale parameter of the bulge potential. With the current choice, the dispersion at small  $r$  is lower than realistic. This issue could be solved by adding an additional potential with smaller scale radius. A multi component bulge model has been suggested for instance by [31].

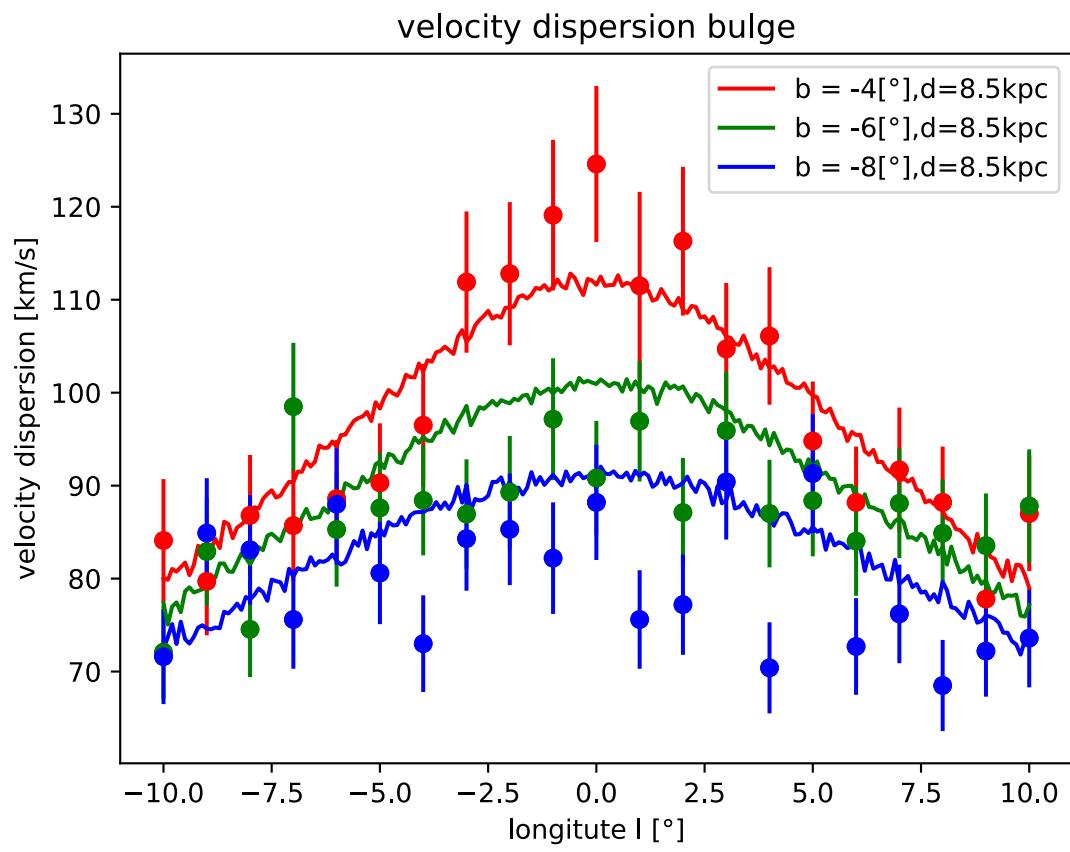


Fig. 5.6: Velocity dispersion of bulge stars



---

CHAPTER  
SIX

---

## INTEGRATION

Numerical integration is needed during initialization and simulation. Various options for time integration have been implemented. The GNU Scientific Library ([14], GSL) is used for integrations performed during initialization. The relevant functions are all based on QAG or QAGI, which have been ported from the Fortran library QUADPACK [52] to C in GLS. The decision trees given on page 79-80 in [52] help with the decision on when and how to use the respective methods.

### 6.1 Quadrature, Adaptive, General-purpose (QAG)

This algorithm makes use of adaptive Gauss-Kronrod quadrature to estimate the definite integral of a given function. Quadrature sums are defined as

$$Q_n[a, b] \equiv \sum_i^n w_i f(x_i) \cong \int_a^b w(x) f(x) dx$$

where  $w_i$  are weights,  $x_i$  nodes,  $w(x)$  a weight function. The highest possible degree of precision is  $2n-1$ . With this maximum precision  $Q_n$  is exact for polynomials of degree smaller or equal to  $2n-1$ .

Using classical Gaussian quadrature formulae, error estimation, by increasing  $n$  to  $n+1$ , requires  $n+1$  evaluations of  $f(x)$  in addition to  $n$  evaluations from calculation of the original sum, since the respective nodes have no common points. By doing so, the degree of precision is only increased from  $2n-1$  to  $2n+1$ . Therefor, the error estimation obtained by subtracting the two sums could be unreliable.

Adding  $n+1$  points to the Gauss-Legendre formula - here  $w(x) = 1$  and the nodes are zeros of the Legendre polynomial - Kronrod introduced the option of increasing the precision to  $3n+1$ , again requiring  $n+1$  additional evaluations of  $f(x)$  ([27]).

$$Q_n^K[a, b] \equiv \sum_i^n w_i f(x_i) + \sum_j^{n+1} w_j^* f(x_j^*) \cong \int_{-1}^1 f(x) dx$$

QAG makes use of this option, bisecting the interval with the largest local absolute error estimate in each step. This division is repeated until either the absolute or relative global error estimate are smaller than required by the caller.

## 6.2 Quadrature, Adaptive, General-purpose, Infinite interval (QAGI)

In case of a semi infinite interval  $(a, \pm\infty)$ , the integration variable is transformed

$$x = a \pm \frac{1-t}{t}$$

leading to

$$\int_a^{\pm\infty} f(x)dx = \pm \int_0^1 f\left(a \pm \frac{1-t}{t}\right) t^{-2} dt$$

For a infinite interval

$$\int_{-\infty}^{\infty} f(x)dx = \int_0^{\infty} f(x) + f(-x)dx = \int_0^1 \left( f\left(\frac{1-t}{t}\right) + f\left(\frac{t-1}{t}\right) \right) t^{-2} dt$$

After the transformation QAGS with the 15-point Kronrod rule is used. QAGS, in addition to the adaptive bisection (see QAG), makes use of the Wynn  $\epsilon$ -algorithm to accelerate the convergence.

## 6.3 Velocity Verlet Algorithm

For cluster members the acceleration is a combination of the force resulting from the presence of all other cluster stars (see Baren Huts Algorithm) and from the milky way potential. The acceleration of field stars solely comes from the milky way potential. In each time step both velocity and acceleration of each star is evaluated.

Since the velocity,  $v(t)$  changes over time, its value at the midpoint between the current ( $t_n$ ) and the next timestep  $t_{n+1}$  is intuitively a better approximation than  $v(t_n)$  or  $v(t_{n+1})$ . The same holds true for the acceleration. This leads to [13]

$$\begin{aligned} x_{n+1} &= x_n + hv_{n+0.5} \\ v_{n+1.5} &= v_{n+0.5} + \frac{h}{m} F(x_{n+1}) \end{aligned} \tag{6.1}$$

which is the Leapfrog algorithm. If one needs  $x$  and  $v$  at the same time, (6.1) can be split into two halve steps.

$$\begin{aligned} v_{n+0.5} &= v_n + \frac{h}{2m} F(x_n) \\ x_{n+1} &= x_n + hv_{n+0.5} \\ v_{n+1} &= v_{n+0.5} + \frac{h}{2m} F(x_{n+1}) \end{aligned}$$

$F(x)$  does not have to be calculated twice, because  $F(x_{n+1})$  can be used as  $F(x_n)$  in the next timestep.

## 6.4 Barnes-Hut Algorithm (BH)

When determining the gravitational force acting on a star which belonging to a cluster, the positions of all other stars in that cluster have to be taken into account.

$$m_i \vec{x}_i = G \sum_{j=1, j \neq i}^N \frac{m_i m_j (\vec{x}_j - \vec{x}_i)}{|\vec{x}_j - \vec{x}_i|^3}$$

Calculating this force for all stars requires  $O(n^2)$  operations. However, the simulated clusters consist of  $\sim 10^4 - 10^5$  stars. While the resulting number of calculations is possible, it is not feasible for a typical desktop computer. Therefor, the Barnes-Hut algorithm (BH) has been implemented which is of order  $O(n \log(n))$ .

The gist of the BH is to approximate a set of stars by their total mass and center of mass (com) if the distance between them and the star, for which the force is to be calculated, is large enough.

The total mass and com of a set of  $m$  stars is

$$m_{com} = \sum_{i=1}^m m_i$$

$$\vec{x}_{com} = \frac{1}{m_{com}} \sum_{i=1}^m m_i \vec{x}_i$$

All cluster stars are stored in an octree. An octree is a data structure where each node in the tree has up to eight child nodes. These nodes split the space represented by their parent node into eight cubes. External nodes are nodes without any children. Each external node contains at most one star. Internal nodes have at least one child. They represent stars stored in their child nodes by storing their total mass and com. The root node contains the whole space occupied by the cluster. Each node stores the following information: total mass, amount and center of mass of stars contained within the cube, two points defining the volume of the cube, one point at the center of the cube, since cpu time is more valuable than ram, as well as links (pointers) to each child node and to the parent node. If a child pointer is null, it does not exist yet.

Stars are added recursively starting at the root node. If the current node is already an internal node, the star is passed to the appropriate child. Mass and com of the internal node are updated. The appropriate child is determined by comparing the position of the star with the center of the node. If the considered node is a external node but already contains a star, both the newly added and already present star are passed down to the appropriate child or children. Consequently, the current node becomes a internal node. Since both stars can lie in the same octant, this can lead to additional recursions until the stars are assigned to different child nodes. If the current node is external and does not yet contain a star, the star is added to the node and the recursion ends.

When calculating the gravitational force on a star, the octree is travelled through recursively beginning with the root node. In case the distance  $d$  between the star and a node is sufficiently large, the stars within that node are approximated by the mass and com of that node, otherwise all child nodes within the current node are considered. Whenever the distance criterion is met, the acceleration vector stemming from the force is calculated, added to the overall acceleration of the star and the recursion for the considered branch stops.

Whether or not  $d$  is big enough, is determined by the quotient  $\theta$ .

$$\theta = \frac{s}{d} < \theta_{max}$$

with  $s$  the side length of the cube and  $\theta_{max}$  a set threshold value. In the special case  $\theta_{max} = 0$ , BH becomes a direct-sum algorithm.  $\theta_{max} = 0.5$  is a commonly chosen value.

doto?: explain smoothing



## COORDINATE SYSTEMS

In the simulation, locations and velocities of stars are stored in galactocentric cartesian coordinates (GCA). However, the observer/scopesim expect heliocentric equatorial polar coordinates. Moreover, observational data is usually given in heliocentric galactic polar coordinates. This data is used for initial cluster positions in the simulation and to compare results. The implementation of transformations is therefor inevitable. The code has been adapted from GalPot ([48]).

### 7.1 Galactocentric Cartesian (GCA)

GCA is a right-handed coordinate system with the galactic center in its origin. The projection of  $\hat{e}_x$  onto the galactic equator (or midplane) points to the initial location of the sun and  $\hat{e}_z$  towards the galactic north pole. Therefor, the direction of galactic rotation at the location of the sun is the negative z axis.

### 7.2 Galactocentric Polar (GCP)

GCP is a spherical coordinate system and similar to GCA. Position and velocity transformation between the two systems goes as follows.

$$\begin{aligned}
 \rho &= \sqrt{x^2 + y^2 + z^2} \\
 \theta &= \text{atan2}(y, x) \\
 \varphi &= \arcsin\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right) \\
 \dot{\rho} &= \frac{x\dot{x} + y\dot{y} + z\dot{z}}{\sqrt{x^2 + y^2 + z^2}} \\
 \dot{\theta} &= \frac{\dot{xy} - x\dot{y}}{x^2 + y^2} \\
 \dot{\varphi} &= \frac{z(x\dot{x} + y\dot{y}) - \dot{z}(x^2 + y^2)}{(x^2 + y^2)\sqrt{x^2 + y^2}}
 \end{aligned} \tag{7.1}$$

## 7.3 Local Standard of Rest (LSR)

Like GCA, LSR is a right-handed coordinate system. The origin of positions is the location of the sun and the origin of velocity is the velocity of a star on a circular orbit with mean velocity of stars in the solar neighborhood.  $\hat{e}_x$  points towards the galactic center,  $\hat{e}_y$  towards the direction of galactic rotation and  $\hat{e}_z$  approximately towards the galactic north pole.

The location of the sun is given by ([25], [5])

$$\vec{x}_{sun,GCA} \cong (8.20, 0, 0.014)[kpc]$$

and the mean velocity ([25])

$$\vec{v}_{mean,GCA} \cong (0, -232.8, 0)[km/s]$$

Since the sun is generally not in the galactic midplane, there is an angle between the planes spanned by  $(x, y)_{GCA}$  and  $(x, y)_{LSR}$ . This angle can be expressed as

$$\begin{aligned}\sin(\alpha) &= \frac{z_{sun,GCA}}{\sqrt{x^2 + z^2}} \\ \cos(\alpha) &= \frac{x_{sun,GCA}}{\sqrt{x^2 + z^2}}\end{aligned}$$

The transformation of position and velocity vectors from GCA to LSR has to contain a rotation by  $-\alpha$  about the y axis.

$$\begin{aligned}x_{LSR} &= \cos(\alpha)(x_{sun,GCA} - x_{GCA}) - \sin(\alpha)(z_{GCA} - z_{sun,GCA}) \\ y_{LSR} &= -y_{GCA} \\ z_{LSR} &= \sin(\alpha)(x_{sun,GCA} - x_{GCA}) + \cos(\alpha)(z_{GCA} - z_{sun,GCA}) \\ u_{LSR} &= -\cos(\alpha)u_{GCA} - \sin(\alpha)w_{GCA} \\ v_{LSR} &= v_{sun,GCA} - v_{GCA} \\ w_{LSR} &= -\sin(\alpha)u_{GCA} + \cos(\alpha)w_{GCA}\end{aligned}$$

## 7.4 Heliocentric Cartesian (HCA)

The only distinction between HCA and LSR is the origin of the velocity. In HCA the origin is the velocity of the sun. The difference between the average velocity in the solar neighborhood and the sun itself, also called the peculiar motion of the sun, is given by ([56])

$$\vec{v}_{sun,LSR} \cong (11.1, 12.24, 7.25)[km/s]$$

Transformation from LSR to HCA can be written as

$$\begin{aligned}\vec{x}_{HCA} &= \vec{x}_{LSR} \\ \vec{v}_{HCA} &= \vec{v}_{LSR} - \vec{v}_{sun,LSR}\end{aligned}$$

## 7.5 Heliocentric Galactic Polar (HGP)

HGP is a spherical coordinate system with origins for position and velocity identical to those defined in HCA. Coordinates given in this system are often called galactic coordinates. The galactic longitude  $l$  is the angular distance in the galactic midplane with  $l = 0^\circ$  towards the galactic center. The galactic latitude  $b$  denotes the angle below and above the galactic midplane ranging from  $-90^\circ$  to  $90^\circ$ . and  $r$  the radial distance.

The transformation from HCA to HGP is identical to the transformation from GCA to GCP (7.1) with  $l \equiv \varphi$ ,  $b \equiv \theta$  and  $\rho \equiv r$

## 7.6 Heliocentric Equatorial Polar (HEQ)

The following positions and the transformation between HGP and HEQ are described in [9] (p 1044-1047), the transformation between different epochs in [33] (p 95-105). The latter is dated. The international celestial reference system (ICRS) should be preferred if highest possible accuracy is of concern.

HEQ, like HGP, is a spherical coordinate system having the same origins for position and velocity. However, angles are given in and normal to the celestial equator which is not parallel to the galactic midplane. The right ascension ( $a$ ) is the angular distance in the equator with  $a = 0^\circ$  towards the northward equinox. The equinox is the intersection of the ecliptic - the plane in which the earth orbits the sun - and the celestial equator. The declination ( $d$ ) is the angular distance above or below in the equator.

Since the ecliptic and the equator are in motion, a reference frame is needed. A reference frame consists of quantities defining the coordinate system at a specific time as well as methods to calculate those quantities for any other date. A commonly used reference frame is defined for the J2000.0 epoch ( $\epsilon_0$ ).

In order to transform between HCA and HEQ at  $\epsilon_0$ , the direction of the north Galactic pole (NGP) and the galactic center (GC) are needed in both basis.

In HCA the NGP is simply  $\vec{x}_{NGP,HCA} = (0, 0, 1)$ . In HGP, since the direction is normal to the fundamental plane,  $b = 90^\circ_{GC,HGP}$ . In HEQ at  $\epsilon_0$  the direction is

$$a_{NGP} = 12^h 51^m 26.28^s$$

$$d_{NGP} = 27^\circ 7' 41.7''$$

The GC defines the x axis of HCA:  $\vec{x}_{GC,HCA} = (1, 0, 0)$ . In GC the same direction is

$$a_{GC,HEQ} = 17^h 45^m 40.0409^s$$

$$d_{GC,HEQ} = -29^\circ 0' 28.118''$$

To express these basis vectors in HCA basis, they can be transformed as follows

$$x_{HCA} = \cos(d) \cos(a)$$

$$y_{HCA} = \cos(d) \sin(a)$$

$$z_{HCA} = \sin(d)$$

The third basis vector is the cross product of  $\vec{x}_{NGP}$  and  $\vec{x}_{GC}$ . With these basis vectors the change of basis matrix is

$$M = [\hat{e}_x, \hat{e}_y, \hat{e}_z]$$

The full transformation from HCA to HEQ consists of the two steps: the multiplication with  $M$  followed by the transformation from cartesian to spherical as given in (7.1).

For the transformation between HGP and HEQ the direction of the north celestial pole (NCP) is required. NCP is perpendicular to the celestial equator, hence  $d_{NCP} = 90^\circ$ . In HGP at  $\epsilon_0$ , NCP is

$$l_{NCP} = 123^\circ 55' 55.2''$$

$$b_{NCP} = 27^\circ 7' 41.7''$$

Using NGP and NCP the transformation from HGP to HEQ at  $\epsilon_0$  is

$$\begin{aligned}\sin(d) &= \sin(d_{NGP}) \sin(b) + \cos(d_{NGP}) \cos(b) \cos(l_{NCP} - l) \\ \cos(d) \sin(a - a_{NGP}) &= \cos(b) \sin(l_{NCP} - l) \\ \cos(d) \cos(a - a_{NGP}) &= \cos(d_{NGP}) \sin(b) - \sin(d_{NGP}) \cos(b) \cos(l_{NCP} - l)\end{aligned}$$

Three angles describing the precision of both planes are needed in order to transform between epochs  $\epsilon_F$  and  $\epsilon_D$ .

$$\begin{aligned}z &= (2306.2181 + 1.39656T - 0.000139T^2)t + (1.09468 + 0.000066T)t^2 + 0.018203t^3 \\ \theta &= (2004.3109 - 0.85330T - 0.000217T^2)t + (-0.42665 - 0.000217T)t^2 - 0.041833t^3 \\ \xi &= (2306.2181 + 1.39656T - 0.000139T^2)t + (0.30188 - 0.000344T)t^2 + 0.017998t^3\end{aligned}$$

where  $t$  and  $T$  are time differences given in units of Julian century.

$$\begin{aligned}T &= \frac{JD(\epsilon_F) - JD(\epsilon_0)}{36525} \\ t &= \frac{JD(\epsilon_D) - JD(\epsilon_F)}{36525}\end{aligned}$$

With these three rotations, a precession matrix  $P$  as well as its inverse can be formalized.

$$P(\epsilon_F, \epsilon_D) = \begin{bmatrix} cz * c\theta * c\xi - sz * s\xi & -cz * c\theta * s\xi - sz * c\xi & -cz * s\theta \\ sz * c\theta * c\xi + cz * s\xi & -sz * c\theta * s\xi + cz * c\xi & -sz * s\theta \\ s\theta * c\xi & -s\theta * s\xi & c\theta \end{bmatrix}$$

where  $cz = \cos(z)$ ,  $sz = \sin(z)$ ,  $c\theta = \cos(\theta)$ ,  $s\theta = \sin(\theta)$ ,  $c\xi = \cos(\xi)$  and  $s\xi = \sin(\xi)$ .

Before multiplication with  $P$ , the target vector has to be expressed in cartesian coordinates.

## 7.7 Heliocentric Telescope Polar (HTP)

This coordinate system aims to describe positions as they are observed with a telescope. As with HEQ the right ascension  $a$  describes the angular distance in and the declination  $d$  the angular distance below and above the midplane. The origin, however is defined by the line of sight vector, that is  $\hat{v}_{los} = [r, a, d] = [1, 0, 0]_{HTP}$ . Hence one image, taken with a telescope, is a circle in the  $ad$ -plane.

It may seem as if the transformation from HGP to HTP and vice versa simply consists of a shift by  $\hat{v}_{los}$  in HGP basis. However for any line of sight vector with  $b \neq 0$ , the result would be an ellipse rather than a circle. Casually speaking, the cone of vision in any spherical coordinate system is a circle when projected onto the unit sphere. While this circle always covers the same range in  $b$ , the range in  $l$  depends on the distance ( $b$ ) between the circle and the midplane. With increasing absolute value of  $b$  the circles around the unit sphere described by  $l$  decrease in size.

While there may be a solution for this issue, a different transformation was devised. Positions can be transformed from HCA to HTP by aligning  $\hat{v}_{los}$  with the x axis  $\hat{a}(x)$  using a rotation matrix generated as described in (5.3) with the rotation axis defined by  $[1, 0, 0]_{HCA}$  and  $\hat{v}_{los,HCA}$ . Subsequent transformation from cartesian to polar coordinates as described in the section GCP above completes the transformation.

---

**CHAPTER****EIGHT**

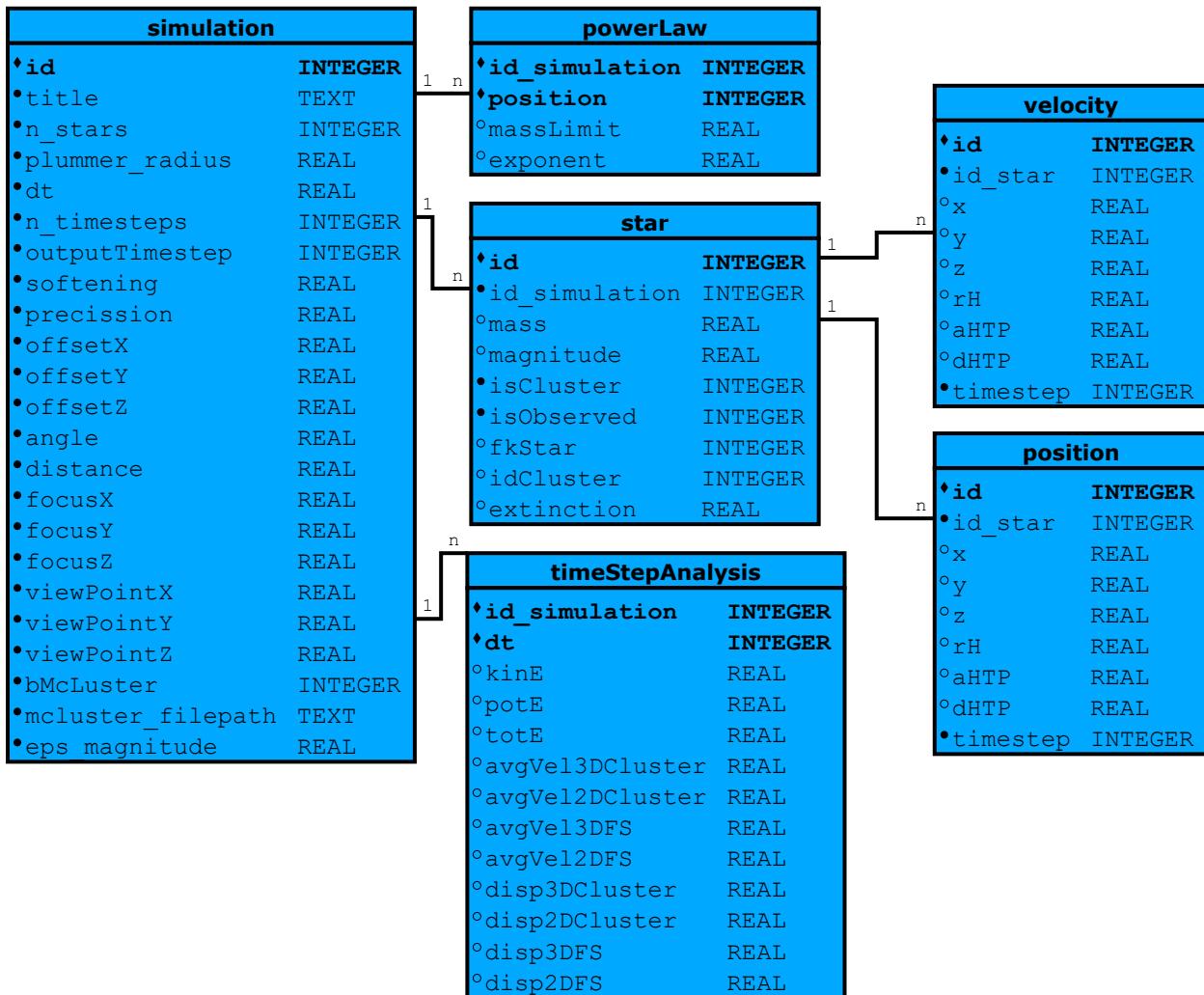
---

**DATABASE**

A SQLite Database is used to store simulations including all relevant parameters, stars with their respective positions and velocities as well as data resulting from analysis. SQLite supports both C++ and python. And is therefore used in most data related parts of the project. Generated data can be loaded and plotted directly from the database. Previously performed simulations can be loaded from the database in order to carry out analysis like energy vs time or average star velocity vs time.

Due to bad performance, no foreign key constraints were set for position.id\_star and velocity.id\_star. A multi-column index is used to improve the execution time of queries containing both velocity.timestep and velocity.id\_star. NULL values are avoided where possible. However, at many stages during execution, some information is unknown, but creating entries is still required. Observed stars contain no information about their mass or GCA phase space coordinates, some can not be mapped to simulated stars, etc.

## 8.1 Entity Relationship Diagram



## MOCK OBSERVATIONS

Images of stars are generated using ScopeSim [22] with the MICADO\_Sci instrument, the ELT telescope and the Armazones location package. Parameters differing from default values are

1. SIM.sub\_pixel.flag - Initial tests indicated the neccesity of integrating sub pixel shifts
2. DET.width and DET.height - Width and height of the detector plane in pix are set ensuring the image contains the whole COV. The pixelscale scale (arcsec/pixel) the ELT is 0.004
3. OBS.dit - The exposure time is adjusted to one hour.
4. scao\_const\_psf.convolve\_mode = same ??? what does this do.
5. scao\_const\_psf.rotational\_blur\_angle - to reflect the rotation of the point spread function (PSF) occuring during one hour, this angle is set to 15 degrees. However, with the approach discussed in *Masking bright stars* this will has little to no impact on the results.

For each simulation two images, one for  $t = 0$  and one for  $t = dt$ , are stored as FITS files.

### 9.1 Creating Source Objects

ScopeSim requires one source object for each simulated star. Source objects contain spatial and spectral information. The spatial information consists of x and y in arcsec relative to the FOV axis, here generated by a chain of transformations from GCA to HTP and stored in the appropriate database table.

The spectral information is aquired using multiple functions from the ScopeSim Templates library involving the following steps:

1. The spectral type of a star is determined by looking up the closest spectral type from A Modern Mean Dwarf Stellar Color and Effective Temperature Sequence [50] given its mass.
2. The spectrum is taken from Pickles catalogue [51] requireing metalicity in range 0.5 - 2 solar and age of about 5 Gyr.
3. That spectrum is scaled such that it has 0 apparent Vega mag in the V filter.
4. The absolute magnitude  $M_V$  of the star in the V filter is determined from its mass using [50] with linear interpolation.
5. The apparent magnitude  $m_V$  is obtained from  $m_V = M_V + 5 \log_{10} (d) - 5$
6. The interstellar extinction  $e$  is taken into account using [24]
7. The weight  $w$  of the spectrum  $s$  is defined as

$$F(x, y) = s * w$$

and calculated via

$$w = \frac{F}{F_0} = 10^{0.4(m_0 - (m_V + e))} = 10^{-0.4(m_V + e)}$$

## DATA REDUCTION

With the help of Photutils [8] the 2D HTP positions and fluxes of stars are detected from the FITS files generated during observation. In a first step all stars are stored in the database, each having exactly one location corresponding to the timestep of the FITS file.

The following method, while yielding decent results, is certainly not optimal. Due to hardware and time constraints, options such as Image Segmentation were not feasible. Moreover, the parameters of the chosen method were not optimized beyond some spot checks.

### 10.1 Detecting Stars

#### 10.1.1 Background

Testing FITS files generated with no input sources still yield some amount of detected sources. With the *Parameters* used for the 25 simulations, 125 sources were detected from an empty background. The FITS files contain a raster of 64 images, sometimes overlapping and sometimes separated by one pixel due to rounding. This leads to wrong detections at the corners. However, this effect only explains a fraction of the detections.

#### 10.1.2 Masking bright stars

PSFs of bright stars were wrongfully detected as stars. Increasing the detection threshold would have meant not detecting real faint stars in different areas. Hence the decision was made to mask square areas around bright stars leading to only one detection within those areas.

The side length  $l_i$  of the square depends on the flux  $F_i$  of the stars:  $l_i = f(F_i)$ .

To find the appropriate function, FITS for single stars at a given distance and with varying mass were generated, the sources in each file detected using DAOStarFinder and their flux and maximum distance between the correct and any wrong sources calculated. A linear fit of the resulting dataset (*Appendix*) lead to the empirical function

$$l_i = \begin{cases} 0 & \text{if } F_i < 100 \\ 0.01 * F_i + 28 & \text{if } F_i \geq 100 \end{cases} \quad (10.1)$$

ToDo?: Add 3 images without mask drawn mask after mask

The DAOStarFinder method is called twice:

1. To find the bright stars and generate the mask. The resulting table contains one row for each source.

This table is sorted by the flux column in descending order and iterated from top to bottom until the current entry has  $F_i < 100$ .

Elements of the mask - a 2D boolean array with the same size as the image - are updated.

All elements inside the box with side length (10.1) are set to true and the current table entry stored in a new table if located outside a masked area.

2. passing the mask parameter generated in the previous step and returning sources outside the masked areas.

Both the bright sources recorded after the first and the faint sources returned from second call are stored in the database.

## CLUSTER ANALYSIS

The objective is to detect all cluster stars present in the observed data with a certain accuracy. One image contains a set of stars  $s_i$  with positions  $\vec{x}_i(t)$  in pixel coordinates and apparent magnitudes  $m_i$ . This information is not enough to detect cluster memberships. The number density of cluster stars in position space is similar to that of the field stars. Furthermore, cluster and field star positions overlap significantly. Moreover, depending on the distance of the cluster to the observer, the cluster size and the field of view angle, cluster stars often cover a significant area of if not the whole image. The proper motion  $\vec{v}_i(t)$  density of cluster stars on the other hand is very high relative to that of field stars and hardly any field stars have  $\vec{v}_i(t)$  close to cluster stars.

### 11.1 Velocity approximation

The proper motion of a star can be approximated from its position at two different timestamps via first order taylor expansion

$$v_i = \frac{x_i(t+dt) - x_i(t)}{dt}$$

However, this is only possible if the positions at two different timesteps can be accurately associated with the same star  $s_i$ . The assumption that  $\vec{x}(t+dt)$  belongs to the same star as  $\vec{x}(t)$  if their Euclidian distance is smaller than the distance of  $\vec{x}(t)$  to any other  $\vec{x}(t+dt)$ , only hold true for small  $dt$ . However, due to the discreteness of images,  $dt$  has to be large enough so that the change in position is detectable between timesteps. The pixel scale - the ratio of arcsec to pixel - dictates a lower bound for  $dt$ .

Since the data results from a simulation, it is trivial to verify whether or not the attribution of the two positions is indeed correct. In fact, tests using only the Euclidian distance as a metric with  $dt = 1\text{ day}$  lead to a small but significant amount of wrong assignments.

An additional condition was introduced. Stars at  $\vec{x}(t)$  and  $\vec{x}(t+dt)$  have to have an similar apparent magnitude  $m$  so be considered the same star.

$$\left| 1 - \frac{m_i(t)}{m_j(t+dt)} \right| < \varepsilon_m$$

where  $\varepsilon_m$  is the maximum relative difference in apparent magnitude.

With this constraint, all positions where correctly assigned during further tests.

## 11.2 DBSCAN

The clustering algorithm of choice is DBSCAN since it is density based and able to detect clusters of arbitrary shape. Additionally and contrary to other algorithms the amount of clusters to find is not a parameter. DBSCAN detects any clusters present in the data based on two parameters:

1.  $\epsilon$ : the maximum distance between points to be considered neighbors
2. nPoints: the minimum amount of neighbors for a point to be classified as core point.

During execution all datapoints are classified as one of the following:

1. core point: a point with at least nPoints within  $\epsilon$ :
2. border point: a point having at least one core point but less than nPoints within  $\epsilon$ :
3. noise/outlier: any other point

The implementation of DBSCAN can be summarized as follows: Iterate the list of points. If the current point is not already classified, check if it meets the requirements to be classified as core point. Once a core point has been found, the neighboring points of that point are tested. If they too have enough neighbors the recursion continues until all neighbors are classified as either core or border points.

In a naive implementation, the distance of each point to every other point is checked. The time complexity of such an implementation is  $O(n^2)$ . Moreover, for large datasets the recursion can lead to stack overflow.

R\*-tree or similar data structures can be used to improve the performance to an average of  $O(n \log n)$  ([12]) The library mlpack ([10]) was used, which includes an implementation of DBSCAN supporting R\*-tree and many other trees.

Initially and in addition to the standard condition  $\epsilon_x$  for the spatial distance between points an supplementary condition  $\epsilon_v$  was introduced. The difference in velocity between two stars has to be smaller than  $\epsilon_v$  to be classified as neighbors:

$$\|\vec{v}_1 - \vec{v}_2\|_2 < \epsilon_v$$

Large  $\epsilon_x$  lead to more accurate membership detection. It turned out, the spatial distance condition does not benefit the results at all and was dropped. For larger areas than used here, constraining the spatial distance, for instance via subdivision, should be beneficial.

## 11.3 Performance

Observed stars are mapped to simulated stars via their proximity in order to measure the performance. If a star is not the closest observed star to any simulated star he remains not mapped. Not mapped stars exist due to Background (Todo: Link) and parts of the PSFs of bright stars being detected as stars. Observed stars can then be attributed one of the following types:

1. True Positive (TP): correctly classified as cluster star.
2. False Positive (FP): wrongly classified as cluster star.
3. True Negative (TN): correctly classified as field star.
4. False Negative (FN): wrongly classified as field star.
5. Unconfirmed Positive (UP): not mapped star classified as cluster star. Treated as FP.
6. Unconfirmed Negative (UN): not mapped star classified as field star. Not taken into account since FN is unlikely.

### 11.3.1 Accuracy

$$A = \frac{TP + TN}{TP + TN + FP + FN}$$

With a large amount of field stars relative to cluster stars, this metric is not ideal as it will give a good rating even if most clusterstars are FN.

### 11.3.2 Precision and Recall

When FPs are more problematic than FNs the precision  $P$  should be high

$$P = \frac{TP}{TP + FP + UP}$$

On the flip side, if FNs are a big concern but FPs tolerable the recall  $R$  is a good metric

$$R = \frac{TP}{TP + FN}$$

### 11.3.3 F1 Score

This metric is a balance between  $P$  and  $R$ . Contrary to  $A$  TN is not taken into account.

$$F_1 = 2 \frac{P * R}{P + R} = \frac{TP}{TP + 0.5(FP + UP + FN)}$$



## EXPERIMENTS

### 12.1 Parameter optimization

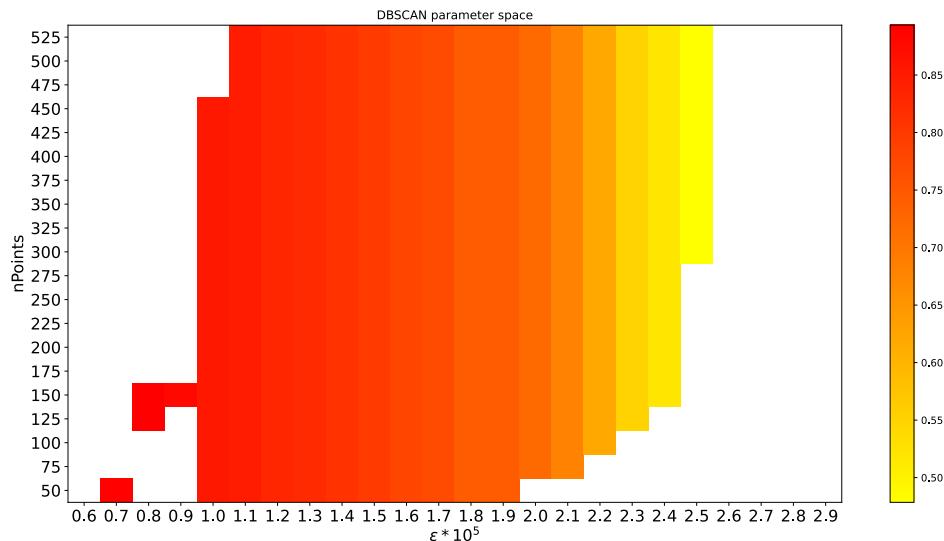


Fig. 12.1: Precision depending on DBSCAN parameters

The quality of the cluster analysis with DBSCAN depends on the choice for its parameters. Fig. 12.1 depicts the precision for a  $1[kM_{\odot}]$  cluster  $300[\text{pc}]$  from the GC. Decreasing  $\epsilon$  leads to a higher precision until it is too small, at which point multiple clusters are detected. Based on this observation  $\epsilon$  was set to  $1.1 * 10^{-5}$  and  $nPoints$  to 200 for all simulations.

## 12.2 Setup

A total of 25 simulations with varying focus and cluster mass were carried out in order to study the effects of these parameters on the cluster detection performance. The focus was set to  $l = 0^\circ$  and  $b \in \{0, 5, 10, 25, 180\} [^\circ]$  HGP and the cluster mass  $\in \{0.64, 1.6, 4.0, 10, 25\} [\text{kM}_\odot]$ . Each simulation was repeated 10 times for error estimation.

### 12.2.1 Parameters

The following parameters remained unchanged between the simulations.

#### General simulation parameters

- FOV angle: 54[arcsec]  
The field of view angle, chosen large enough to cover most of cluster with the highest mass. See  $\alpha$  in Fig. 5.2
- View distance: 9[kpc]  
The height of the COV or line of sight distance reaching behind the GC when looking towards it.
- Cluster distance: 8[kpc]  
The distance between the observer and the cluster. When looking straight at the GC the distance between the cluster and the GC is 300[pc]  
The mean cluster velocity is set to the circular velocity at this location.
- View point:  $[8300 \quad 0 \quad 27]^T_{GCA}$  [pc]  
The position of the observer in GCA.
- Timestep: 28[day]  
Amount of time between the two recorded snapshots. The time per integration is 7[day]. Therefor snapshots are taken every 4 integrations.

#### Cluster specific parameters (McLuster)

- (P) Profile: 0 Plummer density profile
- (R) Radius: -1  
With this setting the radius is calculated by mcluster via a mass to half-mass radius relation as described in [23]
- (Q) Virial ratio: 0.5  
The cluster is in virial equilibrium.
- (f) IMF: 1 Kroupa ranging from 0.08 Msun to 100 Msun
- (C) Output: 3 The resulting mass, position and velocity for each star is written into a file.

## 12.3 Results and Interpretation

Depending on the total cluster mass the amount of simulated CS (SCS) ranges from  $1.3e3$  to  $40.4e3$  while the amount of mapped CS (MCS) lies between  $1.0e3$  and  $16.4e3$ . The decrease in detectability of CS is due to increasing CS density and has no direct impact on the clustering performance. The difference between observed stars (OS) and mapped stars (MS) is negligible, in fact the amount of non mapped stars decreases with increasing number of SCS.

While the number of simulated FS (SFS) remains constant, the number of mapped FS (MFS) decreases with increasing number of CS because bright CS hide FS. This inverse relationship does impact the clustering performance to some degree, less MFS means less potential FPs as well as TNs, the latter are not relevant for Precision and F1 score.

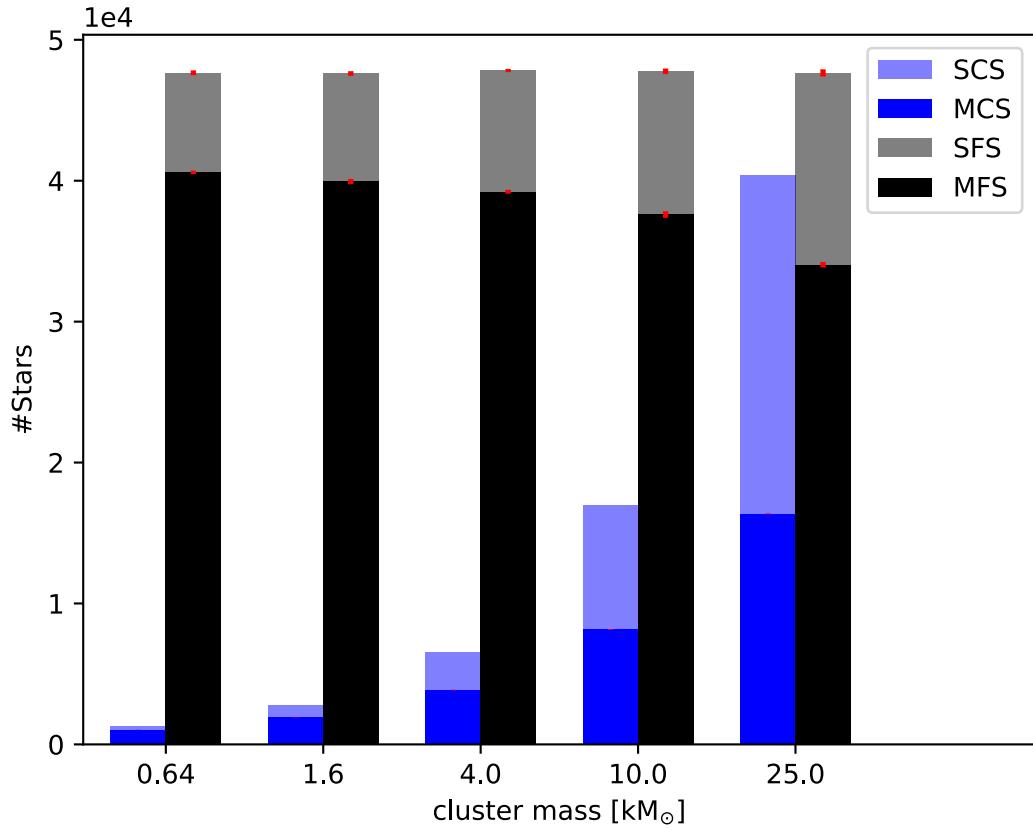


Fig. 12.2: Number of simulated/detected cluster and field stars at  $10^\circ$  depending on the cluster masses

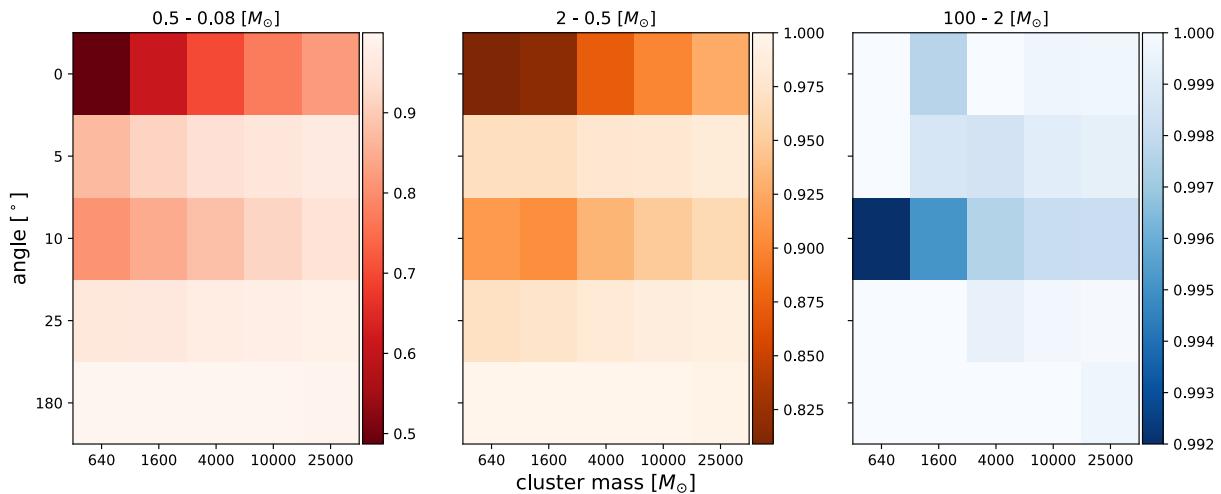


Fig. 12.3: Precision for different mass bins, angles and cluster masses

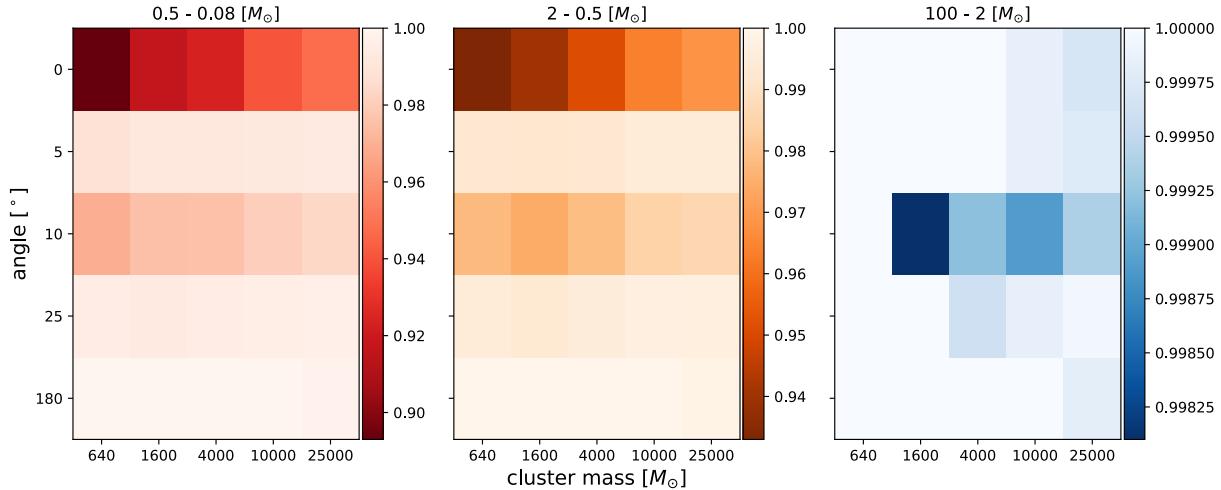


Fig. 12.4: Precision using accurate velocities

Fig. 12.3 displays the precision using the velocity of OS and Fig. 12.4 for the velocity of SS. The only relevant difference between simulated and observed HTP proper motion is the the accuracy of position and consequently of velocity. While both figures show the same relationships, the drop in overall performance due to inaccuracies introduced during observation and source detection are painfully apparent.

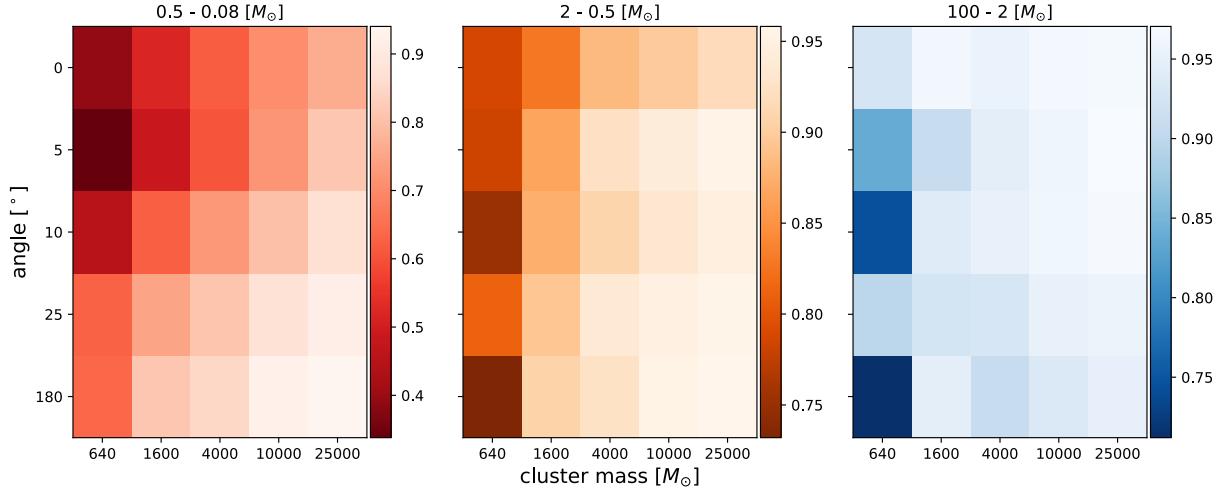


Fig. 12.5: F1 score for different mass bins, angles and cluster masses

As is visible in Fig. 12.3 and Fig. 12.4 the precision, with one exception, is correlated with the angle. Curiously for the same cluster mass the precision is lower at  $10^{\circ}$  than at  $5^{\circ}$ .

Fig. 12.6 provides the explanation for this outlier. At  $10^{\circ}$  the average field star velocity is closer to the average cluster star velocity than at any other angle, making it harder to differentiate between cluster and noise.

The bigger the cluster mass the higher the cluster star velocity density, which implies the second correlation - precision with cluster mass - presented in Fig. 12.3.

Fig. 12.7 displays examples for the 2D HTP velocity space of two simulated clusters near the GC.

In this example the  $0.64[kM_{\odot}]$  cluster only has 1143 stars inside the circle while  $10[kM_{\odot}]$  has 3158. In both cases

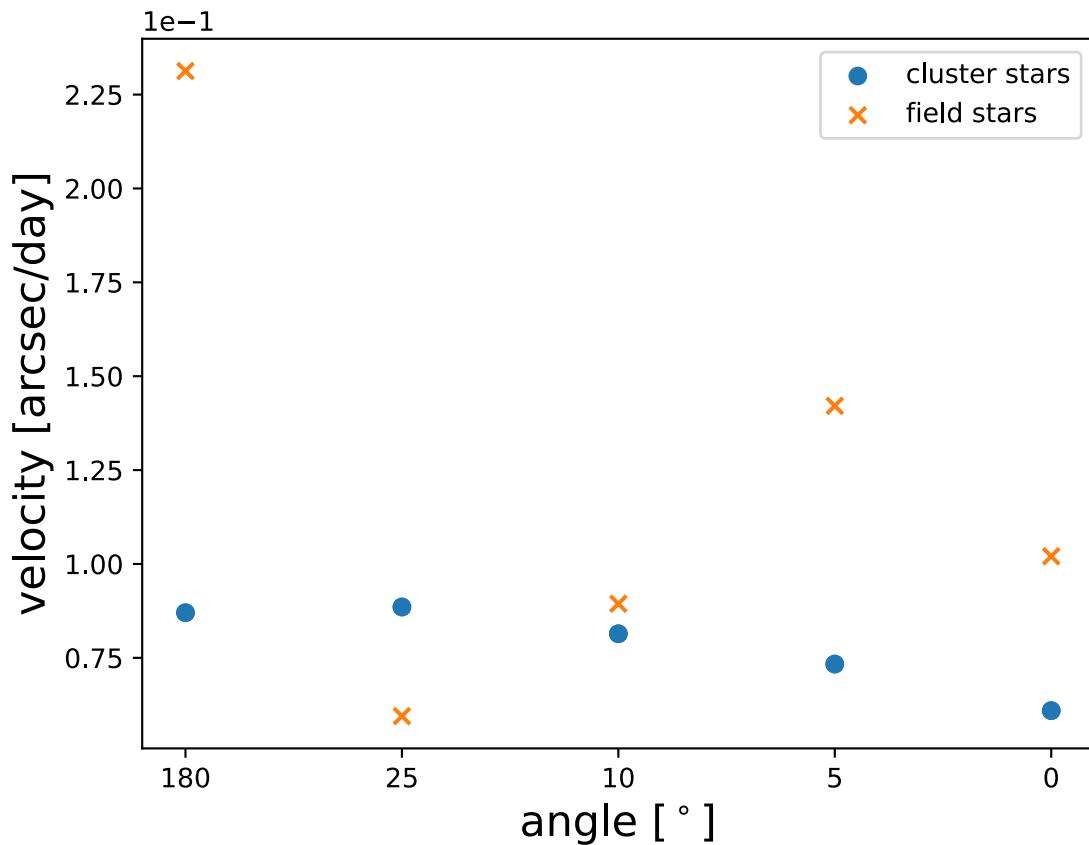


Fig. 12.6: Average cluster and field star velocity at different angles

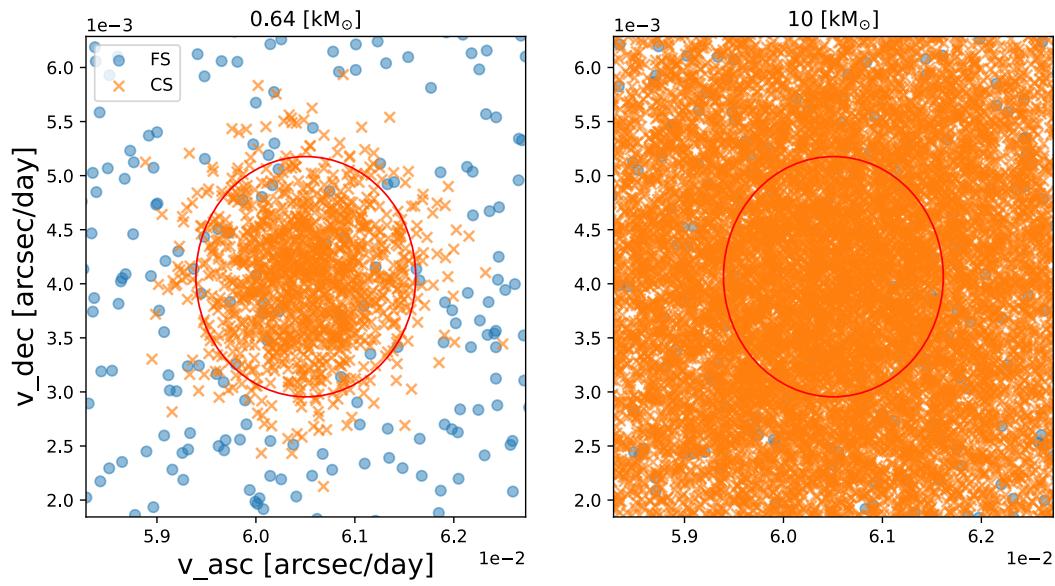


Fig. 12.7: 2D HTP velocity of simulated clusters

statistically the same amount of field stars fall within that area, leading to a higher ratio of FPs and therefore a lower precision for the lower mass cluster.

---

CHAPTER  
**THIRTEEN**

---

## CONCLUSIONS

The results presented, particularly Fig. 12.3, indicate that data gathered with the ELT will indeed be accurate enough to study the IMF of YMCs near the GC. However, it remains questionable if this is still the case for stars with mass  $< 0.08[M_\odot]$ . Information about this mass regime, such as spectral type, MW disc and core IMFs are sparse. Moreover, the precision obtained here in the range  $0.5 - 0.08[M_\odot]$  is already quite low. That is not to say it won't be possible. Many improvements and additions could be made to the presented work such as

1. A more accurate bulge model, for instance [36]
2. Additional snapshots to recover temporarily hidden stars and get a more accurate velocity estimation.
3. A sophisticated method for extracting positional data like point spread function fitting.
4. Accurately determine the ideal time step size.



---

**CHAPTER  
FOURTEEN**

---

**APPENDIX**

Table 14.1: Dataset for empirical mask function

flux	box size	dist	mass
927146.1955	254.3559872	127.1779936	10
890635.019	254.3417547	127.1708773	9.9
857036.8725	254.341936	127.170968	9.8
823955.0858	254.3421943	127.1710971	9.7
793135.9042	254.3424033	127.1712017	9.6
762592.0798	254.3426546	127.1713273	9.5
734504.8423	254.3429091	127.1714545	9.4
707751.1254	254.3431486	127.1715743	9.3
681256.1109	254.3434061	127.1717031	9.2
656346.0444	254.3436538	127.1718269	9.1
633104.9017	254.3439632	127.1719816	9
609936.5628	254.3442026	127.1721013	8.9
587640.2069	254.3445212	127.1722606	8.8
566233.046	254.3448084	127.1724042	8.7
545639.6678	254.3451645	127.1725822	8.6
525443.3219	254.3454953	127.1727477	8.5
506252.3011	254.4948138	127.2474069	8.4
487489.4659	254.3462375	127.1731188	8.3
470008.9293	254.3465892	127.1732946	8.2
462754.3324	254.3470858	127.1735429	8.1
445797.2677	254.5130792	127.2565396	8
429456.9678	254.347868	127.173934	7.9
413760.4129	254.3482648	127.1741324	7.8
398652.2024	254.3487825	127.1743912	7.7
384123.4617	254.4173555	127.2086778	7.6
370111.5523	253.8093122	126.9046561	7.5
356487.63	254.4810594	127.2405297	7.4
343530.5029	255.0658948	127.5329474	7.3
316316.5998	247.8349505	123.9174752	7.2
309163.7247	254.4288941	127.2144471	7.1
302130.3778	254.4303415	127.2151707	7
295295.4188	254.5012123	127.2506062	6.9
288627.7467	254.433333	127.2166665	6.8
282129.2052	254.4345462	127.2172731	6.7
275630.1016	254.436262	127.218131	6.6
269416.7466	254.4377591	127.2188795	6.5

continues on next page

Table 14.1 – continued from previous page

flux	box size	dist	mass
263376.811	247.9665145	123.9832573	6.4
257390.5638	247.9972751	123.9986376	6.3
251577.0397	252.6213071	126.3106536	6.2
245854.0065	248.0135846	124.0067923	6.1
236400.0465	248.0575239	124.0287619	6
227300.7517	252.618173	126.3090865	5.9
218570.2985	252.6201478	126.3100739	5.8
210129.5831	247.9292325	123.9646163	5.7
202041.5421	247.85708	123.92854	5.6
194255.2774	252.6363119	126.318156	5.5
186750.9496	252.6021078	126.3010539	5.4
182535.794	247.9673489	123.9836745	5.3
178401.698	248.0481722	124.0240861	5.2
174325.8134	252.6377075	126.3188537	5.1
170400.5422	252.6334207	126.3167103	5
248519.9168	247.7858731	123.8929365	4.9
242908.6771	247.9706455	123.9853228	4.8
237403.2487	248.0166034	124.0083017	4.7
232060.6523	247.9419667	123.9709833	4.6
218259.9677	254.6442946	127.3221473	4.5
205273.3003	247.9794943	123.9897472	4.4
193150.6701	250.4719141	125.2359571	4.3
181724.6168	252.6310455	126.3155227	4.2
170942.9686	252.6333644	126.3166822	4.1
160789.262	252.6181038	126.3090519	4
168811.5177	247.8041635	123.9020817	3.9
162697.5008	252.6385607	126.3192803	3.8
156860.6692	252.6179306	126.3089653	3.7
151196.1118	247.8991699	123.949585	3.6
145724.8797	248.360639	124.1803195	3.5
150293.5107	248.072207	124.0361035	3.4
130973.0674	255.4786549	127.7393274	3.3
114105.3359	252.5656424	126.2828212	3.2
99358.44939	248.5393443	124.2696721	3.1
86537.23392	251.6269223	125.8134612	3
75357.58044	251.6432071	125.8216036	2.9
65650.99523	251.6614012	125.8307006	2.8
63683.16206	251.6658441	125.8329221	2.7
61750.32575	251.67074	125.83537	2.6
59884.23739	252.6195792	126.3097896	2.5
57772.09156	251.6814637	125.8407319	2.4
50097.62697	253.4222927	126.7111463	2.3
43479.65427	216.2939201	108.14696	2.2
33899.676	248.6107083	124.3053542	2.1
29960.07512	248.6307907	124.3153953	2
24613.56833	150.7541002	75.37705012	1.9
36933.09895	198.6613702	99.3306851	1.8
29129.77414	199.0091892	99.50459462	1.7
27120.64332	216.4733782	108.2366891	1.6

continues on next page

Table 14.1 – continued from previous page

flux	box size	dist	mass
20229.98476	170.1296746	85.06483729	1.5
19179.36077	179.7883487	89.89417436	1.4
15120.70563	151.6853158	75.84265791	1.3
11210.70479	86.37067622	43.18533811	1.2
7990.562518	62.34638126	31.17319063	1.1
6367.416643	103.8925506	51.94627529	1
4117.980781	51.12273987	25.56136994	0.9
3939.712223	52.55001005	26.27500503	0.8
2698.416769	51.4333436	25.7166718	0.7
1046.477108	27.12953968	13.56476984	0.6
727.6929275	27.40087693	13.70043846	0.5
537.6169065	21.17919258	10.58959629	0.4
201.0127308	20.96270907	10.48135453	0.3
96.41247775	0	0	0.2
26.95062209	0	0	0.1

Table 14.2: Number of mapped stars, partly displayed in Fig. 12.2

Angle	Mass	MNCS		MNCS		MNCS		MNCS		MNCS		MNFS		MNFS		MNFS		MNFS		MNFS	
		Tot	Err	>2	>2	2 - 0.5	2 - 0.5	0.5 -	0.5 -	0.08	0.08	Err	Tot	Err	>2	>2	2 - 0.5	2 - 0.5	0.5 -	0.5 -	0.08
180	640	957.5	7.5	36.9	0.32	206.5	2.1	714.1	5.9	655	26	1.2	0.92	120.5	6.3	534	26				
180	1600	1557.	18.2	83.4	0.7	305	1.8	1168.	7.6	644	29	1.2	1.2	113.4	9.7	529	36				
180	4000	2607	11	191.3	0.67	522.3	3.5	1893.	28.5	607	18	0.9	0.99	114.4	6.6	491	18				
180	10000	6175	17	366.8	1.5	1276.	13.4	4532	14	579	16	1	0.94	106.6	7.6	471	22				
180	25000	15565	551	840	5.2	3533	18	1119	230	535	21	1.1	0.74	106.4	6.6	428	26				
25	640	1015	13	39.2	1	249.4	6.1	726.7	8.8	1747	6161	33.1	5.4	3354	37	14088	184				
25	1600	1790	14	88.9	0.88	357.8	5	1344	13	1722	6115	31.9	5.3	3276	55	13918	148				
25	4000	3231	22	184.9	1.4	695.6	9.2	2350	20	1667	972	31.8	5.2	3185	38	13462	283				
25	10000	7522	64	407.4	3.7	1646	21	5468	45	15944	121	30.3	7.2	3159	39	12754	131				
25	25000	16291	37	910.6	4.7	3804	12	11577	29	14855	135	28.4	2.5	3025	25	11802	2143				
10	640	1044	29.9	39.6	1.1	255	2.9	749.6	9.9	40591	1126	77.8	7.9	8012	59	32502	150				
10	1600	1933	21	92.7	1.2	405.4	4.2	1435	20	39948	178	80	8.4	7868	75	32000	204				
10	4000	3860	25	215.6	1.3	857.2	5.7	2787	22	39216	137	71.8	8.1	7686	28	31458	144				
10	10000	8156	29	435.1	1.7	1832	12	5889	22	37598	233	68.8	5.8	7460	46	30069	248				
10	25000	16362	235	917.4	5.3	3950	12	11496	28	34042	192	67.2	9.2	7169	82	26807	216				
5	640	1041	11	40.6	0.52	265.5	2.3	734	10	56466	189	97.4	8.8	10896	65	45474	228				
5	1600	1972	21	90.2	1.4	425.7	6	1456	18	55920	267	88.2	8.1	10803	77	45029	302				
5	4000	4087	23	222.4	2.2	912.4	8.9	2952	17	54850	270	79.3	6.8	10632	77	44139	324				
5	10000	8865	22	434.8	3.7	2033.96		6396	19	52700	231	88	10	10322	78	42289	241				
5	25000	16224	61	946.2	5.8	4118	14	11160	50	47312	303	82.4	6.9	9781	93	37448	319				
0	640	892	17	38.7	1.5	250.4	4	603	15	24946	589	89	8.7	44325	177	20504	728				
0	1600	1676	25	87.6	3.3	402	5.3	1186	21	24614	843	90	10	44216	170	20183	877				
0	4000	3680	24	221.1	3.4	955	12	2504	15	24044	696	88.8	7.4	43939	112	19641	670				
0	10000	8053	51	457.2	4.6	2159	6.1	5437	45	22969	423	84.3	9.5	43348	181	18626	657				
0	25000	15170	86	972	11	4484	31	9715	73	20331	642	80.1	5.8	42162	175	16106	671				

Table 14.3: Number of simulated stars, partly displayed in Fig. 12.2

Angle	Mass	SNC	SSNC	SSNC	SSNC	SSNC	SSNC	SSNC	SSNC	SSNF	SSNF	SSNF	SSNF	SSNF	SSNF	SSNF	SSNF
		Tot	Tot	> 2	> 2	2 - 0.5	2 - 0.5	0.5 -	0.5 -	0.08	0.08	Tot	Tot	> 2	> 2	2 - 0.5	2 - 0.5
180	640	1326	0	43	0	292	0	991	0	687	22	1.2	0.92	125.2	6.3	560	24
180	1600	2777	0	105	0	521	0	2151	0	697	27	1.2	1.2	122	11	574	36
180	4000	6521	0	261	0	1284	0	4976	0	684	21	0.9	0.99	127.2	5.7	556	22
180	10000	169490	0	646	0	3403	0	129000	0	689	23	1.1	0.88	124.3	7.7	563	29
180	25000	404170	0	1592	0	8217	0	306080	0	678	25	1.2	0.63	123.8	6	553	29
25	640	1322	0	44	0	293	0	985	0	19730	161	33.9	5.3	3588	37	16108	187
25	1600	2777	0	106	0	525	0	2146	0	19809	151	33.5	5.1	3558	52	16217	186
25	4000	6528	0	262	0	1287	0	4979	0	19809	97	33.8	6.1	3558	49	16217	119
25	10000	169300	0	648	0	3388	0	128940	0	19690	180	32.4	8.2	3608	34	16050	188
25	25000	404530	0	1599	0	8198	0	306560	0	19845	164	31.9	2.4	3575	28	16238	187
10	640	1326	0	44	0	293	0	989	0	47660	171	79.7	7.6	8547	62	39033	215
10	1600	2775	0	106	0	523	0	2146	0	47606	167	82.5	8.8	8510	70	39014	202
10	4000	6518	0	262	0	1284	0	4972	0	47819	113	76.5	8.6	8475	36	39268	135
10	10000	169460	0	647	0	3395	0	129040	0	47765	195	74.6	6.7	8484	55	39206	242
10	25000	404320	0	1593	0	8202	0	306370	0	47652	255	76.3	9.6	8521	64	39054	298
5	640	1328	0	43	0	293	0	992	0	67394	184	99.7	9.2	11600	66	55695	217
5	1600	2777	0	106	0	522	0	2149	0	67521	188	90.8	7.8	11612	77	55819	242
5	4000	6517	0	262	0	1283	0	4972	0	67488	308	84.6	6.7	11614	79	55790	371
5	10000	169530	0	647	0	3402	0	129040	0	67468	264	94	11	11614	76	55760	289
5	25000	404060	0	1590	0	8207	0	306090	0	67524	319	92.7	8.4	11581	74	55850	365
0	640	1326	0	43	0	292	0	991	0	35132	252	96.6	9.1	50872	171	30035	371
0	1600	2777	0	105	0	521	0	2151	0	35122	245	97	11	50918	148	30020	883
0	4000	6521	0	261	0	1284	0	4976	0	35122	2818	96	7.5	50900	133	30023	428
0	10000	169500	0	646	0	3403	0	129010	0	35106	943	94.8	9.9	50917	159	30005	468
0	25000	404180	0	1592	0	8217	0	306090	0	35115	418	93.5	7.6	50967	162	30009	650

Table 14.4: Average velocity of mapped stars, partly displayed in Fig. 12.6

Angle	Mass	OCV	OCV Err	OFV	OFV Err
180	640	0.00310743	0.00000084	0.00826	0.00022
180	1600	0.00310694	0.00000036	0.00814	0.00017
180	4000	0.00310659	0.00000046	0.0083	0.00026
180	10000	0.00310572	0.00000094	0.00827	0.0002
180	25000	0.004132217	0.00000008	0.00823	0.00024
25	640	0.00316207	0.00000076	0.002124	0.000015
25	1600	0.00315911	0.00000084	0.002133	0.000016
25	4000	0.0031645	0.00000007	0.0021225	0.0000096
25	10000	0.00315688	0.00000034	0.002133	0.000013
25	25000	0.00315863	0.00000021	0.002125	0.000012
10	640	0.0029074	0.0000027	0.003193	0.000014
10	1600	0.0029064	0.0000025	0.003193	0.000012
10	4000	0.0029064	0.0000012	0.0031919	0.0000097
10	10000	0.00290469	0.00000084	0.003195	0.000013
10	25000	0.002904	0.00000063	0.0031957	0.0000092
5	640	0.0026192	0.0000066	0.005075	0.000014
5	1600	0.0026194	0.0000036	0.005077	0.00001
5	4000	0.0026213	0.0000026	0.005072	0.000011
5	10000	0.0026156	0.0000013	0.005072	0.00001
5	25000	0.002615	0.0000014	0.0050729	0.0000086
0	640	0.0021763	0.000009	0.0036446	0.0000056
0	1600	0.0021772	0.000006	0.0036443	0.0000027
0	4000	0.0021795	0.0000041	0.0036421	0.0000038
0	10000	0.0021774	0.0000024	0.0036425	0.0000049
0	25000	0.0021747	0.0000011	0.0036398	0.0000051

Table 14.5: Precision for mapped stars. see Fig. 12.3

Angle	Mass	OP Tot	OP Err	OP > 2	OP > 2 Err	OP 2 - 0.5	OP 2 - 0.5 Err	OP 0.5 - 0.08	OP 0.5 - 0.08 Err
180	640	0.99979	0.00067	1	0	1	0	0.9997	0.00095
180	1600	0.99947	0.00069	1	0	1	0	0.99925	0.00097
180	4000	0.99905	0.00055	1	0	1	0	0.99864	0.00079
180	10000	0.99863	0.00051	1	0	1	0	0.99811	0.00071
180	25000	0.99363	0.00084	0.9996	0.00089	0.9967	0.00095	0.9922	0.0012
25	640	0.961	0.013	1	0	0.9704	0.0095	0.952	0.019
25	1600	0.9654	0.0084	1	0	0.9733	0.0085	0.96	0.011
25	4000	0.976	0.0036	0.9994	0.002	0.9817	0.0051	0.9716	0.0042
25	10000	0.9808	0.0015	0.99973	0.00086	0.9873	0.0039	0.9769	0.0017
25	25000	0.9855	0.001	0.99988	0.00038	0.98962	0.00089	0.9828	0.0014
10	640	0.855	0.019	0.992	0.018	0.913	0.019	0.809	0.025
10	1600	0.8733	0.0089	0.9952	0.0084	0.905	0.017	0.847	0.011
10	4000	0.9015	0.0095	0.9975	0.0027	0.932	0.012	0.879	0.011
10	10000	0.93	0.0029	0.998	0.001	0.9483	0.0062	0.9167	0.0028
10	25000	0.9507	0.002	0.9981	0.0018	0.9616	0.0045	0.9422	0.0024
5	640	0.926	0.013	1	0	0.9683	0.008	0.871	0.024
5	1600	0.9394	0.0095	0.9987	0.0042	0.9683	0.0095	0.912	0.015
5	4000	0.9567	0.0052	0.9985	0.0034	0.9779	0.0061	0.9384	0.0069
5	10000	0.9631	0.0028	0.999	0.0017	0.9788	0.0032	0.9518	0.0041
5	25000	0.9725	0.0016	0.9992	0.0011	0.9839	0.0015	0.9642	0.0021
0	640	0.625	0.023	1	0	0.807	0.027	0.486	0.029
0	1600	0.696	0.022	0.9975	0.0052	0.816	0.027	0.611	0.024
0	4000	0.771	0.011	1	0	0.8721	0.0076	0.698	0.015
0	10000	0.8236	0.0066	0.99953	0.00099	0.8989	0.0056	0.7718	0.0085
0	25000	0.8665	0.0037	0.99967	0.00053	0.926	0.0054	0.8212	0.0044

Table 14.6: Precision for simulated stars. see Fig. 12.4

Angle	Mass	SP Tot	SP Err	Tot	SP > 2	SP Err	> 2	SP 2 - 0.5	SP Err	2 - 0.5	SP 0.5 - 0.08	SP Err	0.5 - 0.08
180	640	1	0		1	0		1	0		1	0	
180	1600	1	0		1	0		1	0		1	0	
180	4000	1	0		1	0		1	0		1	0	
180	10000	1	0		1	0		1	0		1	0	
180	25000	0.99817	0.00022		0.99981	0.00043		0.99847	0.00045		0.998	0.00035	
25	640	0.9946	0.0018		1	0		0.9946	0.004		0.9943	0.0021	
25	1600	0.9934	0.002		1	0		0.9931	0.003		0.9931	0.0024	
25	4000	0.99496	0.00057		0.9996	0.0012		0.995	0.0016		0.9947	0.00067	
25	10000	0.99569	0.00062		0.99984	0.00049		0.99663	0.00089		0.99523	0.00092	
25	25000	0.9962	0.00031		0.99994	0.0002		0.99654	0.00038		0.99591	0.00038	
10	640	0.9719	0.0042		1	0		0.978	0.0093		0.9689	0.0044	
10	1600	0.9756	0.0029		0.9981	0.0059		0.9742	0.0048		0.9749	0.0039	
10	4000	0.9768	0.0021		0.9992	0.0016		0.9787	0.004		0.9751	0.0023	
10	10000	0.9816	0.001		0.9989	0.001		0.9841	0.0023		0.9801	0.0014	
10	25000	0.9845	0.00048		0.99937	0.00089		0.9861	0.0015		0.98332	0.00055	
5	640	0.9892	0.0032		1	0		0.9925	0.0038		0.9878	0.0038	
5	1600	0.9914	0.0012		1	0		0.9918	0.0031		0.9909	0.0021	
5	4000	0.9918	0.0013		1	0		0.9926	0.0029		0.9912	0.0014	
5	10000	0.99294	0.00085		0.99984	0.00049		0.9943	0.0014		0.99224	0.00083	
5	25000	0.99367	0.00039		0.99975	0.00044		0.99459	0.00068		0.9931	0.00046	
0	640	0.905	0.015		1	0		0.933	0.025		0.893	0.016	
0	1600	0.923	0.012		1	0		0.9399	0.006		0.916	0.015	
0	4000	0.932	0.0049		1	0		0.9508	0.0058		0.9239	0.0058	
0	10000	0.9472	0.0023		0.99984	0.00049		0.9636	0.0035		0.9405	0.0023	
0	25000	0.9536	0.0014		0.99969	0.00044		0.968	0.0018		0.9475	0.0018	

Table 14.7: F1 Score for mapped stars. see Fig. 12.5

Angle	Mass	OF1 Tot	OF1 Tot Err	OF1 > 2	OF1 > 2 Err	OF1 2 - 0.5	OF1 2 - 0.5 Err	OF1 0.5 - 0.08	OF1 0.5 - 0.08 Err
180	640	0.66	0.04	0.712	0.073	0.732	0.05	0.635	0.035
180	1600	0.8398	0.006	0.944	0.012	0.9068	0.0044	0.8125	0.0072
180	4000	0.8682	0.0059	0.9075	0.0062	0.924	0.01	0.8476	0.0055
180	10000	0.9315	0.0029	0.934	0.0037	0.9531	0.0038	0.925	0.003
180	25000	0.9449	0.0014	0.9492	0.0019	0.9582	0.0018	0.9403	0.0014
25	640	0.688	0.043	0.897	0.022	0.811	0.056	0.626	0.041
25	1600	0.789	0.019	0.924	0.028	0.893	0.024	0.747	0.018
25	4000	0.8462	0.0043	0.9278	0.0035	0.9355	0.0051	0.8098	0.0059
25	10000	0.8937	0.0035	0.9511	0.0023	0.9476	0.003	0.8719	0.0045
25	25000	0.9277	0.0012	0.9556	0.0024	0.9547	0.002	0.9163	0.0013
10	640	0.547	0.015	0.744	0.033	0.751	0.019	0.45	0.018
10	1600	0.7	0.011	0.9384	0.0097	0.875	0.012	0.621	0.015
10	4000	0.7817	0.0063	0.9507	0.0078	0.9079	0.0081	0.7224	0.0077
10	10000	0.8395	0.0044	0.9594	0.0029	0.9318	0.0034	0.7979	0.0056
10	25000	0.8927	0.0017	0.9671	0.0024	0.9467	0.0026	0.8667	0.0021
5	640	0.501	0.015	0.838	0.014	0.782	0.018	0.338	0.019
5	1600	0.61	0.014	0.907	0.014	0.865	0.014	0.486	0.019
5	4000	0.7122	0.0067	0.9448	0.009	0.9233	0.0076	0.6061	0.0074
5	10000	0.7906	0.0036	0.9585	0.0033	0.941	0.0035	0.7196	0.0046
5	25000	0.8608	0.0029	0.9705	0.002	0.9539	0.0019	0.8117	0.0037
0	640	0.538	0.016	0.926	0.018	0.786	0.02	0.39	0.021
0	1600	0.626	0.014	0.9622	0.01	0.828	0.018	0.518	0.016
0	4000	0.7147	0.0089	0.9536	0.0042	0.8837	0.0068	0.619	0.012
0	10000	0.7749	0.0048	0.9641	0.0036	0.8978	0.0038	0.7053	0.0066
0	25000	0.824	0.0032	0.9657	0.0021	0.9174	0.0034	0.7637	0.004

---

CHAPTER  
**FIFTEEN**

---

## **INDICES AND TABLES**

- genindex
- modindex
- search



## BIBLIOGRAPHY

- [1] Anisa Bajkova and Vadim Bobylev. Parameters of six selected galactic potential models. *Open Astronomy*, 26(1):72–79, 2017. URL: <https://doi.org/10.1515/astro-2017-0016>, doi:doi:10.1515/astro-2017-0016.
- [2] Nate Bastian, Kevin R. Covey, and Michael R. Meyer. A universal stellar initial mass function? a critical look at variations. *Annual Review of Astronomy and Astrophysics*, 48(1):339–389, Aug 2010. URL: <http://dx.doi.org/10.1146/annurev-astro-082708-101642>, doi:10.1146/annurev-astro-082708-101642.
- [3] H. Baumgardt and S. Sollima. The global mass functions of 35 galactic globular clusters – ii. clues on the initial mass function and black hole retention fraction. *Monthly Notices of the Royal Astronomical Society*, 472(1):744–750, Aug 2017. URL: <http://dx.doi.org/10.1093/mnras/stx2036>, doi:10.1093/mnras/stx2036.
- [4] James Binney. Dynamics for Galactic Archaeology. *New Astron. Rev.*, 57:29–51, 2013. arXiv:1309.2794, doi:10.1016/j.newar.2013.08.001.
- [5] James Binney, Ortwin Gerhard, and David Spergel. The photometric structure of the inner Galaxy. *Monthly Notices of the Royal Astronomical Society*, 288(2):365–374, 06 1997. URL: <https://doi.org/10.1093/mnras/288.2.365>, arXiv:<https://academic.oup.com/mnras/article-pdf/288/2/365/2841167/288-2-365.pdf>, doi:10.1093/mnras/288.2.365.
- [6] James Binney and Scott Tremaine. *Galactic Dynamics: Second Edition*. Princeton University Press, Princeton, rev - revised, 2 edition, 2011. ISBN 9780691130262.
- [7] Jo Bovy. Galpy: a python library for galactic dynamics. *The Astrophysical Journal Supplement Series*, 216(2):29, Feb 2015. URL: <http://dx.doi.org/10.1088/0067-0049/216/2/29>, doi:10.1088/0067-0049/216/2/29.
- [8] Larry Bradley, Brigitta Sipocz, Thomas Robitaille, Erik Tollerud, Zé Vinícius, Christoph Deil, Kyle Barbary, Tom J Wilson, Ivo Busko, Axel Donath, Hans Moritz Günther, Mihai Cara, krachyon, Simon Conseil, Azalee Bostroem, Michael Droettboom, E. M. Bray, P. L. Lim, Lars Andersen Bratholm, Geert Barentsen, Matt Craig, Shivangee Rathi, Sergio Pascual, Gabriel Perren, Iskren Y. Georgiev, Miguel de Val-Borro, Wolfgang Kerzendorf, Yoonsoo P. Bach, Bruno Quint, and Harrison Souchereau. Astropy/photutils: 1.2.0. September 2021. URL: <https://doi.org/10.5281/zenodo.5525286>, doi:10.5281/zenodo.5525286.
- [9] Bradley W. Carroll and Dale A. Ostlie. *An Introduction to Modern Astrophysics*. Pearson Addison-Wesley, 2nd (international) edition, 2007.
- [10] Ryan R. Curtin, Marcus Edel, Mikhail Lozhnikov, Yannis Mentekidis, Sumedh Ghaisas, and Shangtong Zhang. Mlpack 3: a fast, flexible machine learning library. *Journal of Open Source Software*, 3(26):726, 2018. URL: <https://doi.org/10.21105/joss.00726>, doi:10.21105/joss.00726.
- [11] Luc Devroye. *Non-Uniform Random Variate Generation*(originally published with. Springer-Verlag, 1986. URL: <http://cg.scs.carleton.ca/~luc/rnbookindex.html>.
- [12] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, KDD'96, 226–231. AAAI Press, 1996.

- [13] R.P. Feynman, R.B. Leighton, M. Sands, and EM Hafner. *The Feynman Lectures on Physics; Vol. I.* Volume 33. AAPT, 1965.
- [14] M. et al Galassi. Gnu scientific library reference manual. 2018. URL: <https://www.gnu.org/software/gsl/>.
- [15] James E. Gentle. *Matrix Algebra: Theory, Computations, and Applications in Statistics*. Springer Publishing Company, Incorporated, 1st edition, 2007. ISBN 0387708723.
- [16] Michael Y Grudi, Dávid Guszejnov, Philip F Hopkins, Astrid Lamberts, Michael Boylan-Kolchin, Norman Murray, and Denise Schmitz. From the top down and back up again: star cluster structure from hierarchical star formation. *Monthly Notices of the Royal Astronomical Society*, 481(1):688–702, 08 2018. URL: <https://doi.org/10.1093/mnras/sty2303>, arXiv:<https://academic.oup.com/mnras/article-pdf/481/1/688/25692710/sty2303.pdf>, doi:10.1093/mnras/sty2303.
- [17] S. Harfst, S. Portegies Zwart, and A. Stolte. Reconstructing the arches cluster - i. constraining the initial conditions. *Monthly Notices of the Royal Astronomical Society*, 409(2):628–638, Sep 2010. URL: <http://dx.doi.org/10.1111/j.1365-2966.2010.17326.x>, doi:10.1111/j.1365-2966.2010.17326.x.
- [18] Pavel Kroupa. The initial mass function and its variation. *ASP Conf. Ser.*, 285:86, 2002. arXiv:[astro-ph/0102155](https://arxiv.org/abs/astro-ph/0102155).
- [19] Pavel Kroupa and Carsten Weidner. Galactic-field initial mass functions of massive stars. *The Astrophysical Journal*, 598(2):1076–1078, dec 2003. URL: <https://doi.org/10.1086%2F379105>, doi:10.1086/379105.
- [20] Andrea Kunder, Andreas Koch, R. Michael Rich, Roberto de Propris, Christian D. Howard, Scott A. Stubbs, Christian I. Johnson, Juntai Shen, Yougang Wang, Annie C. Robin, John Kormendy, Mario Soto, Peter Frinchaboy, David B. Reitzel, HongSheng Zhao, and Livia Origlia. THE BULGE RADIAL VELOCITY ASSAY (BRAVA). II. COMPLETE SAMPLE AND DATA RELEASE. *The Astronomical Journal*, 143(3):57, feb 2012. URL: [https://doi.org/10.1088%2F0004-6256%2F143%2F57](https://doi.org/10.1088%2F0004-6256%2F143%2F3%2F57), doi:10.1088/0004-6256/143/3/57.
- [21] Byung-Uk Lee. *Stereo Matching of Skull Landmarks*. PhD thesis, Stanford University, Stanford, CA, USA, 1991.
- [22] Kieran Leschinski, Hugo Buddelmeijer, Oliver Czoske, Miguel Verdugo, Gijs Verdoes-Kleijn, and Werner W. Zeilinger. Scopesim: a flexible general purpose astronomical instrument data simulation framework in python. *Software and Cyberinfrastructure for Astronomy VI*, Dec 2020. URL: <http://dx.doi.org/10.1117/12.2559784>, doi:10.1117/12.2559784.
- [23] M. Marks and P. Kroupa. Inverse dynamical population synthesis. *Astronomy & Astrophysics*, 543:A8, Jun 2012. URL: <http://dx.doi.org/10.1051/0004-6361/201118231>, doi:10.1051/0004-6361/201118231.
- [24] D. J. Marshall, A. C. Robin, C. Reylé, M. Schultheis, and S. Picaud. Modelling the galactic interstellar extinction distribution in three dimensions. *Astronomy & Astrophysics*, 453(2):635–651, Jun 2006. URL: <http://dx.doi.org/10.1051/0004-6361:20053842>, doi:10.1051/0004-6361:20053842.
- [25] Paul J. McMillan. The mass distribution and gravitational potential of the milky way. *Monthly Notices of the Royal Astronomical Society*, 465(1):76–94, Oct 2016. URL: <http://dx.doi.org/10.1093/mnras/stw2759>, doi:10.1093/mnras/stw2759.
- [26] Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White. The structure of cold dark matter halos. *The Astrophysical Journal*, 462:563, May 1996. URL: <http://dx.doi.org/10.1086/177173>, doi:10.1086/177173.
- [27] Sotirios E. Notaris. Gauss-kronrod quadrature formulae - a survey of fifty years of research. *Electron. Trans. Numer. Anal.*, 45:371–404, 01 2016.
- [28] Alicia Porras, I. Cruz-González, and Luis Salas. Imf in perseus arm yscs. *ASP Conference Series*, 287:98–103, 01 2003.
- [29] Simon F. Portegies Zwart, Stephen L.W. McMillan, and Mark Gieles. Young massive star clusters. *Annual Review of Astronomy and Astrophysics*, 48(1):431–493, Aug 2010. URL: <http://dx.doi.org/10.1146/annurev-astro-081309-130834>, doi:10.1146/annurev-astro-081309-130834.

- [30] Adrian M. Price-Whelan, David W. Hogg, Kathryn V. Johnston, and David Hendel. Inferring the gravitational potential of the milky way with a few precisely measured stars. *The Astrophysical Journal*, 794(1):4, Sep 2014. URL: <http://dx.doi.org/10.1088/0004-637X/794/1/4>, doi:10.1088/0004-637x/794/1/4.
- [31] A. C. Robin, D. J. Marshall, M. Schultheis, and C. Reylé. Stellar populations in the milky way bulge region: towards solving the galactic bulge and bar shapes using 2mass data. *Astronomy & Astrophysics*, 538:A106, Feb 2012. URL: <http://dx.doi.org/10.1051/0004-6361/201116512>, doi:10.1051/0004-6361/201116512.
- [32] A. Saro, S. Borgani, L. Tornatore, K. Dolag, G. Murante, A. Biviano, F. Calura, and S. Charlot. Properties of the galaxy population in hydrodynamical simulations of clusters. *Monthly Notices of the Royal Astronomical Society*, 373(1):397–410, Nov 2006. URL: <http://dx.doi.org/10.1111/j.1365-2966.2006.11046.x>, doi:10.1111/j.1365-2966.2006.11046.x.
- [33] P. Kenneth. Seidelmann, Great Britain., and United States Naval Observatory. *Explanatory supplement to the astronomical almanac / prepared by the Nautical Almanac Office, U.S. Naval Observatory ; with contributions from H.M. Nautical Almanac Office, Royal Greenwich Observatory ... [et al.] ; edited by P. Kenneth Seidelmann*. University Science Books Mill Valley, Calif, [rev. ed.]. edition, 1992. ISBN 0935702687.
- [34] John Vince. *Mathematics for Computer Graphics*. Springer, 01 2006. ISBN 978-1-84628-034-4. doi:10.1007/1-84628-283-7.
- [35] D. L. Walker, S. N. Longmore, N. Bastian, J. M. D. Kruijssen, J. M. Rathborne, R. Galván-Madrid, and H. B. Liu. Comparing young massive clusters and their progenitor clouds in the Milky Way. *Monthly Notices of the Royal Astronomical Society*, 457(4):4536–4545, 02 2016. URL: <https://doi.org/10.1093/mnras/stw313>, arXiv:<https://academic.oup.com/mnras/article-pdf/457/4/4536/18517186/stw313.pdf>, doi:10.1093/mnras/stw313.
- [36] Yougang Wang, Hongsheng Zhao, Shude Mao, and R. M. Rich. A new model for the Milky Way bar. *Monthly Notices of the Royal Astronomical Society*, 427(2):1429–1440, 12 2012. URL: <https://doi.org/10.1111/j.1365-2966.2012.22063.x>, arXiv:<https://academic.oup.com/mnras/article-pdf/427/2/1429/2984088/427-2-1429.pdf>, doi:10.1111/j.1365-2966.2012.22063.x.
- [37] Denis Yurin and Volker Springel. An iterative method for the construction of N-body galaxy models in collisionless equilibrium. *Monthly Notices of the Royal Astronomical Society*, 444(1):62–79, 08 2014. URL: <https://doi.org/10.1093/mnras/stu1421>, arXiv:<https://academic.oup.com/mnras/article-pdf/444/1/62/18503149/stu1421.pdf>, doi:10.1093/mnras/stu1421.
- [38] S. J. Aarseth, M. Henon, and R. Wielen. A Comparison of Numerical Methods for the Study of Star Cluster Dynamics. *\aap*, 37(1):183–187, December 1974.
- [39] Barros, D. A., Lépine, J. R. D., and Dias, W. S. Models for the 3d axisymmetric gravitational potential of the milky way galaxy - a detailed modelling of the galactic disk. *A&A*, 593:A108, 2016. URL: <https://doi.org/10.1051/0004-6361/201527535>, doi:10.1051/0004-6361/201527535.
- [40] Gilles Chabrier. Galactic Stellar and Substellar Initial Mass Function. *\pasp*, 115(809):763–795, July 2003. arXiv:[astro-ph/0304382](https://arxiv.org/abs/astro-ph/0304382), doi:10.1086/376392.
- [41] Rebecca A. W. Elson, S. Michael Fall, and Kenneth C. Freeman. The Structure of Young Star Clusters in the Large Magellanic Cloud. *\apj*, 323:54, December 1987. doi:10.1086/165807.
- [42] Espinoza, P., Selman, F. J., and Melnick, J. The massive star initial mass function of the arches cluster \*\*\*. *A&A*, 501(2):563–583, 2009. URL: <https://doi.org/10.1051/0004-6361/20078597>, doi:10.1051/0004-6361/20078597.
- [43] Sanghamitra Goswami, Stefan Umbreit, Matt Bierbaum, and Frederic A. Rasio. Formation of Massive Black Holes in Dense Star Clusters. II. Initial Mass Function and Primordial Mass Segregation. *\apj*, 752(1):43, June 2012. arXiv:[1105.5884](https://arxiv.org/abs/1105.5884), doi:10.1088/0004-637X/752/1/43.
- [44] Lars Hernquist. An Analytical Model for Spherical Galaxies and Bulges. *\apj*, 356:359, June 1990. doi:10.1086/168845.
- [45] Lars Hernquist. N-Body Realizations of Compound Galaxies. *\apjs*, 86:389, June 1993. doi:10.1086/191784.

- [46] Ivan R. King. Density Data and Emission Measure for a Model of the Coma Cluster. *\apjl* , 174:L123, June 1972. doi:10.1086/180963.
- [47] A. H. W. Kuepper, Th. Maschberger, P. Kroupa, and H. Baumgardt. McLuster: A Tool to Make a Star Cluster. July 2011. arXiv:1107.015.
- [48] Paul J. McMillan. GalPot: Galaxy potential code. November 2016. arXiv:1611.006.
- [49] M. Miyamoto and R. Nagai. Three-dimensional models for the distribution of mass in galaxies. *\pasj* , 27:533–543, January 1975.
- [50] Mark J. Pecont and Eric E. Mamajek. Intrinsic Colors, Temperatures, and Bolometric Corrections of Pre-main-sequence Stars. *\apjs* , 208(1):9, September 2013. arXiv:1307.2657, doi:10.1088/0067-0049/208/1/9.
- [51] A. J. Pickles. A Stellar Spectral Flux Library: 1150–25000 Å. *\pasp* , 110(749):863–878, July 1998. doi:10.1086/316197.
- [52] R. Piessens, E. de Doncker-Kapenga, and C. W. Ueberhuber. *Quadpack : a subroutine package for automatic integration*. Springer series in computational mathematics ; 1. Springer, Berlin [u.a.], 1983. ISBN 3540125531.
- [53] H. C. Plummer. On the problem of distribution in globular star clusters. *\mnras* , 71:460–470, March 1911. doi:10.1093/mnras/71.5.460.
- [54] L. Portinari. The Role of the IMF in Modelling Clusters of Galaxies. In A. Vallenari, R. Tantalo, L. Portinari, and A. Moretti, editors, *From Stars to Galaxies: Building the Pieces to Build Up the Universe*, volume 374 of Astronomical Society of the Pacific Conference Series, 131. December 2007.
- [55] Edwin E. Salpeter. The Luminosity Function and Stellar Evolution. *\apj* , 121:161, January 1955. doi:10.1086/145971.
- [56] Ralph Schönrich, James Binney, and Walter Dehnen. Local kinematics and the local standard of rest. *\mnras* , 403(4):1829–1833, April 2010. arXiv:0912.3693, doi:10.1111/j.1365-2966.2010.16253.x.
- [57] A. Toomre. On the gravitational stability of a disk of stars. *\apj* , 139:1217–1238, May 1964. doi:10.1086/147861.
- [58] P. C. van der Kruit. The three-dimensional distribution of light and mass in disks of spiral galaxies. *\aap* , 192:117–127, March 1988.
- [59] P. C. van der Kruit and L. Searle. Surface photometry of edge-on spiral galaxies. I - A model for the three-dimensional distribution of light in galactic disks. *\aap* , 95:105–115, February 1981.