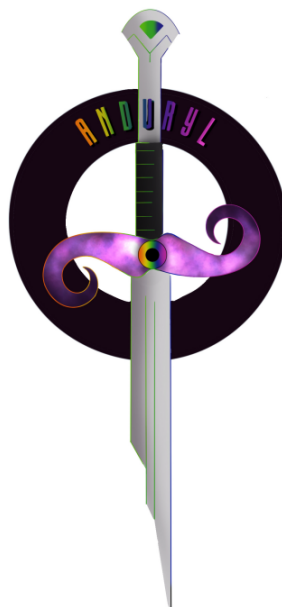# A Python-based Bayesian code to perform photometric decompositions of galaxies

Carlos Marrero de la Rosa

MASTER'S DEGREE IN ASTROPHYSICS
FINAL MASTER'S DEGREE PROYECT
UNIVERSIDAD DE LA LAGUNA
ACADEMIC COURSE: 2022-2023

Tutors:

Jairo Méndez-Abreu , Andrés Asensio Ramos

ANDURYL logo by Alexia Arbelo Vega

"It is absolutely impossible to write this science with the precision with which I understand it in my heart. However, though words are insufficient, the principles must be self-evident".

Miyamoto Musashi, The Book of Five Rings

Acknowledgments

Resumen

La clasificación de entidades u objetos presentes en la naturaleza es uno de los primeros pasos en un análisis científico, esta acción lleva a la búsqueda de patrones comunes que permitan entender los fundamentos físicos que se encuentran detrás de dicha clasificación. En el contexto de la astrofísica, y en concreto de las galaxias, la clasificación de las mismas es una necesidad de cara a entender su contexto evolutivo y su formación. La clasificación morfológica de las galaxias se basa en la búsqueda de las estructuras que las componen: el bulbo (la parte central), el disco (zona con simetría axial en torno a un plano), la barra (un objeto que se mueve como un sólido rígido en torno al centro galáctico) y los brazos espirales (zonas de alta densidad de estrellas, gas y polvo en el disco). El primer tipo de clasificación fue la visual, que se basaba en detectar a ojo estas estructuras; luego, con el paso del tiempo, han ido apareciendo nuevos tipos de clasificación como el CAS (Concentración Asimetría y Suavidad), que se basa en la medición de estos parámetros, y las descomposiciones fotométricas, que se basan en recrear las estructuras anteriores con distintas formas funcionales como los perfiles de Sérsic (Sérsic 1963), exponenciales (Freeman 1970) y de Ferrers (Ferrers 1870), teniendo cada uno de ellos un numero determinado de parámetros libres.

Es en este ultimo tipo de clasificación en el que se centra este trabajo, ya que actualmente hay una variedad de códigos que permiten hacer descomposiciones fotométricas en 2D como GALFIT (Peng et al. 2010), GASP2D (Méndez-Abreu et al. 2008) o IMFIT (Erwin 2015). Pero estos emplean métodos de minimización de mínimos cuadrados, basados en Levenberg–Marquardt, que pueden presentar algunos problemas tales como: soluciones no físicas, complicación a la hora de buscar condiciones iniciales, inversión de componentes, etc. Para tratar de solventar todas estas dificultades, el objetivo de este trabajo es construir un nuevo código que se cimiente en la estadística Bayesiana, ANDURYL, que permita solventar estos problemas. Una vez desarrollado el código, se aplicó a un caso científico analizando las galaxias del cúmulo de galaxias Abell 2142; en el contexto de realizar un trabajo preliminar del proyecto WEAVE Nearby Cluster Survey (WEAVE 2023), que analizará poblaciones de galaxias enanas en cúmulos de galaxias, al fin de ahondar en el conocimiento acerca del comportamiento, evolución y formación de estos sistemas.

Los datos para llevar a cabo el trabajo han sido extraídos de SDSS (Sloan Digital Sky Survey); un proyecto cuyo objetivo es mapear la mayor cantidad de cielo posible usando datos proporcionados por imágenes y espectros de millones de objetos. Para ello se ha hecho un cruce entre la base de datos de SDSS y la información espectroscópica de Liu et al. (2018), de manera que podemos descargar aquellas imágenes de galaxias que fueron clasificados como miembros del cúmulo por Liu et al. (2018). SDSS emplea cinco filtros fotométricos que funcionan en distintos intervalos de longitudes de onda; en concreto para este trabajo se ha empleado la banda "i", que opera en una longitud de onda efectiva de 7481 Å.

A la hora de analizar las imágenes de las distintas galaxias de SDSS se han creado máscaras para no incluir los objetos que se encuentren en la imagen, a excepción de la galaxia de estudio, empleando para ello el código SExtractor[1]. Entre otros efectos ANDURYL, tiene en cuenta que al crear los modelos a partir de los perfiles fotométricos hay que convolucionar el modelo obtenido con la Point Spread Function (PSF), para poder comparar el modelo con la observación.

La versión actual de ANDURYL incluye un perfil de Sérsic y un perfil Sérsic+exponencial. Teniendo esto en cuenta, para poder ajustar los parámetros libres se han estudiado dos métodos Bayesianos diferentes, el Hamiltonian Monte Carlo (HMC; Duane et al. 1987) y Nested Sampling (Skilling 2004). Ambos métodos se encargan de muestrear el espacio de parámetros, dando una

---

[1]https://www.astromatic.net/software/sextractor/

descripción completa de la distribución de probabilidad posterior de cada parámetro, de donde se puede inferir el valor más probable de los mismos. El primer método se basa en realizar un mapeo del espacio de parámetros empleando una dinámica hamiltoniana. Mientras que el segundo, a partir de cambios de variable, analiza el hipervolumen del espacio de parámetros bajo una curva unidimensional.

Usando ANDURYL hemos comparado ambos métodos Bayesianos, encontrándose que el Nested Sampling proporcionaba un mejor muestreo del espacio de parámetros, razón por la cual se eligió como método de inferencia principal del código. Se emplearon 500 muestras de simulaciones de galaxias creadas a partir de un perfil de Sérsic y 500 creadas a partir de un Sérsic+exponencial para testear la robustez del código. Se analizó ambas muestras con los dos perfiles disponibles, encontrándose que los resultados mostraban que el código funcionaba correctamente.

Se procedió entonces al análisis de la muestra de galaxias de A2142, en donde se midieron distintas relaciones de escala como la Faber-Jackson (Faber & Jackson 1976), que relaciona la dispersión de velocidad de las estrellas con la masa de las galaxias, y la relación masa-tamaño. En ambos casos se encontró que las galaxias analizadas seguían las tendencias derivadas para otras muestras de la literatura, dentro de los márgenes de error.

A partir de los resultados obtenidos se puede decir que el desarrollo del código ha resultado satisfactorio y que se han cumplido los objetivos propuestos; teniéndose una nueva herramienta para poder explorar la formación, evolución y comportamiento de las galaxias en distintos entornos.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Galaxy classification methods

The classification of objects, animals or any entity helps us to organize, order, and sequence the information. From here, common characteristics can be found in the classified entities, patterns, that help us to understand, among other things, the physical foundation of that classification.

The objects of interest of this work are galaxies, which are systems composed of millions of stars, gas and dust, gravitationally bound and embedded in a dark matter halo. They have a very varied and rich morphology, but with elements in common. The bulge, is the central group of stars with spheroidal symmetry; is usually composed of old stars (~ 10 Gyr), although it depends on the type of bulge. The bar, is a central bar-shaped structure composed of stars that moves as a rigid body around the galaxy center. There is also the disc, which is a part with axial symmetry around the rotation axis. And there are the spiral arms, which are structures with a high density of stars, gas and dust that agglutinate material in the disc in a spiral shape. An example of a galaxy containing all of them can be seen in **Fig. 1**. All of these structures can be found in galaxies, but some may or may not be present. And from here arises the first method of classification.
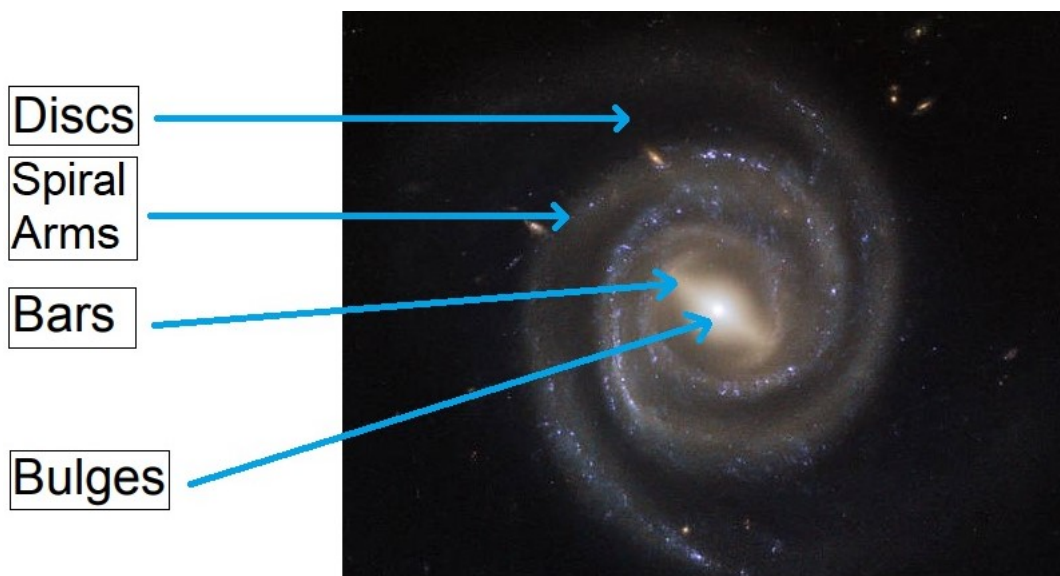
Fig. 1: Image of a galaxy named UGC 6093 containing bulge, disc, bar, and spiral arms[1].

In 1926, Edwin Hubble presented the morphological classification scheme known as Hubble's sequence (Hubble 1926). This system categorize galaxies based on their visual appearance, as

---

[1] https://www.nasa.gov/image-feature/goddard/2018/hubbles-barred-and-booming-spiral-galaxy

a way of trying to understand their formation, evolution, and intrinsic properties. At the core of Hubble's classification is the recognition that galaxies come in various shapes and sizes. The system classifies galaxies into three main categories: ellipticals, spirals, and irregulars. Each category further branches out into subtypes, forming a comprehensive framework to describe the morphological characteristics of galaxies, as can be seen in **Fig. 2**.
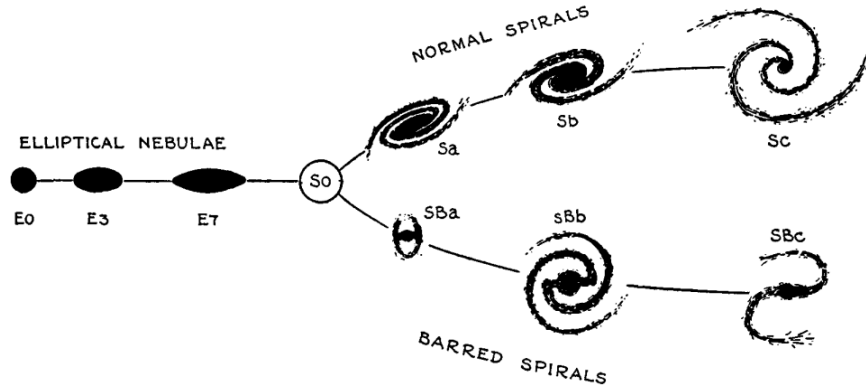
Fig. 2: Diagram of the Hubble sequence of galaxies, taken from Hubble (1936).

Elliptical galaxies, denoted by the letter "E", are typically round or elongated without prominent structure. They range from E0 (more circular) to E7 (more elongated). Lenticular galaxies (S0) have a disc structure with a central concentration of stars (bulge). Spiral galaxies, labeled with the letters "S" or "SB" (for barred spirals), showcase a central bulge surrounded by distinct spiral arms. The arms can be loosely wound (Sc), tightly wound (Sa), or intermediate (Sb). Irregular galaxies, denoted by the letter "I," defy the well-defined structures seen in elliptical and spiral galaxies. They exhibit a diverse range of shapes, often irregular or peculiar, lacking any clear symmetrical pattern. Notably, the 1936 version of Hubble's diagram excluded irregular galaxies, as their lack of rotational symmetry rendered their placement uncertain at that time.

With the passage of time, this type of morphological classification has been further developed, with works such as de Vaucouleurs (1959) or van den Bergh (1976). But these classification methods, based purely on qualitative visual properties, suffers from several problems: it is superficial since it does not depend on quantifiable physical properties, subjective since it depends on the person who will carry out the classification, it depends on the orientation of the galaxy, the visual classification of the galaxy varies depending on the spectral range in which it has been taken the image, etc.

As a result of this, new ways of classifying galaxies have been created over time, such as the CAS (Concentration, Asymmetry, Smoothness) parametric classification (Abraham et al. 1994; Conselice 2003). This method is based on the measurement of these three parameters when classifying, and each of them is a number that quantifies some aspect of the galaxy's structure. Concentration measures the central clump of light on the object, asymmetry measures how symmetrical the object is, and smoothness measures the fraction of light that can be found on small-scale objects. Since this method has not been used in this work, it will not be discussed in greater detail. We only add that recent advances in Machine Learning and Deep Learning techniques have led to a consistent automation of this method (Huertas-Company et al. 2008; Domínguez Sánchez et al. 2018).

There is another type of classification of interest, and on which this work is based, the classification by photometric decompositions. It is based on parametric techniques that use functional forms to describe the surface brightness distribution of the different structures that make up or can make up a galaxy.

The analysis of the distribution of light in galaxies is complicated by the wide variety of sizes, shapes, and structures. However, it provides a large amount of information on, for example, the evolution of the galaxy disc, the cosmic evolution of galactic morphology, the central structure of early galaxies, etc. When performing brightness profile modeling, residuals can be calculated as the difference between the model and the real image, and this will provide information about unfit components of the galaxy, such as unresolved nuclear components, spiral arms, etc.

To model the surface brightness of these components, different types of profiles are used, such as the Sérsic profile (Sérsic 1963), which is usually used for the bulge part, the exponential profile, that is usually used for the disc part (Freeman 1970), and the Ferrers profile (Ferrers 1870), which is used for the bars.

1.2 Photometric decompositions algorithms

To model the different components of the galaxies and obtain the parameters that define them, in order to infer scale relation, among other things. This can be achieved from the photometric decompositions described above.

Initially, the photometric decompositions were carried out in one dimension (Freeman 1970), adjusting the surface brightness profile simultaneously for the different components; this gave rise to non-linear fits with a multitude of free parameters, where a change in the initial conditions led to very different fits (Gao & Ho 2017). Nowadays, it is possible to carry out photometric fits in two dimensions, fitting directly all pixels.

For this, there are numerous codes that perform 2D photometric decompositions, such as GASP2D (Méndez-Abreu et al. 2008), GALFIT (Peng et al. 2010) or IMFIT (Erwin 2015), among others. All these codes use minimization methods usually a least-squares fitting algorithm that uses the Levenberg-Marquardt technique to find the best solutions when fitting the galaxies to the different models. Moreover they suffer from serious drawbacks (Lange et al. 2016): (I) local minima trapping; (II) un-realistic solutions; (III) reversal of components (Allen et al. 2006); (IV) indecisiveness as to which model to use; and (V) bad representation of final errors. To try to overcome these inconveniences, photometric decompositions can be made within the framework of Bayesian statistics, being able to solve the problem of entrapment in local minima by the exploration of parameter space, prevent unreal solutions and the inversion of components providing prior information, and one of the most important aspects, the posterior distribution resulting from performing this fit allows an estimation of the uncertainty in the parameters, including covariances, something that is more difficult and expensive to achieve in conventional algorithms.

In this sense, a code to perform photometric decompositions that uses Bayesian statistics represents a considerable improvement; which would lead to a better analysis of the photometric parameters of galaxies and, therefore, to a better understanding of the physical processes that occur behind their formation and evolution. In Argyle et al. (2018), a Bayesian code called PHI (https://github.com/SEDMORPH/PHI/) was developed with the purpose of obtaining a robust morphological analysis of galaxies. However, the code is developed in IDL programming language,

and this imposes some limitations when it comes to sharing it with the community, among other things.

1.3 Scaling relations and galaxy clusters

The currently concordant cosmological model, the ΛCMD, is quite successful in reproducing observations of large-scale structures in the universe. The dark matter halos convert their potential energy into kinetic energy, virializing; this allows galaxies to form as baryonic matter cools down. The understanding of the evolution of light structures is necessary to explain the relations between dark matter and baryonic structures. Furthermore, the physical properties of galaxies such as their gas content, stellar age, or star formation rate are correlated with galactic morphology (Clauwens et al. 2018; Tacchella et al. 2016). So, a precise quantitative description of the galactic structures is fundamental to be able to advance in the understanding of galactic evolution. There are a number of correlations between some properties of galaxies, such as: the effective radius, the average effective surface brightness and the central velocity dispersion of galaxies. If they could be represented in 3D, they would be distributed along a plane, which is known as the fundamental plane (Djorgovski & Davis 1987).

From here, several other relations can be obtained that are nothing more than a projection of the fundamental plane. The projection of the velocity dispersion and the effective surface brightness gives rise to the Faber-Jackson relation (Faber & Jackson 1976). This relation can be inferred from the Virial Theorem, $2K + U = 0$ ($K \equiv$ kinetic energy, $U \equiv$ potential energy). Therefore, those galaxies that are virialized systems will follow the Faber-Jackson relationship. This trend represents the "cooling diagram" on which the theories of galaxy formation are based (Costantin et al. 2020). Expressing this relation in terms of the luminosity L, and the velocity dispersion σ, it can be state that

$$L \propto \sigma^4. \tag{1}$$

There is also a relation between the mass-size of the bulges and discs of the galaxy, although this is a bit more complex than the previous one, since spheroidal and discy galaxies show different trends (Méndez-Abreu et al. 2021). However, the study of this relation helps to understand the evolution of galaxies in different mass ranges.

In this framework, it could be said that knowledge about galactic evolution is based on the observation and modeling of different systems with different characteristics. There is one particular object in the universe, galaxy clusters, which are the highest density regions known, and mark the largest and most massive virialized features (Kravtsov & Borgani 2012). As they are regions of high density, they serve to study the interactions between the objects that compose them and their neighborhoods.

Given the current theoretical framework, galaxy clusters are formed from mergers of smaller dark matter halos that are virialized, and their study would allow us to see how the environment influences the shape of galaxies. In addition, early-type galaxies are more common in these clusters than in the general field, which makes them excellent regions to investigate the origin and evolution of this type of galaxies and what phenomena may lead to their formation, and the investigation of the physical processes that occur and that lead to the evolution of galaxies in these environments is a current reason for discussion (Boselli & Gavazzi 2014).

One of the associated problems is the evolution of dwarf galaxies, those that have between 10^7 and 10^9 solar masses. Despite the fact that these objects are the most numerous types of galaxies

in the local system, their formation and evolution are still not fully understood. This ignorance becomes clear when, despite being the largest group in the local universe, the numbers still do not agree with those obtained in the models (Moore et al. 1999), among other problems. Then it is necessary to have information and measurements of dwarf galaxies to be able to solve these problems, and this is where the WEAVE Nearby Cluster Survey comes in.

1.4 WEAVE Spectrograph Project

The WEAVE instrument (William Herschel Telescope Enhanced Area Velocity Explorer, WEAVE (2023)) is a new multi-object spectrograph that will be located on the 4.2 m William Herschel Telescope, which is located at the Roque de los Muchachos Observatory, on La Palma in the Canary Islands. With its new technology, it will allow obtaining up to 1000 simultaneous spectra, where the fiber-fed spectrograph comprises two arms, one optimized for blue wavelengths and the other one for red wavelengths. It will be able to operate in two resolutions of 5000 and 20000 and it will be located at the primary focus of the WHT. In February 2023 the integral field unit of WEAVE was commissioned, in this same year the verification observations should begin.

The WEAVE Nearby Cluster Survey will aim to study, in particular, low-mass galaxies, with the purpose of shedding light on the different aspects of their evolution and formation. To understand these physical processes it will be necessary to have an appropriate data set, which is what is intended to be obtained with this survey. Among others, the survey has the following scientific objectives: to study the orbital structure of low-mass galaxies in clusters, stellar populations, metallicities, and the star formation history of cluster dwarf, scaling relations of dwarf galaxies, etc. In order to cover this range, a surface brightness limit will be used, in the r band, of $\mu_{\mathrm{r}} = 20$ mag arcsec^{-2}. This survey will observe clusters of galaxies that are in an environment of 150 Mpc, which emit X-ray flux. One of the targets of these observations will be the galaxy cluster Abell A2142, an X-ray luminous galaxy cluster located near the constellation Corona Borealis, at ~ 381 Mpc with a redshift of ~ 0.09 (Kopylova & Kopylov 2022).

1.5 Objectives

In line with the Bayesian framework, one of the initial objectives of this work was to translate PHI from IDL to Python, improving those aspects that could have become obsolete. But, finally, it was decided to create a completely new code from scratch, ANDURYL[2] (A bayesiaN Decomposition code for Use in photometRY Labors); which would use a modern Bayesian method as the core of its photometric fitting engine. Once the code is created, it was applied to mock and real samples of galaxies, to test its reliability and analyze the results in search of possible errors. The sample of real galaxies chosen to carry out the first science with ANDURYL was the cluster of galaxies A2142. The galaxies of this cluster will be the subject of future analysis with the new instrument WEAVE, the current study is design as a preparatory work to be applied to the full WEAVE Cluster Survey.

Once the real galaxies have been analyzed, different scale relations have been measured, such as Faber-Jackson and mass-size, in order to verify that the results produced by ANDURYL for real galaxies follow these relations, which would support the results themselves. So the objectives of this project can be summarized as follows:

1) Creation of the ANDURYL code with different model options available.

2) Application of the code to a sample of mock galaxies that contain only bulges and bulge+disc galaxies, and the comparison of the parameters with which the samples have been created and those returned by ANDURYL.

3) Application of the code to a sample of real galaxies, to carry out the first science with ANDURYL and prepare the tools to be ready for the WEAVE Cluster Survey.

---

[2] https://github.com/CarlosMDLR/ANDURYL

Chapter 2

Data

Two different types of data have been used, depending on their use for the testing of the code or the scientific analysis. Regarding the former, 500 bulge mock galaxies were used, created from a Sérsic profile, and another 500 bulge-disc mock galaxies, created from a combination of a Sérsic profile and an exponential one. These mock galaxies were created using the GASP2D code (Méndez-Abreu et al. 2008), having been used in other works such as Méndez-Abreu et al. (2017). It should be noted that both, ANDURYL and GASP2D are codes that use a different methodology, and the functions used for the photometric profiles are also different.

For the scientific analysis, data from 234 galaxies belonging to the cluster A2142, have been used. To obtain the catalog of real galaxies that has been used, a crossmatch was made between the Sloan Digital Sky Survey (SDSS) catalog for this cluster and the Liu et al. (2018) catalog, imposing as conditions that the galaxies were classified as members by Liu and that the velocity dispersion was measured in the SDSS spectra database.

For both types of mock galaxies it was not necessary to do a pre-processing of the data, all of them had the same noise and the same Point Spread Function (PSF) parameters incorporated, which was modeled from a Moffat function. However, for the A2142 cluster galaxies it was necessary to carry out different actions to obtain: (i) the images of each galaxy, (ii) the measurement of the corresponding PSF, and (iii) the creation of masks for each case.

2.1 Sloan Digital Sky Survey

SDSS is a project whose goal is to map as much of the sky as possible, providing enough information for cosmological scale studies, as well as other astrophysics studies. This project has measured, of the order of millions of spectra and images in a range of wavelength bands, from 3551Å to 8931Å, a range observed using five filters, detailed below.

To carry out these observations SDSS has a series of telescopes and instruments, beginning with the Sloan Foundation Telescope, a 2.5 m telescope at Apache Point Observatory with a 1.08 m secondary mirror, more details of which can be seen at Gunn et al. (2006). For the fourth phase of SDSS, observations of the southern hemisphere from the Irénée du Pont Telescope at Las Campanas Observatory, which is another 2.5 m telescope, were included. Details of the telescope are described in detail in Bowen & Vaughan (1973).

Currently SDSS also has several spectrographs, like BOSS. They are two identical spectrographs where each one uses two cameras, red and blue, with a dichroic splitting the light at roughly 6000 Å and a full wavelength range from 3600 to 10400 Å. The high resolution APOGEE spectrograph is a multi-object, with a spectral resolution around R=22500.

SDSS has had various phases during its years of operation, and for each phase they publish a certain number of data releases. They are cumulative, that is, each new one that comes out includes all the sky coverage of prior releases. Each one incorporates new techniques, observation methods, etc. Therefore, we proceed to give details of Data Release 16 (DR16; Ahumada et al. 2020), which is the one that has been used in this work.

DR16 includes six types of data: images, optical spectra, infrared spectra, Integral Field Unit (IFU) spectra, stellar library spectra, and catalog data. The ones of interest for this work are mainly the images. All of them are accessible and downloadable from their website[1]. It includes SDSS data taken until August 2018, and encompasses more than one-third of the entire celestial sphere. Specifically, the number of unique, primary sources amounts to 469,053,874, of which 208,478,448 are galaxies.

The camera built for collecting imaging data for SDSS consists of two arrays, one that uses 30 CCDs of 2048x2048 pixels and another array of 24 CCDs of 400x2048, both with the same pixel size, with the pixel scale being 0.396 [arcsec/px] (Gunn et al. 1998).

In order to handle the data generated by the imaging camera, the SDSS data processing factory employed a series of interconnected processes. These processes, functioning as pipelines, undertook the tasks of data processing and calibration. Their primary objective is to eliminate instrumental signatures and refine the data to generate a range of valuable data products. Among the outcomes are enhanced images, purged of any artifacts caused by the instruments, a photometric solution for nocturnal conditions, and a comprehensive catalog of the diverse objects discovered within the dataset.

After this, the Science Archive Server provides the survey images, called "corrected frames", as "frames-*.fits.bz2" files. The units of these images are not given in usual counts, but in nanomaggies per pixel, and that they are already subtracted from the sky brightness. A nanomaggy is 10^{-9} times the flux F of the source relative to the standard source F_0 (which defines the zeropoint of the magnitude scale). SDSS uses five particular filters, called (u, g, r, i and z), each one operates in a determined range of wavelength, where the effective wavelength of each of them can be seen in Table. 1.

Table 1: Effective wavelength and magnitude limit for the different filters used by DR16 of SDSS.

Filters:	u	g	r	i	z
Effective wavelength (Å)	3551	4686	6165	7481	8931
Magnitude limit	22.0	22.2	22.2	21.3	20.5

2.2 Obtaining galaxy maps and PSF synthesis

Of the galaxies resulting from the crossmatch between SDSS and Liu et al. (2018), the data is downloaded from the SDSS "i" filter. Once this is done, the right ascension and declination coordinates of each galaxy are used to find them in their corresponding frame. We use the radius R_{90}, which encompasses 90 % of the Petrosian flux, as a measure to make a rectangular clip around
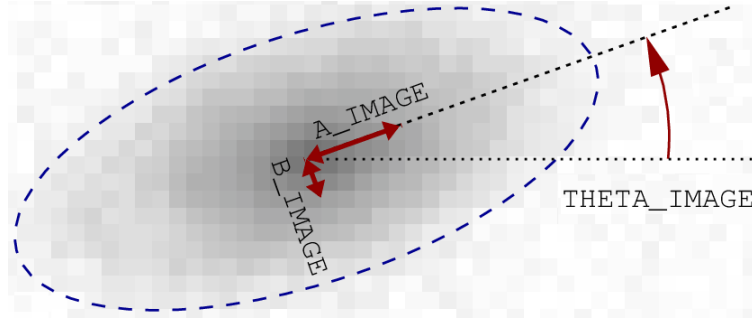
---

[1]https://www.sdss.org/

Fig. 3: Example of an ellipse around an object using SExtractor. [2]

the center of the galaxy of radius $10 \times R_{90}$.

The Petrosian flux (Petrosian 1976), as described in SDSS, for any band is defined as the flux that is within a certain number of times n_p the Petrosian radius r_p. The latter being the point at which the ratio between the local surface brightness in an aperture and the mean surface brightness reaches a certain value. Therefore, the expression for the Petrosian flux can be written as

$$F_p \equiv \int\limits_0^{n_p r_p} 2\pi r' I(r') dr', \tag{2}$$

where $I(r')$ is the azimuthally averaged surface brightness profile.

Now that the images of each galaxy have been downloaded and resized, the PSF must be measured for each field. The PSF is nothing more than the response of an optical system to a point light source, and in observational astrophysics, from relatively large ground-based telescopes, atmospheric turbulence (seeing) is the dominant contribution to the PSF. Using the "pyraf"[3] package, a program was developed to find the stars in each field with a number of counts greater than five times the sky noise. Once the stars were found, each one of them was fitted with a Moffat function (see Eq. (3)), obtaining the Full Width at Half Maximum (FWHM) and the parameter β.

$$f_{\text{Moffat}}(r; \alpha, \beta) = \frac{\beta - 1}{\alpha^2} \left[1 + \left(\frac{r}{\alpha} \right)^2 \right]^{-\beta} \; ; \; \alpha = \frac{\text{FWHM}}{2\sqrt{2^{\beta^{-1}} - 1}}. \tag{3}$$

2.3 Creating masks using SExtractor

When analysing the galaxies from the SDSS fields in which they are found, it is highly possible that stars, or other galaxies, are found within the cutout field of view. It is necessary to mask these other sources and leave only the galaxy of interest, so that the subsequent photometric fit is not compromised. As there are a large number of galaxies to be analyzed, making the masks individually was ruled out; making therefore necessary to create an algorithm that allows the 234 necessary masks to be automatically made. This is where SExtractor[4] comes in, which is a program that builds a catalog of objects from an astronomical image.

With the PSF information obtained in the previous section, the SExtractor configuration files can be created for each galaxy image. In this way, SExtractor returns information about the

---

[2]https://astromatic.github.io/SExtractor/Position.html#basic-shape-parameters-a-b-theta
[3]https://iraf-community.github.io/pyraf.html
[4]https://www.astromatic.net/software/sextractor/

objects in the image. Specifically, the position parameters of each object (x,y) are saved, as well as the ellipticity, the position angle of the semi-major axis (A) with respect to the x-axis of the image, the semi-major and minor axes themselves (A, B), all of them for an ellipse that encloses most of the light of the found object, as can be seen in **Fig. 3**.

With these obtained parameters, ellipses can be generated that mask all the objects found in an image, so mask files can be generated for each galaxy image. The way in which the program that has been designed to build the masks works is that, with the information from the ellipses, these are generated in a mesh with the size of the galaxy image, magnifying the size of the ellipses by a certain value in order to ensure that it covers the desired object. The contour and interior of these ellipses is associated with a null value, while the exterior of these ellipses has the unit value. Therefore, they can be used so that the pixels covered by these ellipses are not taken into account in the subsequent Bayesian treatment. It must be taken into account that the object to be analysed, which is the galaxy located near the center of the image, must not be covered. So, the ellipse that is associated with the galaxy of interest is also assigned with the unit value, and it is superimposed on the mesh. In this way, if there are objects close to the galaxy of interest, the masks of these objects will not affect the surroundings of the galaxy, as can be seen in **Fig. 4**.
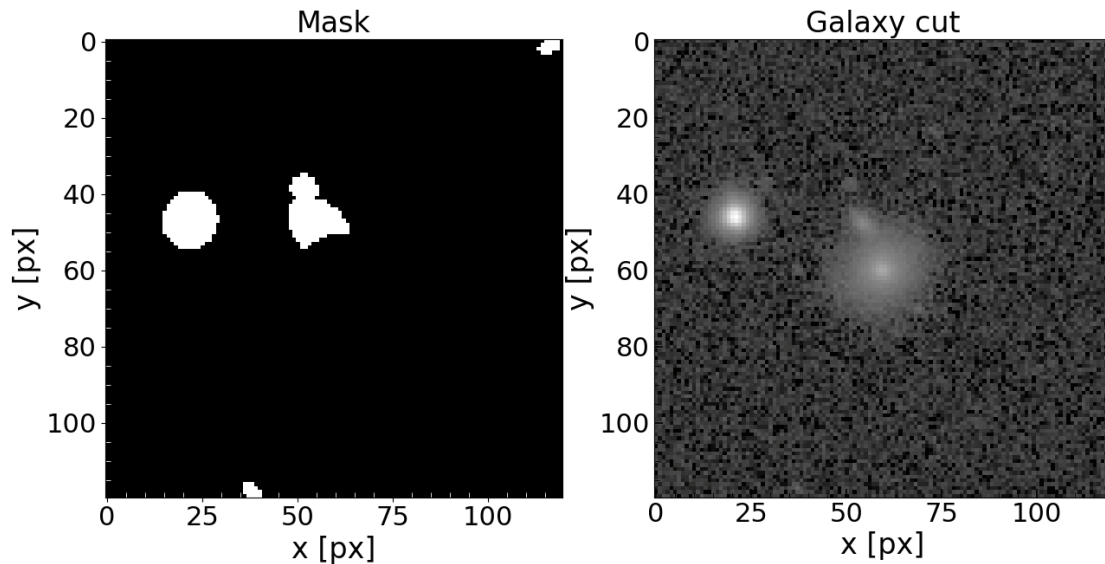

Fig. 4: Example of a mask created for one of the sample galaxies.

Chapter 3

Methodology

This section will deal with the internal functioning of ANDURYL, as well as the photometric decomposition models used by it, and the different Bayesian inference methods that have been used throughout the work. The way in which ANDURYL works can be summarized in a few short steps:

1) The necessary information is passed to it through a configuration file, such as the directories where the images to be analyzed are located, the file with the PSF information for each image, the PSF and surface brightness models to be used, etc.

2) The program reads the configuration file and reads the image files (FITS files), then proceeds to perform the photometric fit following a bulge or a bulge-disc model, as indicated. This generated model is then convolved with the PSF through the Fast Fourier Transform (FFT).

3) Once the model has been created, the likelihood and the posterior functions are calculated through the given Bayesian inference method that has been used, inferring all the parameters of the model that are available. Once this is done, the information of all the parameter space sampling carried out by the method is saved in a FITS file for each galaxy.

4) If it is set as an option, the program will make the plots where the galaxy can be seen together with the created model and its residuals, as well as the triangular diagrams for the posteriors of the free parameters.

3.1 Photometric decomposition models

To model the different components of a galaxy, via photometric decompositions, there are a great variety of functional shapes. In this work two functional shapes has been used, the Sérsic profile and the exponential profile, which give a good account of the bulge of a galaxy and its disc, respectively. However, the program is designed so that other components can be easily incorporated in the future. The Sérsic profile (Sérsic 1963), as it has been incorporated into this work, has the following form

$$I(r) = I_e \; 10^{\left\{ -b_n \left[\left(\frac{r}{R_e} \right)^{n^{-1}} -1 \right] \right\}} \; ; \; b_n = (0.868 \cdot n) - 0.142, \tag{4}$$

being I_e the intensity at radius R_e, which encloses half of the total intensity of the model, b_n is a function that ensures that R_e contains half of the light and depends on n, which is the Sérsic index, which accounts for how concentrated is the profile.

On the other hand, the exponential profile (Freeman 1970) is such that

$$I(r) = I_0 \; \exp\left(-\frac{r}{h} \right), \tag{5}$$

beeing I_0 the central intensity and h the disc scale lenght of the profile.

Apart from all the free parameters that both profiles have, another two parameters must also be taken into account for each model: the ellipticity ε and the angle PA. These two parameters are introduced because the intensity profile is characterized by elliptical concentric isophotes, where the ellipticity is defined as $\varepsilon = 1 - (b/a)$, beeing b and a the semi-minor and semi-major axis of the ellipse, just like in **Fig. 3**. But here the angle, PA, is measured in degrees and counter-clockwise from the vertical axis of the image of the galaxy, unlike in the figure cited.

With these definitions, the expression for the radius can then be written, based on Argyle et al. (2018). First, the reference system is redefined, so that it is centered in the center of the object (x_0, y_0). The angle is also taken as the one between the ordinate axis of the image and the semi-major axis of the analyzed object, as can be seen in **Fig. 5**.
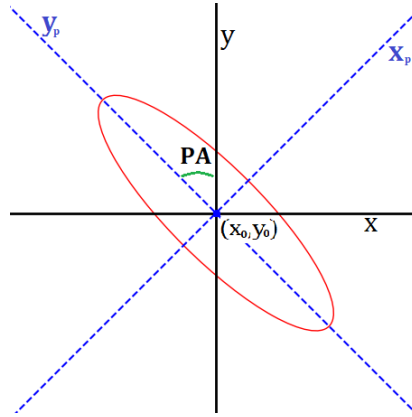
Fig. 5: Illustration of the reference frame centered on the center of the object (x_0, y_0).

Therefore, the coordinates x_{p} and y_{p} of the new reference system can be defined as

$$x_{\mathrm{p}} = -(x - x_0)\sin(\mathrm{PA}) + (y - y_0)\cos(\mathrm{PA})$$
$$y_{\mathrm{p}} = -(x - x_0)\cos(\mathrm{PA}) - (y - y_0)\sin(\mathrm{PA}). \tag{6}$$

Then, the projected radius is such that

$$r = \sqrt{x_{\mathrm{p}}^2 + \left(\frac{y_{\mathrm{p}}}{1 - \varepsilon}\right)^2}. \tag{7}$$

So, with all this information, the functional forms that have been used in this work are constructed to fit the different types of objects.

3.2 Oversampling around the central zone

When using a model such as the Sérsic profile on a pixel mesh, it must be taken into account that the values obtained for each pixel are referenced to its center. Due to this, when there is a Sérsic profile with a very pronounced profile, that is, one that varies very quickly as we move away from the center, it is possible that the function varies from one extreme of the pixel to the other, so that the central value is not representative of the real value of the profile in the environment of that pixel. These pronounced profiles could appear, for example, for Sérsic index values of $n > 6$, approximately.

To illustrate this more clearly, let's imagine that we have a grid mesh with a size of (50,50) pixels, and our object is in the center, at 25. If we simulate here a Sérsic model with a Sérsic index $n = 8$, we will have, for example, around pixel 26, the graph that can be seen in **Fig. 6**. The value of the profile in the center of pixel 26 differs from the mean value of the profile within the pixel. To solve this consistency problem, oversampling is used.

The idea is to oversample each pixel by a predefined factor. In this work each pixel is divided into another 10 pixels in a mesh of (20,20) pixels around the center of the model. In this way, the values of the model are evaluated in these new pixels, and then it is rebinned to the original size, but saving, as the new value of the pixel, the mean of the subgrid created with the oversampling.
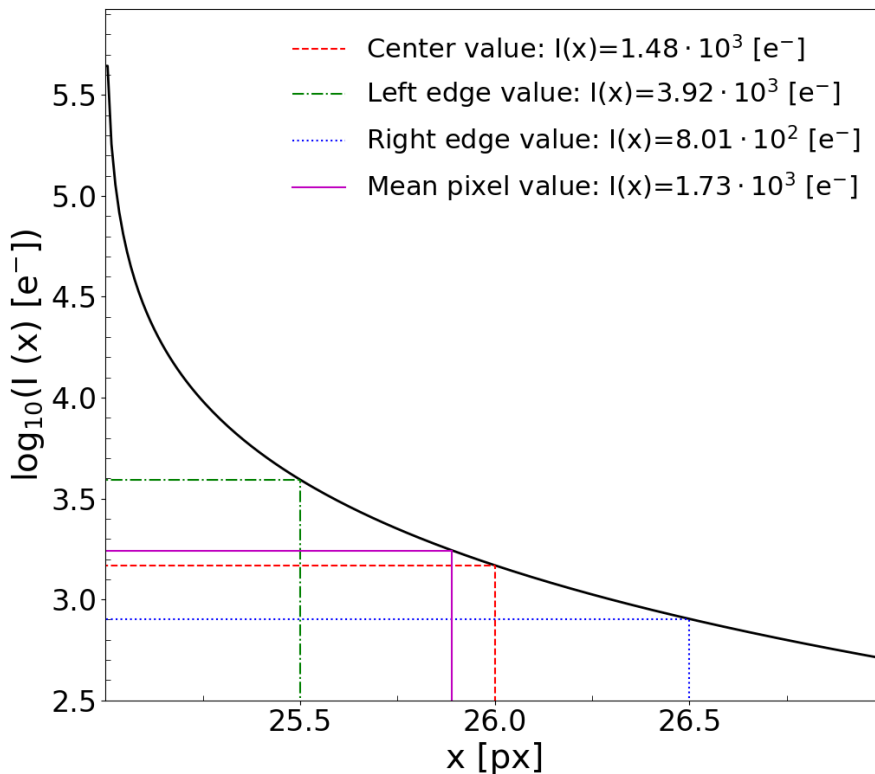
Fig. 6: Illustration of a Sérsic profile in one pixel with an index of $n = 8$.

3.3 Convolution with the PSF

Once a model has been generated, it must be convolved with the PSF. This is required to take into account the atmospheric contribution and compare the model with the observed data on an equal footing. A PSF model is created from a Moffat function, like the one seen in section. 2.2, in a mesh of a certain size. The number of pixels on each side must be an odd number, so that the

center of the PSF is in the central pixel of the mesh.

To carry out the convolution between the PSF and the created model, since the images can be considered periodic, the Fast Fourier Transform (FFT) has been used. The FFT is an algorithm that performs the Discrete Fourier Transform (DFT) for a sequence (Cooley & Tukey 1965). Specifically, in this work the FFT in two dimensions has been used, so as the DFT, which has an expression such that, being X a matrix of size (m, n)

$$\text{DFT}(X[m,n]) = \sum_{k=0}^{N-1}\sum_{l=0}^{N-1} \text{A}[k,l] \cdot \exp\left(i2\pi mk/N\right)\exp\left(i2\pi nl/N\right), \tag{8}$$

where $\text{A}[k,l]$ are complex coefficients such that

$$\text{A}[k,l] = \frac{1}{N^2}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} X[m,n]\exp\left(-i2\pi mk/N\right)\exp\left(-i2\pi nl/N\right). \tag{9}$$

Then, once it is clear how the FFT is defined, it is known that, by the convolution theorem, and having the image model, F and the PSF it can be written that

$$\text{F} \circledast \text{PSF} = \text{FFT}^{-1}\left\{\text{FFT}(F) \cdot \text{FFT}(\text{PSF})\right\}. \tag{10}$$

But, when using these definitions in Python, one detail must be taken into account, and is that the FFT has the first element of the image as the origin of its calculation, that is, the pixel in the upper left corner of the image. And the PSF that has been created has as its center at the center of the image. Then, it is necessary to "shift" the PSF before passing it through the FFT mechanism. The center of the PSF must be in the upper left corner, and it is required to do a zero padding of the PSF until it has the same dimensions as the image of the model before doing the padding. Now, it is possible to convolve the model with the PSF using Eq. (10), although only the real part of this result is retained. Once the convolution with the PSF is finished, a model is available that can be compared with the real image to be adjusted, now it is the turn of the Bayesian statistics.

3.4 Bayesian framework

The cornerstone of ANDURYL is based on how to build a model that can correctly reproduce the observation, based on the methods and profiles previously described. This is where Bayesian statistics comes in. The Bayesian inference is a statistical approach, based on Bayes Theorem (Bayes & Price 1763), that updates our knowledge about an observation as new information is acquired. This gives a probabilistic description from Bayes Theorem, as can be seen in Argyle et al. (2018)

$$p(\vec{x}|d) = \frac{\mathcal{L}(d|\vec{x})\, p(\vec{x})}{\int d\vec{x}\, \mathcal{L}(d|\vec{x})\, p(\vec{x})}, \tag{11}$$

where $p(\vec{x}|d)$ is the posterior probability distribution of the set of the model parameters that has been defined as $\vec{x}$. This is the result of the combination of other concepts, as the likelihood function, $\mathcal{L}(d|\vec{x})$, which is a joint probability density of the data as a function of the model parameters, that is, it measures the probability of obtaining the observed data given a specific model. The other term, $p(\vec{x})$, is the prior function of the vector of parameters $\vec{x}$, that represents the knowledge that we have a priori about each parameter.

Therefore, the objective is to calculate the posterior probability distribution and, based on it, calculate the marginal posterior distribution for each parameter, which is defined as

$$p(x_i|d) = \int dx_1, ..., dx_{i-1}, dx_{i+1}, ..., dx_n \; p(\vec{x}|d). \tag{12}$$

The marginal posterior is the probability distribution of a subset of the model parameters, after having integrated with respect to the rest of the parameters. In this way, inferences can be made from individual parameters.

3.4.1 Likelihood

In the astrophysical context in which this work is framed, in general, it can be assumed that the number of photons detected in each pixel of the image is high enough so that the Poisson distribution approaches a Gaussian distribution. Then, according to this, the expression for the individual per-pixel likelihood can be written as

$$p_i(d_i|m_i) = \frac{1}{\sqrt{2\pi}\sigma_i}\exp\left[\frac{-(d_i - m_i)^2}{2\sigma_i^2}\right], \tag{13}$$

where d_i are the pixels of the galaxy image, and m_i are the pixels of the generated model. The expression for the noise, that is a combination of different factors, is given by $\sigma_i = \sqrt{d_i + S + \sigma_{\text{RN}}^2}$; being S the sky background noise and σ_{RN} the readout noise related to the amplification of the chip's charge into an analogue voltage. From here, an expression for the likelihood can be obtained, summing over all N pixels

$$-\ln\mathcal{L} = -\sum_{i=1}^{N} \ln p_i \Rightarrow -\ln\mathcal{L} = \frac{\chi^2_{\text{unnorm}}}{2} = \frac{N\chi^2}{2} = \frac{1}{2}\sum_{i=1}^{N} \frac{(d_i - m_i)^2}{2\sigma_i^2}. \tag{14}$$

3.4.2 Priors

When doing Bayesian inference, it is extremely important to use all the physical information that is available a priori, in order to constrain the parameter space and that the results produced by the model are valid. That is, one can impose a constraint on the values of the parameters and, in turn, add a probability distribution to indicate which values are most likely to be obtained. However, this must be done with caution and noting that results might be biased for some a-priori information.

In this work, the studies given in Argyle et al. (2018) have been followed, using uniform priors for most of the adjusted parameters. Details about all the parameters that can be adjusted with ANDURYL and the priors used for each of them can be seen in the **Table. 2**. Apart from the parameters introduced by the photometric models, the parameters corresponding to the Moffatian PSF are incorporated as part of the model, which are the FWHM and β. The parameter σ_{add} is also added to the generated model, so it gives information of how well the background sky noise has been extracted from the image. The combined parameter Re/h is also included, which serves to guarantee that there are no component inversion problems, that is, so that the program does not try to adjust the disc with a Sérsic profile or the bulge with a Sérsic+exponential profile.

Table 2: Free parameters that are adjusted by ANDURYL together with the different priors and ranges for each of them. N_x and N_y are the number of pixels on each side of the image.

Sérsic parameters		
Parameter	Prior	Parameter space
Effective intensity, I_e [e$^-$]	Uniform in log (I_e)	[0,6]
Effective radius, R_e [px]	Uniform in $\log(R_e)$	[0,6]
Sérsic index, n	Uniform in n	[0.5,10]
Ellipticity, $\varepsilon_{\mathrm{ser}}$	Uniform in $\varepsilon_{\mathrm{ser}}$	[0,0.8]
Angle, PA$_{\mathrm{ser}}$ [°]	Uniform in PA$_{\mathrm{ser}}$	[0,180]
Exponential parameters		
Central intensity, I_0 [e$^-$]	Uniform in $\log(I_0)$	[0,6]
Scale length, h [px]	Uniform in $\log(h)$	[0,6]
Ellipticity, $\varepsilon_{\mathrm{exp}}$	Uniform in $\varepsilon_{\mathrm{exp}}$	[0,0.8]
Angle, PA$_{\mathrm{exp}}$ [°]	Uniform in PA$_{\mathrm{exp}}$	[0,180]
Common parameters		
Central coordinates, x_0, y_0 [px]	Uniform in (x_0/N_x) &(y_0/N_y)	[0.1,0.9]
FWHM Moffat, FWHM	Uniform in FWHM	[0,10]
β Moffat, β	Uniform in β	[0.5,12]
Noise, σ_{add}	Uniform in σ_{add}	[0,10]
Combined parameter R_e/h	Uniform in R_e/h	[0,1.678]

3.5 Bayesian engine

When evaluating the posterior marginal distribution, the equation Eq. (12), in general, an analytical solution will not be available. It is necessary to approximate the integral. There are many options to carry out this operation, and in this work two of them have been explored: Hamiltonian Monte Carlo and Nested Sampling.

3.5.1 Hamiltonian Monte Carlo

The most traditional methods for sampling posterior distributions are based on performing a random walk over the parameter space, which has been defined as $\vec{x}$ previously. The Hamiltonian Monte Carlo (HMC) (Duane et al. 1987) is an algorithm that is based on the analogies of the trajectories of a physical system and the probability distributions. Specifically, it uses Hamiltonian dynamics to generate new samples of the probability distribution, using Hamilton's equations to infer how positions and momentums evolve over time in the parameter space. All this is combined with a Monte Carlo sampling, which accepts only proposals that improve the probability distribution.

In this method, a new variable is introduced in the denominator of Eq. (11) (henceforth denoted as $f(\vec{x})$), the momentum $\vec{p} \in \mathbb{R}^D$. So it can be written that

$$f(\vec{x}) = p(\vec{x})\mathcal{L}(d|\vec{x}) \Rightarrow f(\vec{x})p(\vec{p}) = p(\vec{x})\mathcal{L}(d|\vec{x})p(\vec{p}). \tag{15}$$

This momentum increases the space and leads to defining a joint probability $f(\vec{x})p(\vec{p})$. Therefore, being D the dimensions of the space of parameters analyzed, with this a negative log joint probability is defined as

$$H(\vec{x},\vec{p}) = -\log(f(\vec{x})) + \frac{1}{2}\log\left[(2\pi)^D|\vec{M}|\right] + \frac{1}{2}\vec{p}^{\,\mathrm{T}}\vec{M}^{-1}\vec{p}, \tag{16}$$

beeing $\vec{M}$ the mass matrix. This matrix is used to control the relations between positions and momentums, influencing how the HMC explores the parameter space. In other words, it is used to

adjust the scale of the momentums in the parameter space. In practice this matrix can be chosen, some of the available options being the identity matrix (assumes a unitary mass for all dimensions of the parameter space) or a diagonal matrix, with different masses to adapt to the different scales of the parameters. The last member of Eq. (16) is a kinetic energy term and the first term is associated with potential energy. Trajectories moving in this parameter space follow each other along constant energy levels. Then, to move through this parameter space preserving the total energy and volume, the following information is required

$$\frac{d\vec{x}}{d\tau} = \frac{\partial H}{\partial \vec{p}} \; ; \; \frac{d\vec{p}}{d\tau} = -\frac{\partial H}{\partial \vec{x}}, \tag{17}$$

being τ the step iteration. Therefore, the sampling process basically consists of drawing the momentum. Then, a new set of parameters $\vec{x}^*$ is defined, through for example a leapfrog algorithm, and repeating this process until sampling the set of parameters of interest.

There are several methods that are derived from the HMC, but the one that has been used in this work is the HMC-NUTS (No-U-Turn Sampler; Hoffman & Gelman 2011; Cobb et al. 2019). It uses a recursive algorithm to construct a collection of probable candidate points that covers a broad range of the target distribution. The algorithm automatically halts when it detects backtracking and retracing of steps. It also has the ability to adapt its step size, τ, to a more suitable one.

All of this Bayesian method was implemented within the ANDURYL core via the "hamiltorch"[1] package (Cobb et al. 2023), a Python package built on top of PyTorch[2] that works with tensors. PyTorch is a Python-based framework that offers tools for model optimization, performance evaluation, etc. "hamiltorch" takes advantage of the automatic gradient calculation and optimization capabilities of PyTorch to perform efficient calculation of the gradients needed in the HMC method. It allows Bayesian inference on models with continuous, discrete and mixed parameters. It is for this reason that, within the program, the modules are also built using PyTorch.

3.5.2 Nested Sampling

The second sampling method that has been used in this work is called Nested Sampling (Skilling 2004). Nested Sampling employs a different method than other Bayesian inference methods. Here, the key idea is to enclose regions of the parameter space that have a higher probability than the regions of the previous iteration. To describe how this method works, we also start from the denominator of Bayes' Theorem, which is called the marginal likelihood

$$\mathcal{Z} = \int d\vec{x} \; \mathcal{L}(d|\vec{x}) \; p(\vec{x}). \tag{18}$$

The objective of Nested Sampling is to calculate $\mathcal{Z}$, generating samples of $\vec{x}$ according to the posterior distribution. To deal with this integral, Nested Sampling employs an alternative, where a variable change is applied such that $d\vec{\theta}(\vec{x}) = p(\vec{x})d\vec{x}$. In this way $\mathcal{Z}$ is redefined as a sum over prior mass and it is expressed as the area under a monotonic one-dimensional curve. The mass represents the fraction of total probability of the prior that is contained within said curve, therefore

$$\mathcal{Z} = \int_0^1 d\vec{\theta} \; \mathcal{L}(\vec{x}(\vec{\theta})), \tag{19}$$

each volume element of the original prior in its parent parameter space is now mapped to a one-dimensional scalar element. So, given a number of points, the area under the curve can be

---
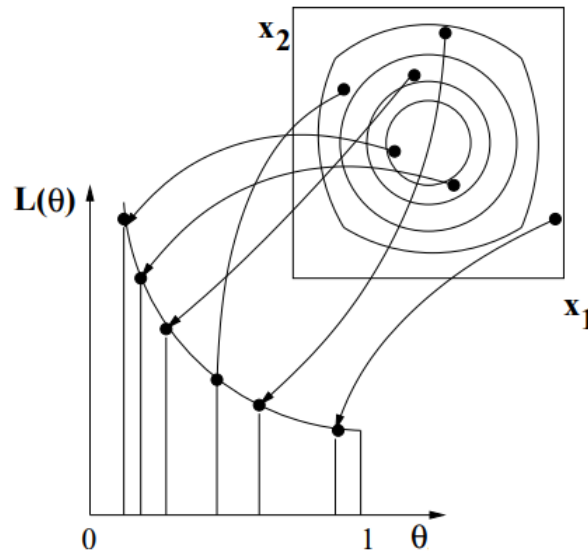
[1] https://github.com/AdamCobb/hamiltorch
[2] https://pytorch.org/

Fig. 7: Illustration of the iso-contours on which Nested Sampling is based to integrate the likelihood. Taken from Mukherjee et al. (2006).

approximated. These points are generated from the known $p(\vec{\theta})$ distribution (Murray et al. 2005). Then, Nested Sampling integrates the likelihood over the volume of the prior into iso-contours of equal likelihood, as can be seen in **Fig. 7**.

Nevertheless, this methodology describes the simplest method of Nested Sampling. There may be occasions in which the characteristics of the problem studied require another type of approach, as is the case in this work.

Ellipsoidal Nested Sampling

Drawing samples from the prior is a complicated step. To try to solve this drawback, ellipsoidal sampling tries to approximate the contours that have the same likelihood through a D-dimensional ellipsoid, using the covariance matrix of the points that are active in the iteration. And the new sets of points are selected from those that are inside this same ellipsoid, not as in the previous case where the selection could result in a decrease in the volume of the prior (Feroz et al. 2009).

The new points that are selected through ellipsoidal sampling guarantee that the new enclosed volume has a better likelihood than in the previous volume, that is, an ellipsoid is sought that encloses a number of points such that this condition is met.

Although even this method is not without drawbacks, drawing a single ellipsoid may not work well for distributions where the posterior is multimodal, as can be seen in **Fig. 8**. To solve this problem of high degeneration, the number of active points can be divided into different sub-ellipses, so that the degeneration can be followed more easily. The volume of these ellipses can be arbitrarily increased or decreased in order to find the solution more quickly.

When implementing these Nested Sampling methods, the multi-ellipse method has been chosen, through the Python package "nestle" from MIT (Barbary et al. 2023). This package allows high flexibility by allowing the choice of the number of active points, as well as a multiplicative factor of the volume of the drawn ellipses. For practical reasons, in this work the original code of "nestle" has been modified by adding a watchdog; this means that an element has been added that allows
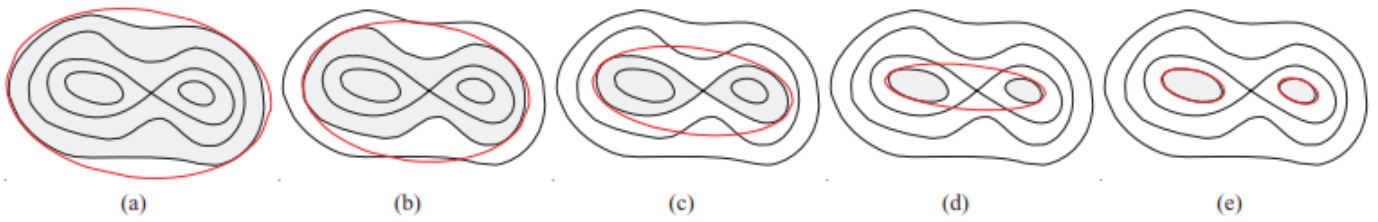
Fig. 8: Illustration of how ellipsoidal nested sampling works, for a simple bimodal distribution. From figure (a) to (d) we can see how with a single ellipse it is not possible to characterize the two regions well. While in figure (e) it can be seen that both regions can be characterized if a multi-ellipsoidal model is used.

choosing how much time can be invested in the fit. If the method takes too long to converge, it can skip it and go to the next one, if there is one.

3.6 Obtaining and visualizing the results

Once the posterior sampling process is finished, a vector of parameters is obtained with a size equal to the number of iterations that have been necessary to sample the posterior distribution. In this way, the result is summarized taking the mean marginal value for each parameter.

When the optimal values for each parameter are obtained, the final model can be built, using a Sérsic profile or a combination of Sérsic+exponential, depending on the model that has been used. Now the residuals can be calculated, which are nothing more than subtracting the model from the data. The residuals provide information about the unadjusted components of the galaxy, that is, if the galaxy has, for example, a bulge, a disc and a bar and the model only include the bulge and the disc, the bar will appear in the residuals.

The posterior distribution of the analysis produced by ANDURYL can also be displayed, where the marginal distributions of the parameters can be seen together with the correlations that may exist between them. These diagrams are known as triangular plots, and an example of one of them can be seen in **Fig. 9**. The shape of the marginal distributions of pairs of parameters gives an indication of the covariance between those parameters. Thus, for example, if there is a gaussian form of it, it seems to indicate that there is no direct correlation between them; but if the shape of the plot has, for example, the shape of a straight line, it would indicate a linear correlation between both parameters.

With the mean of the marginal distributions of each parameter, the final model can then be generated. In this way it can be compared with the observation and represent the residuals, which account for those components not resolved by the model. An example can be seen in **Fig. 10**.
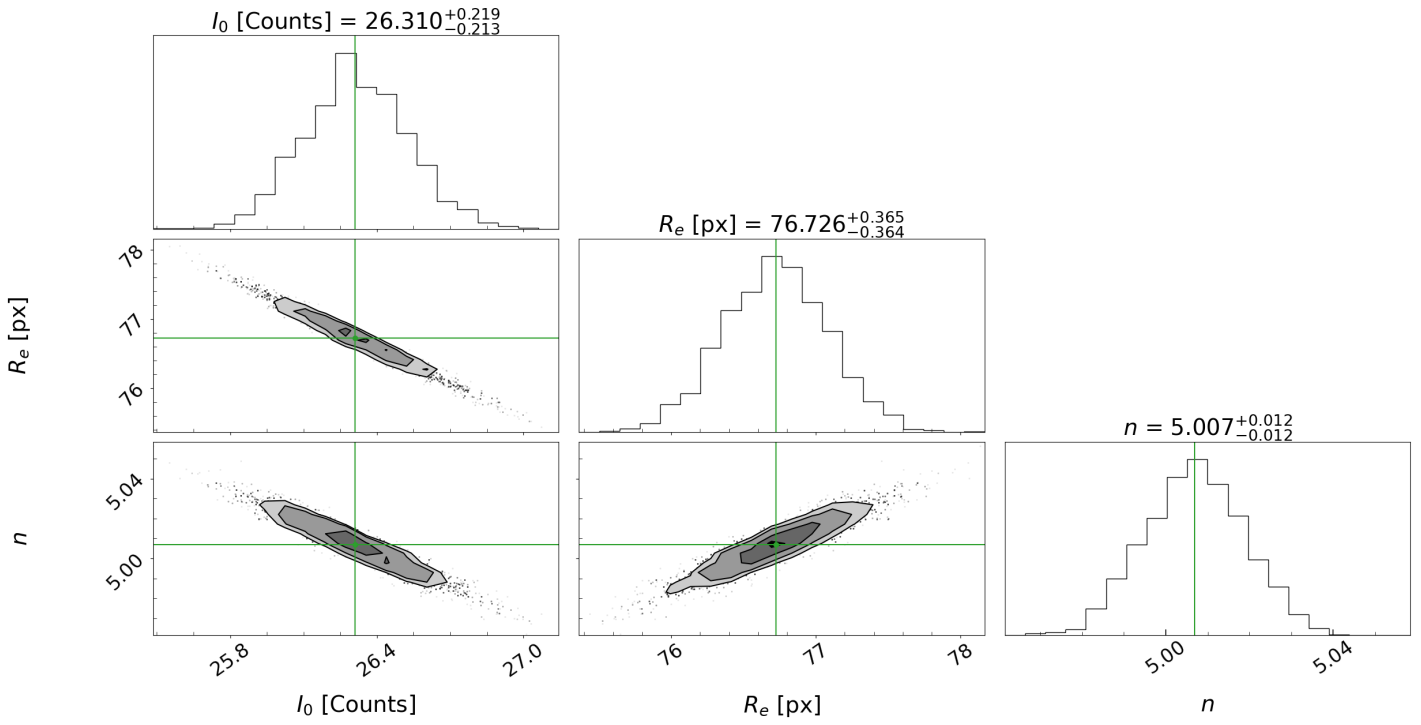
Fig. 9: Example of a triangular diagram showing the posterior marginal distributions for a Sérsic fit of a galaxy. The marginal distribution of each parameter can be seen in the upper diagonal. While in the rest of the panels the joint marginal distribution of pairs of parameters is shown. The grayscale contours are shown at 0.5, 1, 1.5, and 2 σ. And the green solid line indicates the mean value of the posterior distribution.

Fig. 10: Example of a photometric decomposition using ANDURYL. The observed galaxy can be seen on the upper left, the generated model in the upper right, and the residuals on the bottom.

Chapter 4

Calibration and test using mock galaxies

4.1 Comparing HMC and Nested Sampling

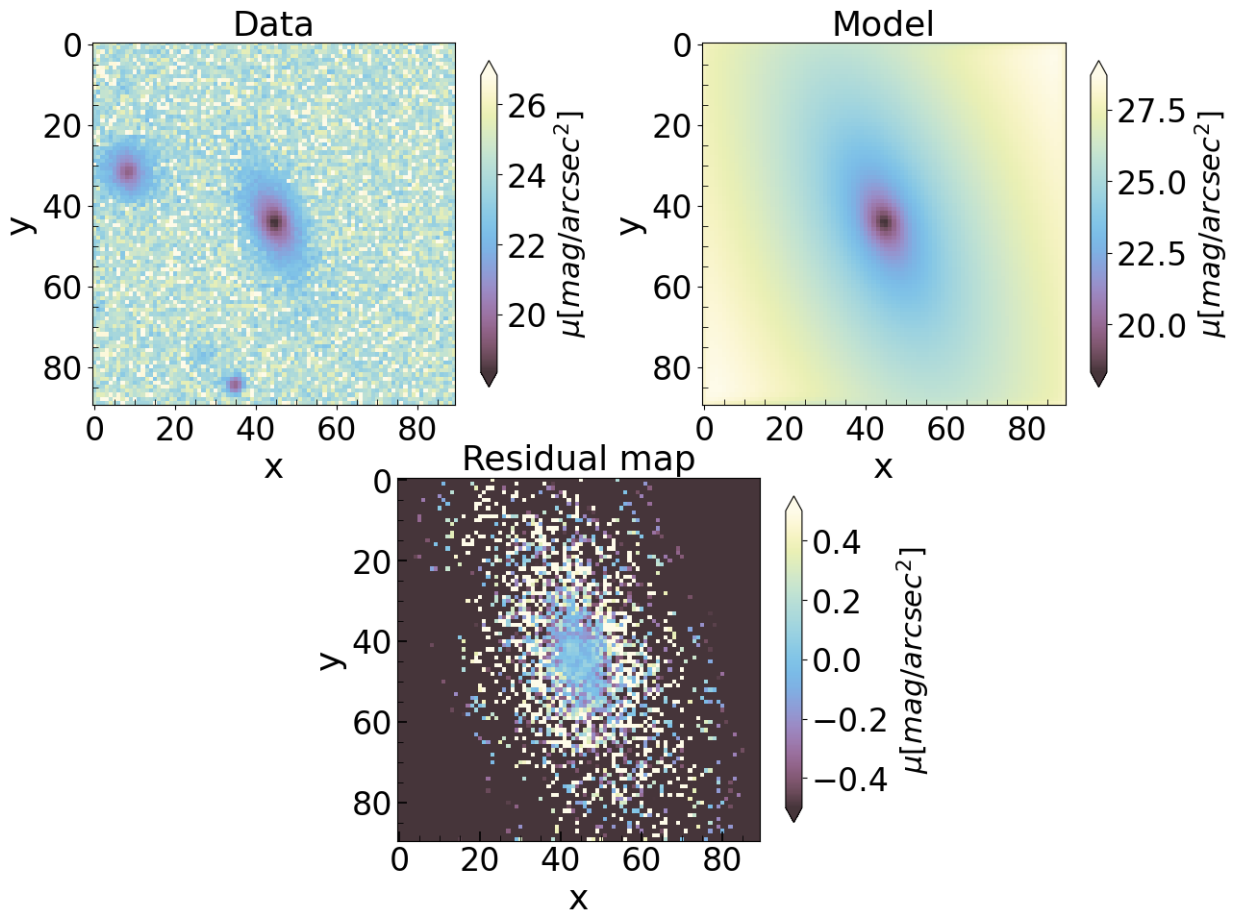To test which one of the Bayesian methods was more appropriate for carrying out this work, a preliminary test was carried out with the galaxy NGC0155. This have been proved to be adequately fitted using only a Sérsic profile, as demonstrated in Méndez-Abreu et al. (2017). As can be seen from the results in the Table. 3, the values obtained by GASP2D and by ANDURYL using both methods are certainly similar, this could lead us to think that both methods, HMC and Nested Sampling are equally valid.

Table 3:   Comparison between the values obtained by the GASP2D program and the values obtained by ANDURYL using HMC and Nested Sampling for the galaxy NGC0155 using a Sérsic profile.

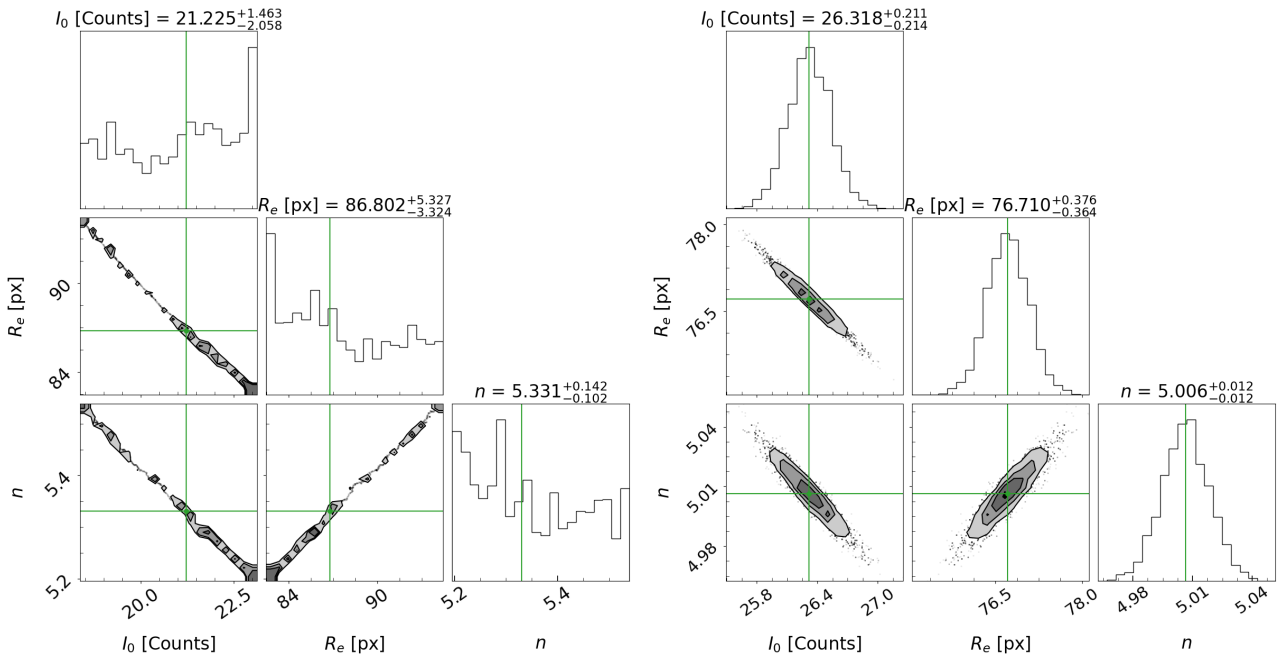| Method/Program	I_e [Counts]	R_e [px]	n	$\varepsilon_{\mathrm{ser}}$	$\mathrm{PA}_{\mathrm{ser}}[°]$	x_0 [px]	y_0 [px]
GASP2D	24.51	80.05	5.10	0.23	84.68	182.67	154.95
HMC	21.23	86.80	5.33	0.24	95.39	182.68	154.95
Nested Sampling	26.30	76.74	5.01	0.23	95.36	182.67	154.95

Fig. 11: In the left part the triangular plot associated with the fit of the galaxy NGC0155 with the HMC method can be seen and in the right part with the Nested Sampling method, both from the ANDURYL code.

However, if triangular plots are displayed for both cases (see **Fig. 11**), the figure corresponding to the HMC seems to indicate that it has not converged. This can be deduced from the form of the marginal distributions that have non-Gaussian forms. In contrast the marginal distributions given by Nested Sampling have Gaussian shapes and the correlations have shapes that are more in line with what is expected, keeping a Gaussian shape in the case of being random correlations.

The reason why HMC sampling seems not to converge can be deduced from Figure 3 of Cobb et al. (2019). In the present work HMC-NUTS was used, and it can be seen that by adapting the step-size this makes the algorithm incapable of explore the narrowest areas of the model's geometry. And, precisely, the likelihood that is analyzed in this work seems to have a very narrow shape, given the sampling that the Nested Sampling method shows. This does not mean that HMC is not a valid method, but rather that it would be necessary to explore what mechanisms would be necessary to achieve proper functioning.

After detecting the better reliability of Nested Sampling, this was implemented in the code and it was parallelized using the Python "joblib" package [1]. This way, different cores of the computer system processor can be used. In order to reduce the computational time necessary for the analysis of large data sets.

4.2 Mock galaxies sample

After choosing the Bayesian sampling engine, the analysis of the mock galaxy samples was carried out, comparing the values of the parameters with used to create the mock models with those obtained using ANDURYL. This step aims to test the reliability of the code, look for possible systematic errors and correct them and work on optimizing it.

4.2.1 Inference of the bulge mock galaxy sample

It can be seen in the results of **Fig. 12** that the inferred parameters shown with quartiles of the marginal posteriors, for the bulge mock galaxy sample (with a Sérsic profile) are quite consistent with the input parameters given by GASP2D. Although the relative errors are small, being in most cases a maximum of $2 \sim 6$ %, some slight systematic errors can be seen in the effective intensity and effective radius plots. In the case of effective intensity, it is seen as an upward trend as the value of the effective intensity increases, and in the case of effective radius, the values seems to present a certain offset between $0 \sim 1$ %. Given the dimension of this error, it could be associated with a multitude of variables; from a numerical difference in the precision with which mock galaxies have been made (since they have been made in IDL) and with which they have been analyzed in Python, to a difference due to the way in which oversampling is done (explained in section. 3.2). This is important, since oversampling significantly affects the values obtained for these quantities. In the rest of the parameters, there do not seem to be appreciable trends.

For the case of the angle PA, it can be seen that, at low values, the difference extends to higher errors than for the higher values of the angle. The PA has two extremes in which singular behaviors appear. One happens close to 0 degrees, since this angle is equivalent to 180 degrees, and the posterior is bimodal. The other one is in the case of low ellipticity, for which the PA is not well defined.

---

[1] https://joblib.readthedocs.io/en/stable/
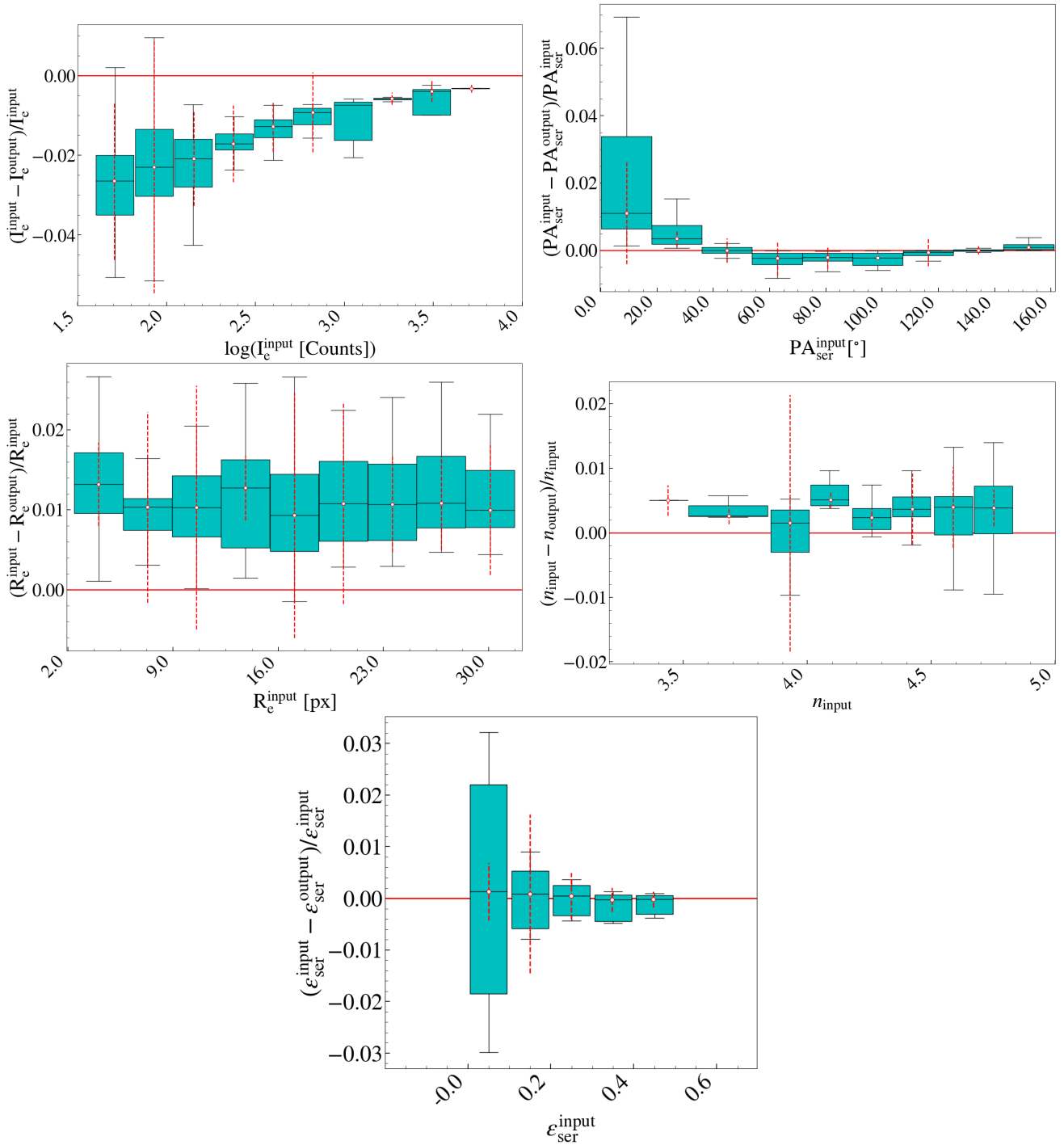
Fig. 12: Different box plots for the parameters resulting from fitting the sample of bulge mocks with a Sérsic profile. The box extends from the first quartile (Q1) to the third quartile (Q3) of the data, with a black line at the median. The whiskers extend from the box by a factor of 1.5 the inter-quartile range (IQR, between Q1 and Q3). The red dotted lines represent the mean of the error, given by ANDURYL, of all the points that are inside each box.

Similarly, when inferred the sample of bulge mock galaxies with a Sérsic+Exponential profile, it is found, as can be seen in **Fig. A.1**, that the relative error of all the parameters is also very small compared to the typical values of each parameter. In addition, they are consistent with the error provided by ANDURYL. The objective of fitting bulge mock galaxies with a profile that can model a bulge and a disc is to check that ANDURYL can correctly model the bulge, minimizing as much as possible (within the limits imposed by the priors) the contribution of the disc. Analogously to the previous case, a certain trend can be seen in the calculations of the parameters I_e and R_e which, again, might be a precision difference. It has not been possible to carry out this type of plots for the parameters given for the disc since the disc does not exist.

It has been possible to verify that the mean values of the marginal distributions tended towards the lower limits of the priors, for the parameters related to the disc. In this way, the influence of the disc in the model is minimal, affecting in a negligible way the inference of the bulge parameters. This can be verified by seeing that the relative errors obtained in this case are quite similar to those when only a Sérsic profile was used.
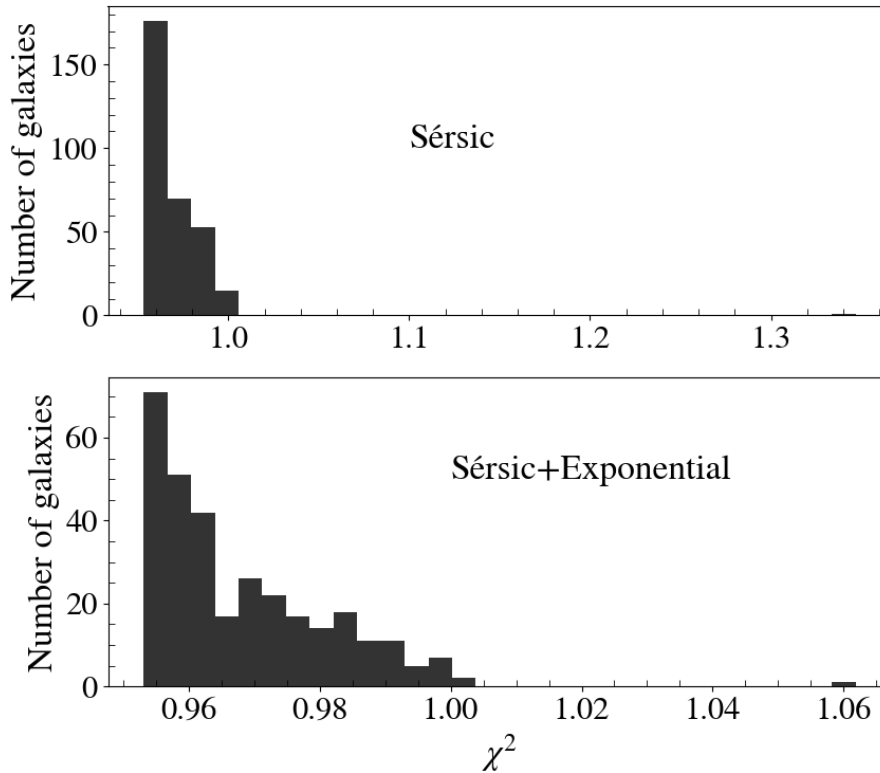

Fig. 13: Distribution of the values of χ^2 for the bulge mock samples, the samples in the upper panel being those created with a Sérsic profile and those in the lower panel with a Sérsic+Exponential profile.

Fig. 13 shows the distribution of the values of χ^2 obtained from the inference of the parameters of bulge mock galaxies. It can be seen that, being analyzed with any of the two models, either Sérsic or Sérsic+exponential, acceptable results are obtained. It is considered that a value of χ^2 is acceptable when it is around the unit, lower values could indicate an overestimation of the weights/noise, and higher values could indicate an underestimation of the same, or that the fit has been unsuccessful. Here, in both cases, the values are consistent with a successful inference.

4.2.2 Inference of the bulge-disc mock sample

In the case of the bulge-disc mock samples, when analyzed only with a Sérsic profile, the values obtained for the different bulge parameters do not coincide with those used to create the models (see

Fig. A.2). It can be clearly seen that the relative errors are around two orders of magnitude higher than those obtained in previous cases. This does not mean that the program is malfunctioning, but rather it is a consequence of trying to adjust two components using only one. Depending on how the characteristics of the discs are, sometimes a successful inference can be achieved. However, it provides an erroneous value of the parameters that would be expected to be obtained from the Sérsic profile, since it is have been readjusted to try to follow the profile of the disc.

However, if these mocks galaxies are inferred with a Sérsic+Exponential profile, a good agreement is obtained between the input values and those obtained by ANDURYL (see **Fig. A.3**). Although keep in mind that when adjusting two components it is essential to use combined parameters to, among other things, avoid the component inversion problem. The imposition of more priors that provide physical limits to the adjusted quantities is essential to achieve a correct fit. Relative errors are slightly higher than the cases of the bulge mock galaxy samples, but even so, they are within a reasonable range, which is sometimes compatible with the statistical dispersion provided by ANDURYL.
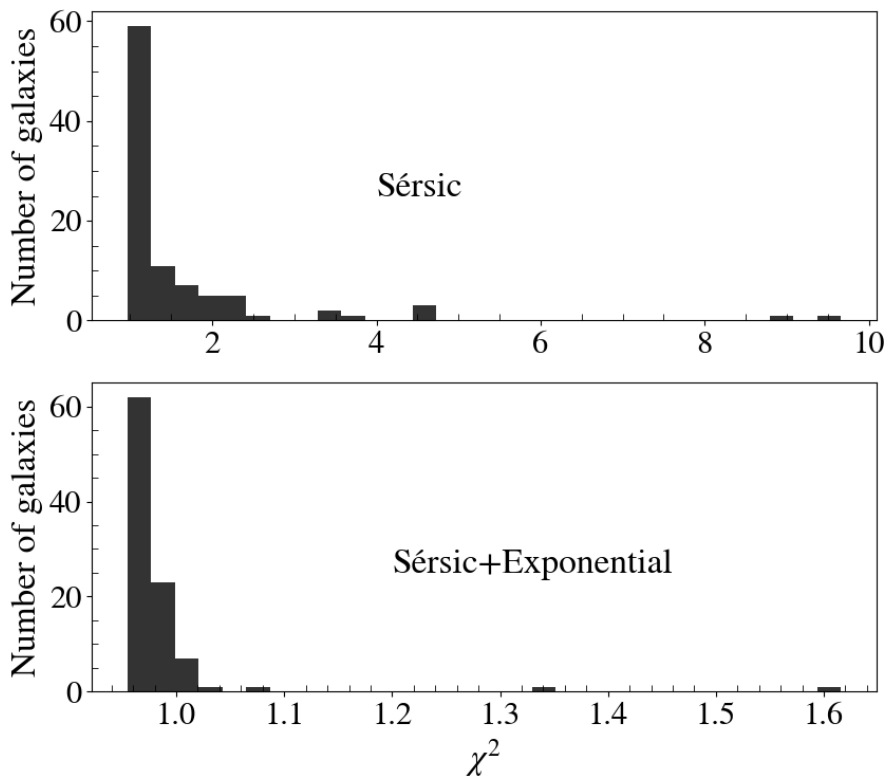

Fig. 14: Distribution of the values of χ^2 for the bulge-disc mock samples, the samples in the upper panel being those created with a Sérsic profile and those in the lower panel with a Sérsic+Exponential profile.

Fig. 14 shows the distribution of the values of χ^2 obtained for the bulge-disc mock galaxies. It can be seen that, when they are analyzed with a Sérsic profile, there is a part, around tens, that exceeds the value of the unity in settings; while when they are analyzed with a Sérsic+exponential profile, values around unity are obtained. This is again consistent with the values obtained for the relative error plots, which indicated anomalies in the analysis with the Sérsic profile. However, it is observed that even so there is a reasonable number of galaxies that have a χ^2 around unity, despite the fact that it has been seen that the parameter settings differed greatly with the input parameters. This can be a problem when analyzing real galaxies, since it is possible that the χ^2 criterion is not sufficient to differentiate a good fit from one that is not, and for this reason it is necessary another more robust criteria.

4.3 Selection criteria

Once the sample of mock galaxies have been analyzed, we ask ourselves how to distinguish which of the models is the most appropriate to adjust a galaxy.

The usual criterion that has been used to carry out this separation is the Bayesian Information Criterion (BIC; Schwarz 1978)). This criterion uses the χ^2_{unnorm} (Eq. (14)) that is inferred from each fit made, and it adds a penalty for the number of parameters used in the fit proportional to the number of adjusted points. So, the equation is

$$\text{BIC} = \chi^2_{\text{unnorm}} + k\ln(m), \tag{20}$$

being k the number of free parameters used in the fit and m the number of adjusted points, that is, the number of points that remain in the image once the mask has been applied. However, as noted in Simard et al. (2011) and Méndez-Abreu et al. (2018), the pixels of the analyzed image are not independent, since when convolved with the PSF, information from different pixels is mixed. Therefore, following the steps given in the literature, the number of pixels has been replaced by the number of resolution elements $m_{\text{res}} = m/A_{\text{PSF}}$, where A_{PSF} is the area of the PSF at Full Width Half Maximum (FWHM). Therefore, the expression for the BIC is finally given by

$$\text{BIC} = \frac{\chi^2_{\text{unnorm}}}{A_{\text{PSF}}} + k\ln\left(\frac{m}{A_{\text{PSF}}}\right). \tag{21}$$

If a fit is made to a galaxy with a Sérsic profile and a Sérsic+Exponential profile, the BIC given for each fit can be calculated, and from them the Δ BIC quantity can be calculated, which is nothing more than Δ BIC = BIC (Sérsic)-BIC(Sérsic+Exponential).

Then, from this amount, it follows that if Δ BIC < 0 then the most appropriate BIC would be the fit with the Sérsic profile, and otherwise the Sérsic+Exponential profile. But this number, Δ BIC < 0, is arbitrary and it is necessary to calibrate it using mock galaxies. This step must be done, since the separation criteria between components are not unique. Being a complex case, it is necessary to do this calibration to know in what order of Δ BIC the program can distinguish between one type of galaxies or another. In the paper of Méndez-Abreu et al. (2018) it can be seen that the separation value obtained is around Δ BIC=-18, slightly higher than the one obtained in this work, thus showing the need for calibration. To do this, histograms of this quantity are made for the mock galaxies samples, for the bulge samples analyzed with Sérsic and Sérsic+Exponential profiles and for the bulge+disc sample analyzed with the same profiles. In this way, a logical limit can be inferred for the value of Δ BIC that allows the separation between both profiles. Once this threshold is obtained from the sample of mock galaxies, it can be used as a criterion for real galaxies, allowing the selection of the most appropriate model for each one.

4.4 Separation of components in mock galaxies

Through the process described in section. 4.3, the BIC has been calculated for the four sets of mock galaxies analyzed with the different photometric profiles. The results for Δ BIC can be seen in **Fig. 15**. In these results, except for some outlayer galaxy, it can be seen that the sample of mock galaxies created with a Sérsic profile form a narrow distribution below a certain value of Δ BIC=-24.2, and that the sample of mock galaxies created with a Sérsic+Exponential profile are mostly grouped above this value. Since the sample of mock galaxies created gives a good account of the behavior of real galaxies, this criterion is considered as the separation criterion between elliptical and disc galaxies. In this work only two components have been adjusted, so really all those real galaxies that could have some more components are classified as a disc galaxy.
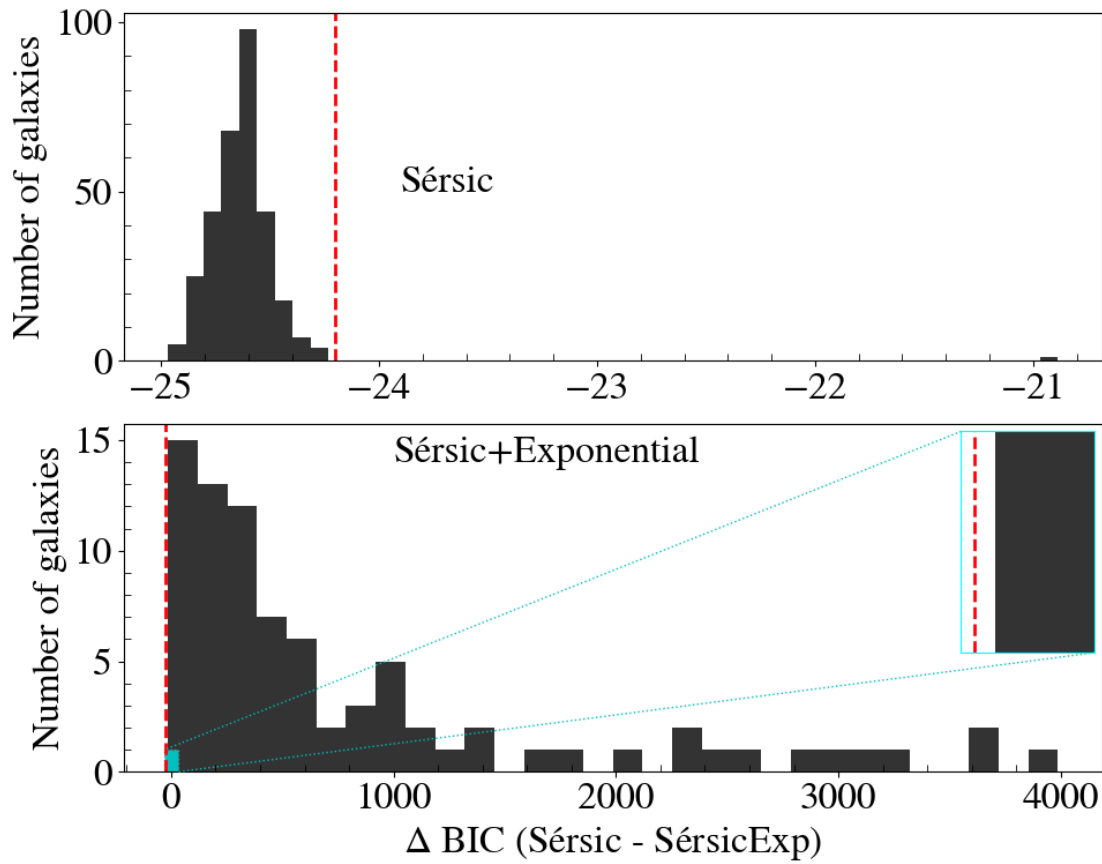

Fig. 15: Distribution of the values of the Δ BIC for the mock galaxies samples, the samples in the upper panel being those created with a Sérsic profile and those in the lower panel with a Sérsic+Exponential profile. The vertical dotted lines symbolize the limit that has been considered Δ BIC = -24.2 as best fit separator.

Chapter 5

Results and discussion

Once the code has been assembled, and it has been calibrated with the use of mock galaxy samples, we apply ANDURYL to real galaxies. For this a bulge fit and a bulge-disc fit have been performed on the sample of real galaxies from the cluster A2142. Of the 234 galaxies that were originally taken from the A2142 sample, 175 galaxies have finally been successfully fitted. The reason for discarding the rest of the galaxies is based on convergence problems, since the program did not converge within the imposed time limit of 80 minutes, and automatically proceeded to another galaxy. A measurement of scaling relations such as Faber-Jackson and mass-size has been made; in order to verify that our photometric-decomposition results follow these relations and discuss it.

5.1 Separation of components in real galaxies

If the same treatment as in section. 4.4 is carried out for for the sample of real galaxies, the results shown in **Fig. 16** are found. It can be seen that the majority of the real galaxies analyzed have been catalogued, according to the previous criteria, as elliptical galaxies, around 141; while 34 have been classified as disc galaxies. Galaxies classified as disc galaxies may have some other components that have not been adjusted. It would not be correct to say that they are purely disc galaxies, simply that they have a bulge and a disc. These may also have some other unresolved component, which should be detected when calculating the residuals.

The fractions of galaxies obtained for each case hide something more than just a number, $\sim$ 80 % of elliptical galaxies and $\sim$ 20 % of galaxies that have at least one disc have been obtained. It is necessary to highlight again the galaxy sample in which the study is being carried out, which is in a cluster of galaxies located in the local Universe; where the late-type galaxies are the least representative morphologically (less than 20% of mass; Vulcani et al. 2011). ANDURYL is not yet equipped with the necessary tools to automatically distinguish galaxies with spiral arms or other components beyond the bulge and disc, but the numbers obtained seem to indicate that most of the galaxies in our study are of early type, which is compatible with the literature. It should also be taken into account that there is still a number of galaxies to analyze, which could play a significant change in these numbers.

5.2 Distribution of galaxy parameters

Now that the models that best fit for each galaxy are available, the distribution of parameters of all of them can be analyzed, separating between the Sérsic profiles and the Sérsic+exponential profiles. Parameters such as the center, ellipticity or position angles have not been represented in this case, for the sake of clarity. Thus, the effective intensity, the effective radius, and the Sérsic index have been represented for the case of the Sérsic profiles; while for the Sérsic+Exponential profile, the central intensity and the scale length of the disc are also added. Both intensities have
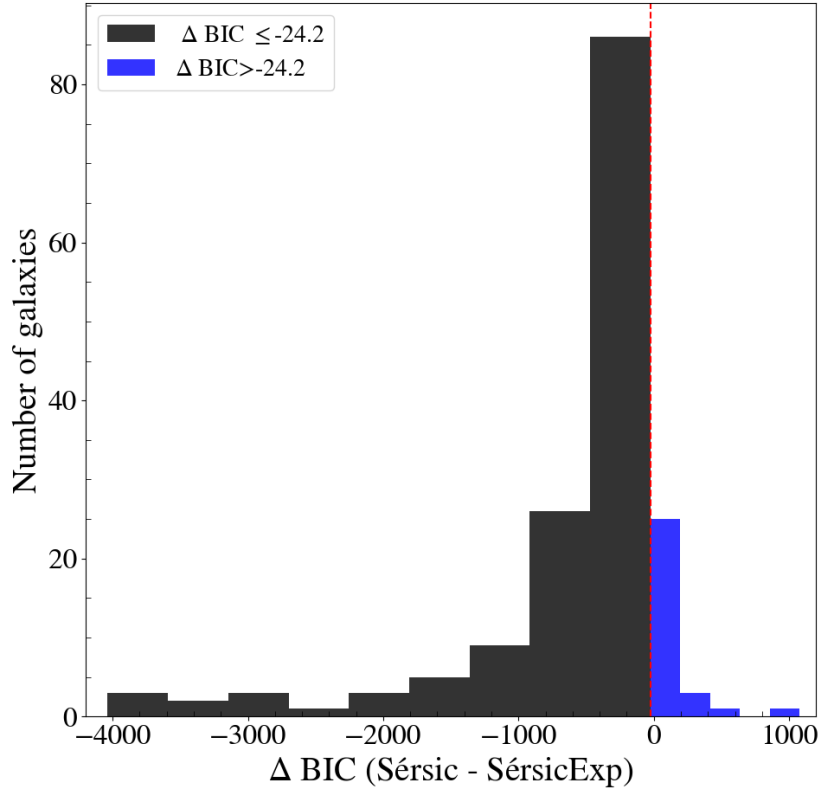
Fig. 16: Distribution of the values of the Δ BIC for the real galaxies. The vertical dotted lines symbolize the limit that has been considered Δ BIC = -24.2 as a separator for a galaxy to be considered suitable to be represented only with a Sérsic profile.

been transformed to surface brightness, measured in [mag/arcsec^2] :

$$\mu_{e/0} = -2.5\log(I_{e/0}) + z_{cal}, \tag{22}$$

where z_{cal} is a calibration value that depends on the scale of the CCD plate.

Fig. 17 shows the histograms obtained for the parameters of the galaxies fitted with a Sérsic profile. It can be seen that there are of the order of 9 galaxies, either with a low effective brightness around $\mu_e = 24 \sim 26$ mag/arcsec^2. And some 15 galaxies with an anomalously high Sérsic indices, around $n = 8 \sim 10$. After analyzing the triangular plots of these galaxies, we found that the marginal posterior distribution of the Sérsic index does not converge properly, but rather collapsed towards the upper edge of the prior. This indicates that the fit is highly likely not to have converged for these galaxies. But, apart from that small percentage of galaxies, the rest seem to group around an average value, which in a way justifies having used uniform priors for these parameters.

For those galaxies analyzed with a Sérsic+exponential profile **Fig. B.1** shows the histograms of the parameters. In this case, it can be seen that there are a number of galaxies with a fairly low central surface brightness, around $\mu_e = 24 \sim 25$ mag/arcsec^2. The value of Δ BIC for these galaxies has been analyzed and they were close to the value of cut of Δ BIC= -24.2; with which it is highly probable that these galaxies could be galaxies of the elliptical type. In addition, in this case the trend around the mean seems to fade with respect to the previous case, although it is true that the sample of galaxies is much smaller. And similarly to the previous case, some Sérsic indices are also observed in the limits of the associated prior, for the same reason.
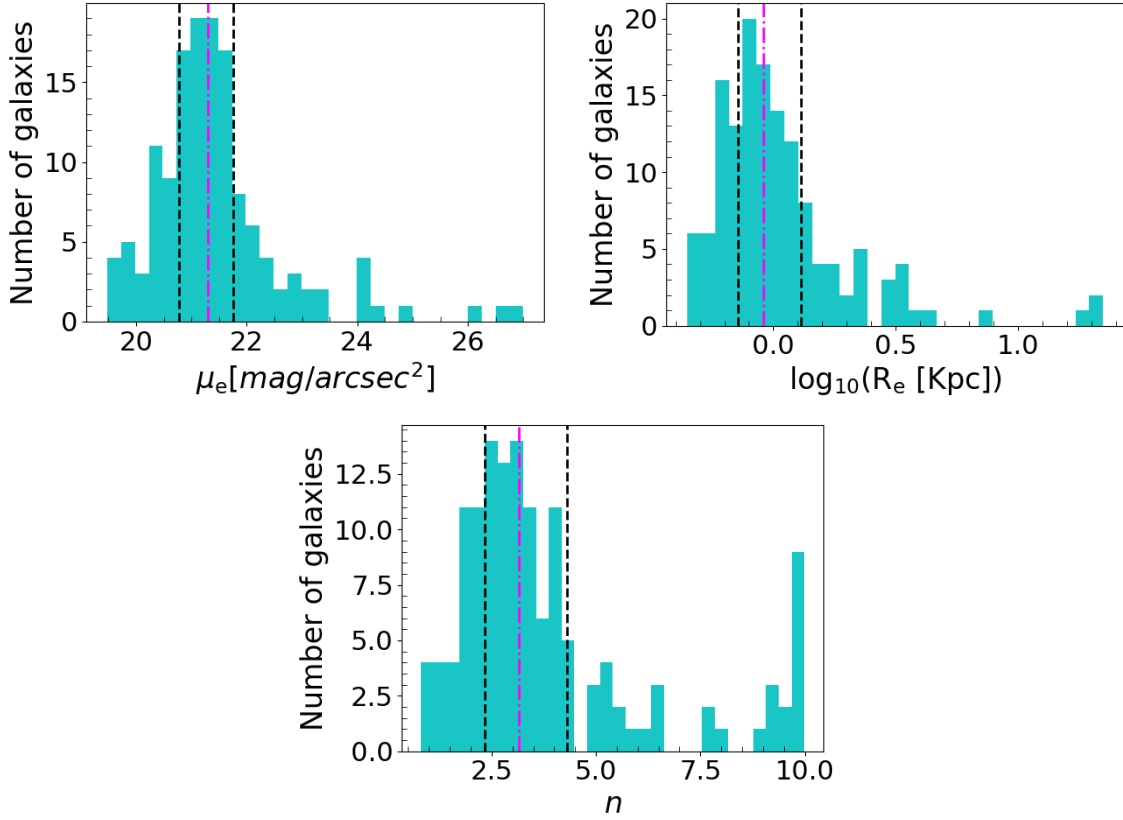
Fig. 17: Distribution of the parameters of the galaxies fitted with a Sérsic profile. In order from left to right, effective surface brightness μ_e, effective radius R_e, and Sérsic index n. The magenta vertical lines represent the mean of the distribution, and the black vertical lines the Q1 and Q3 quartiles.

5.3 Scaling relations and magnitudes

The Faber-Jackson relation (Faber & Jackson 1976), as seen in section. 1.3, connect the luminosity, or magnitude of a galaxy, with the velocity dispersion. This follows from the virial theorem as

$$L \propto \sigma^4 \Rightarrow M_B^i \propto \sigma^4, \tag{23}$$

where the superscript "i" refers to the "i" band of the SDSS filters. The magnitudes can be measured through the Sérsic profile equation, integrating it following Graham & Driver (2005), over an area such that $A = \pi r^2$, so the following integral can be written

$$F = \int_0^R I(r) 2\pi r \ dr, \tag{24}$$

where the intensity term is given by the Sérsic profile, redefined in exponential form to make its integration easier, also redefining the term of b_n

$$I(r) = I_e \exp\left\{-b_n \left[\left(\frac{r}{R_e}\right)^{n^{-1}} - 1\right]\right\} \ ; \ b_n = (2 \cdot n) - 0.327. \tag{25}$$

By making the change of variable $x = b_n(r/R_e)^{n^{-1}}$ in Eq. (24), it yields the following result

$$F = 2\pi n \ q \ I_e R_e^2 \frac{e^{b_n}}{(b_n)^{2n}} \Gamma(2n), \tag{26}$$

where $q = 1 - \varepsilon$. It is important to note that the units of I_e in this process have to be in
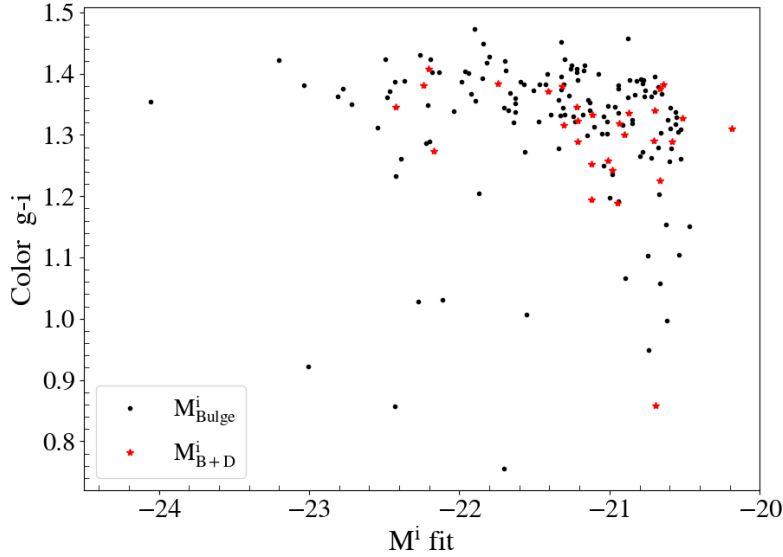
Fig. 18: Color-magnitude diagram for the analyzed galaxies of the cluster A2142.

nanomaggy to be consistent with the next step. According to the SDSS photometric system, the apparent magnitudes can be obtained from the following expression

$$m_{\mathrm{B}}^{i} = 22.5 \ [\mathrm{mag}] - 2.5\log_{10}(F). \tag{27}$$

So, the absolute magnitude can be obtained from

$$M_{\mathrm{B}}^{i} = m_{\mathrm{B}}^{i} + 5 - 5\log_{10}(d). \tag{28}$$

Where d are the distances of the galaxies. They can be easily calculated from the Hubble-Lemaître law, since these galaxies are in a cluster found in the local universe and the dominant metric allows us to ignore the contributions of the cosmological parameters, so

$$d[\mathrm{Mpc}] = \frac{c}{H_0}z, \tag{29}$$

where $c = 3 \cdot 10^5$ km/s is the speed of light in a vacuum, $H_0 = 71.39$ (km/s)/Mpc is the Hubble constant and z is the redshift of the cluster. We point out that we use a common redshift for all galaxies, in order to minimize the peculiar velocity effects of each one of them.

To calculate the masses of the galaxies, the work of Costantin et al. (2020) is followed, taking into account the empirical relation between the color $(g - i)$, the luminosity of the band-i and the stellar mass

$$\log_{10}(M_*/M_\odot) = -0.68 + 0.7(g - i) - 0.4(M_{\mathrm{B}}^{i} - 4.58). \tag{30}$$

The SDSS model magnitude is obtained by comparing two two-dimensional models, a de Vaucouleurs model, defined as

$$I(r) = I_0 \exp\left\{-7.67[(r/R_e)^{1/4}]\right\}, \tag{31}$$

and an exponential profile, and it is compared between both models and they are left with the most suitable one.

Considering the results obtained in **Fig. 18** it can be seen that most of the galaxies, by the definition of the selected sample, are red. Therefore, in addition to morphology (as seen
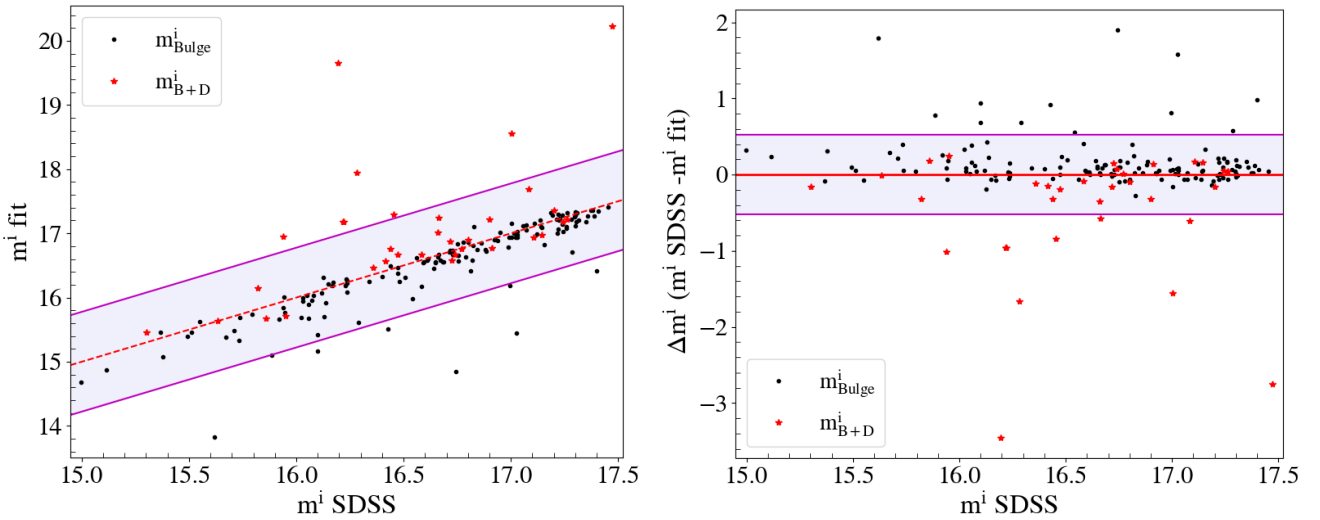
Fig. 19: On the left, comparison between the apparent magnitudes adjusted with ANDURYL vs the apparent model magnitudes given by SDSS, in the "i" filter for real galaxies. On the right the dispersion of these same magnitudes can be seen. The lavender shade indicates the range of the standard deviation.

in section. 5.1), the colors also support that most of the galaxies in our study are of early type. **Fig. 19** compares the magnitudes obtained through the models with the SDSS model magnitudes, and it can be seen that the magnitudes that have been obtained are in good agreement with those of the SDSS, within the range of the standard deviation. However, it seems that for some galaxies fitted with a Sérsic+exponential profile, these turn out to be less bright than they should; while for some galaxies fitted with the Sérsic profile they are brighter than they should be. The galaxies that are within the standard deviation range can be understood to have different values because the calculation of the magnitudes has been carried out in different ways, with different profiles, and also applying different PSFs, so it is reasonable to think that these factors may lead to a slight deviation from the values. However, for those galaxies that deviate beyond the range of the standard deviation, it must be remembered that there are a number of galaxies whose posterior on some parameters, such as the Sérsic index, is very close to the limits of the established priors, making the inference dubious. These galaxies therefore have a greater dispersion when compared to the SDSS values.

With this result it seems appropriate to carry out the conversion to absolute magnitudes, and from there to masses, following the procedure previously explained. Once this is done, the Faber-Jackson relation, using the mass of the bulges of the fitted components, obtained in this work can be shown together with the one obtained in Costantin et al. (2020) in **Fig. 20**. The data provided by this article ranges from globular clusters and nuclear stellar clusters, compact early type galaxies and ultra compact dwarfs to ellipticals and large bulges. We have chosen to represent all the data together to have a wide range of masses in which to represent the Faber-Jackson relation, and to see if the results we have obtained in this work agree with what would be expected from this relation. **Fig. 20** shows that the galaxies studied in cluster A2142 are mostly massive galaxies, with ranges between 10^{10} and 10^{12} solar masses, which is compatible with the sample selection, since SDSS and Liu et al. (2018) selected only the brightest galaxies, and with the fact that most of the galaxies analyzed have been classified as ellipticals, which are the most massive type of galaxies. Some points, marked in red in the figure, fall outside the ranges of standard deviations provided, and generally outside the different groups of galaxies. These points show velocity dispersions below the spectral resolution limit of SDSS (~ 70 km/s). It is probable that this has biased the information of these galaxies and we considered appropriate not to include them in the treatment or in the analysis.
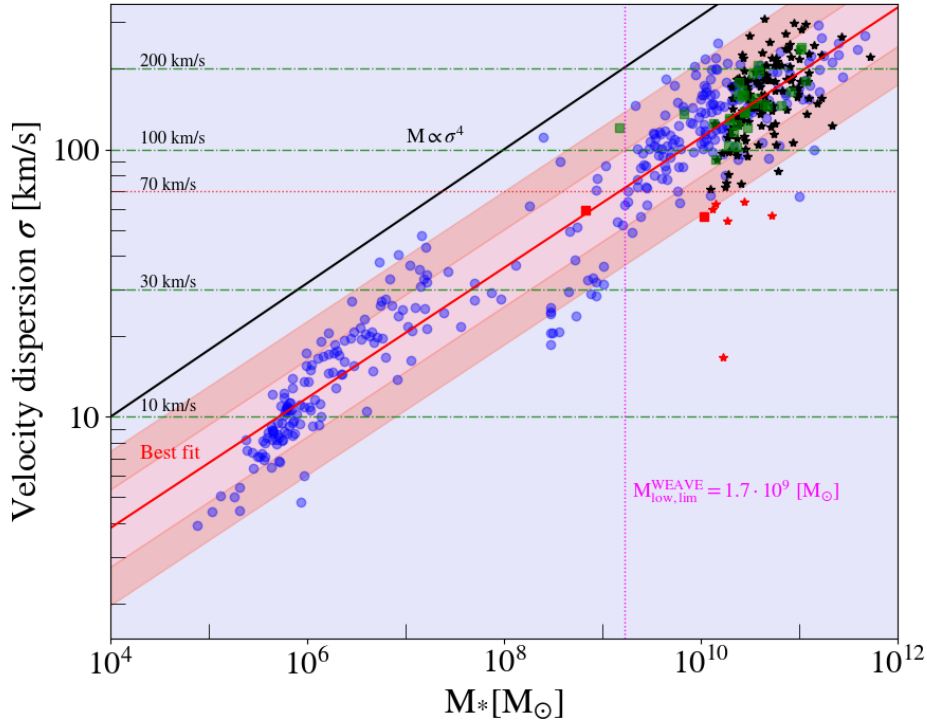
Fig. 20: Representation of the Faber-Jackson relation. The solid black line represents the relation $M \propto \sigma^4$ and the red dotted line the best fit for all the data taken together. The velocity dispersion resolution limit of SDSS is also indicated, which is 70 km/s and the mass of a galaxy with apparent magnitude 20 is represented in a vertical line, which will be the lower limit of WEAVE Cluster Survey exploration. The round blue dots indicate the Faber-Jackson of Costantin et al. (2020) work, the black star-shaped dots are the galaxies fitted with a Sérsic profile and the green square-shaped dots are the galaxies fitted with a Sérsic+exponential profile. Red dots indicate those that were hearing below the SDSS resolution limit. The red shadows around the best fit indicate the errors for each point, corresponding to σ_{err} and $2\sigma_{err}$.

By fitting all the points, a straight line has been obtained whose slope is $m = 0.243 \pm 0.03$, which is in good agreement with what is expected, which is a slope of 0.25 ($\sigma \propto M^{1/4}$). Therefore, our inferences follow the Faber-Jackson relation with good agreement, so that that they represent virialized systems. Since this work also represent a preparatory analysis for the forthcoming WEAVE Cluster Survey, the lower limit in mass for WEAVE, which correspond to an apparent magnitude of 20 in the "i" band (this has been chosen as an estimate of the limit in magnitude of WEAVE) obtaining that $M_{low,lim}^{WEAVE} = 1.7 \cdot 10^9 \, M_{\odot}$. As a consequence, WEAVE is in position to be able to observe the lowest mass galaxies. In order to calculate the mass limit, a color $(g - i)$ representative of the entire sample has been chosen. It must be taken into account that this limit may vary according to the redshift in which the studied cluster is located.

Using the inferred mass of the bulges, **Fig. 21** shows the mass-size relation. First of all, it can be seen that the points analyzed with ANDURYL follow the proper trend observed in the Méndez-Abreu et al. (2021) data. It must be taken into account that the study in Méndez-Abreu et al. (2021) used galaxies from the CALIFA survey that were field galaxies, and not cluster galaxies, so they are galaxies with different morphological types. Even with this difference, it is seen that the galaxies studied for cluster A2142 follow the same scale relationship as the field galaxies. However, it can be seen that there is a slight offset and that the points fitted with ANDURYL are slightly deviated from the mass-size relation, although within the deviation ranges of the fits. If the parameters of the obtained fit lines are analyzed, and they are compared with the results obtained in Méndez-Abreu et al. (2021) (see **Table. 4**) it can be seen that reasonably similar numbers are
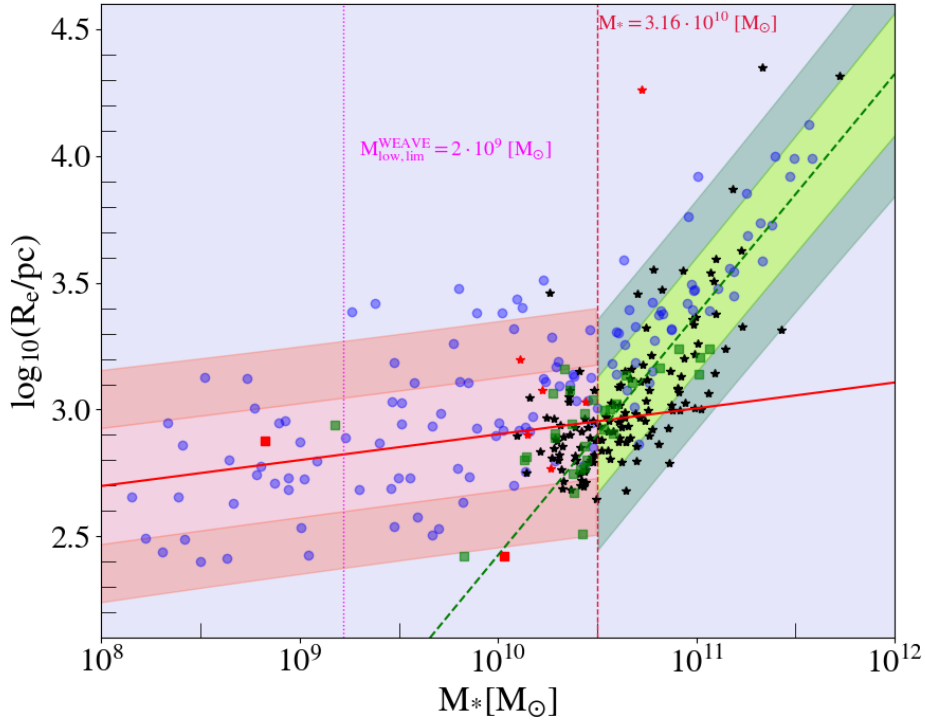
Fig. 21: Representation of the mass-size relation. It has been decided to divide the analysis into two mass bins, taken from Méndez-Abreu et al. (2021), that are $8 < \log_{10}(M_*/M_\odot) < 10.5$ and $10.5 < \log_{10}(M_*/M_\odot) < 12$, so that the changes in the slope can be appreciated. The solid red line represents the best fit for the first mass bin, and the dotted green line the best fit for the second bin. The mass of a galaxy of apparent magnitude 20 is represented in a vertical magenta line in the "i" filter, which will be the lower limit of WEAVE exploration. The round blue dots indicate the galaxies from Méndez-Abreu et al. (2021) work, the black star-shaped dots are the galaxies fitted with a Sérsic profile and the green square-shaped dots are the galaxies fitted with a Sérsic+exponential profile. Red dots indicate those that were hearing below the SDSS resolution limit of 70 km/s. The red and green shadows around the best fits indicate the errors for each point, corresponding to σ_{err} and $2\sigma_{err}$.

obtained between both works, within the margins of error. From this mass-size relation it can be inferred that early-type galaxies, generally more massive, have larger bulges; although a study at lower masses would be needed to be able to detect if there is another change in slope in the environment of dwarf galaxies. This change in slope is the reason why it was decided to have split the data into two mass bins, since a full fit would be unsuccessful.

Table 4: Comparison of best fit values for the mass-size relation between ANDURYL and Méndez-Abreu et al. (2021)[1]. a and b are the coefficients of the fits $\log_{10}(R_e) = a + b\log_{10}(M_*/M_\odot)$.

$\log_{10}(M_*/M_\odot)$	a	b
8-10.5	1.88 $\pm$0.27	0.10 $\pm$0.03
10.5-12	-7.06 $\pm$0.78	0.95 $\pm$0.08
$\log_{10}(M_*/M_\odot)$[1]	a	b
8.0 - 10.5	0.88$\pm$0.48	0.20$\pm$0.05
10.5 - 12.0	-6.2$\pm$0.39	0.87$\pm$0.04

Another analysis that can be carried out is how the morphological fractions are distributed along the clustercentric radius, that is, the distance of the galaxies with respect to the center of the cluster, which is taken as the place where the Brightest Cluster Galaxy (BCG) is located (Dressler 1979). This analysis can be seen in **Fig. 22**. Despite the fact that there are several
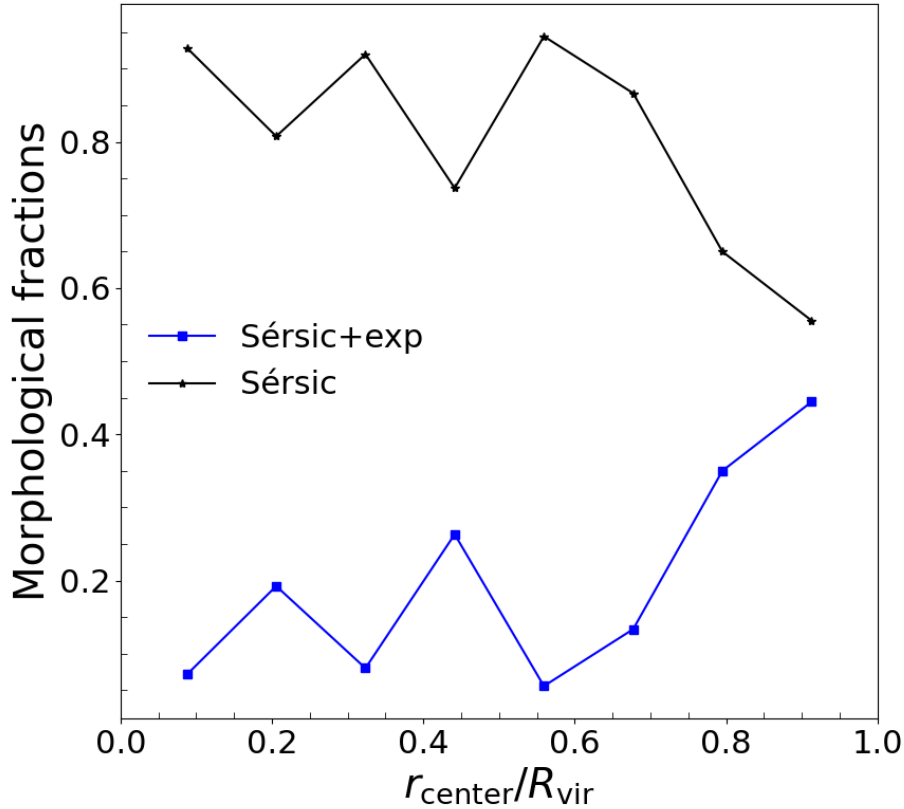
Fig. 22: Representation of the morphological fractions obtained for the galaxy cluster A2142 versus the quotient between the clustercentric distance and the virial radius of the cluster, $R_{\mathrm{vir}} = 0.9$ Mpc h^{-1}. In black stars the galaxies adjusted with a Sérsic profile are represented, and in blue squares those adjusted with a Sersic + exponential profile are shown.

caveats due to the selection made, a certain tendency can be seen in which the fraction of elliptical galaxies decreases with respect to the disc galaxies as the distance is greater with respect to the center of the cluster. This study has focused, on radial scales, within the virial radius of the cluster, which is $R_{\mathrm{vir}} = 0.9$ Mpc h^{-1} (Einasto et al. 2018), since this is the densest region, and it is not expected to find galaxies further away, results are expected mainly in this range of radius. Moreover, the physical mechanisms that could explain this trend seem to require this density to take place (Vulcani et al. 2023).

The results that have been obtained are consistent with the literature, since it is seen that at distances closer to the center, that is, in the densest regions, elliptical galaxies dominate (Dressler 1980). Even with a majority of red galaxies, the separation between disc and elliptical galaxies still allows us to observe some differences.

Chapter 6

Conclusions and future work

In this work we have developed from scratch a new photometric decomposition method based on Bayesian statistics. This is a major step forward in this kind of analysis since it overcomes a number of problems inherent to least squares fitting codes. Some of the most notable improvements are:

- We have now access to the full posterior distribution. This translates into a better understanding and quantification of the uncertainties. There is a high degree of adaptability to new information, and inferences can be made flexibly. The errors on each parameter have been calculated from the sampling of the posterior distribution using a Bayesian method, and therefore they fully include covariances.

- It includes all the observational effects necessary for a photometric decomposition: convolution of the PSF, oversampling of the grid, inclusion of noise in the fit, etc.

- It has been possible to eliminate the need to provide initial conditions to the fitting. Therefore, it is less dependent on the user and, in addition, this results in a release of the workload of carrying out a previous analysis of each galaxy in order to estimate an initial value close to the solution.

- The problems related to the inversion of components has been minimized, thanks to the use of combined (bulge and disc) parameters in the priors, therefore avoiding unrealistic solutions.

Therefore, the work has fulfilled the objective of creating a code that could face the usual problems of a least squares fitting code. The computational time takes around 2-5 minutes to fit a Sérsic profile and around 15 minutes to fit a Sérsic+exponential profile, analyzing a galaxy in each processor core, although this also depends on the particular case of each galaxy. The code is already available on GitHub[1], having a modular structure that facilitates the integration of other photometric components.

The code has been solidly tested using mock galaxies, having a set of mock galaxies for each photometric profile that can be adjusted with ANDURYL. It was found that the code can fit models practically identical to the mock galaxies, while it was verified that the BIC separation criterion can distinguish the components correctly, indicating in which cases a galaxy can be better modeled with one model or another. The first science results derived with ANDURYL have also been satisfactory. It has been possible to carry out a preliminary study of the galaxies in the A2142 cluster and calculate the WEAVE Cluster Survey limit mass for this case. Our conclusions are:

- Most of the galaxies analyzed in cluster A2142 are massive galaxies with masses greater than 10^{10} solar masses. This is related to the sample selection, based on spectroscopically selected

---

[1] https://github.com/CarlosMDLR/ANDURYL

galaxies, in relatively shallow clusters. The WEAVE Cluster Survey will allow to explore the low mass galaxies in clusters, therefore extending the sample up to the regime of dwarf galaxies.

- Of the 175 galaxies successfully fitted, approximately 80% fit with a Sérsic profile and 20% with an Sérsic+exponential profile, indicating that most of the population of this cluster are early-type galaxies, which is in good agreement with literature.

- The magnitudes derived from the fitting coincided quite well with those given by SDSS, which represent another test that probes that our new method is robust.

- The Faber-Jackson and the mass-size relations have been derived, finding that the galaxies analyzed with ANDURYL follow the general trends also find in the literature.

- The morphological fractions of the cluster have also been calculated, finding that the elliptical-type galaxies decreased as we moved away from the center of the cluster, that is, from the densest areas. This is in good agreement with the predictions from the morphology-density relation.

The future prospects of this work come in different ways: firstly, add more stuctural components to the code, such as a Ferrers profile to be able to fit bars. We will also test different priors and physical limits that make the model more in accordance with observations, investigate other possible sampling methods beyond Nested Sampling and improve the computation time by adapting the code to work in supercomputer environments. The implementation of a Hierarchical Bayesian method is also proposed to derive the scaling relations in a self-consistent manner with the fit of each individual galaxy. The analysis of larger samples of galaxies (~ 1000) of a larger sample of clusters (~ 50) reaching the regime of dwarf galaxies, using the WEAVE Cluster Survey, is proposed.

On a personal level, doing this work has meant an improvement in my knowledge of Bayesian statistics, in how to program more optimally in Python, and in how to deal with code bugs. And it has improved my knowledge about the morphological analysis of galaxies.

References

Abraham, R. G., Valdes, F., Yee, H. K. C., & van den Bergh, S. 1994, , 432, 75

Ahumada, R., Allende Prieto, C., Almeida, A., et al. 2020, , 249, 3

Allen, P. D., Driver, S. P., Graham, A. W., et al. 2006, , 371, 2

Argyle, J. J., Méndez-Abreu, J., Wild, V., & Mortlock, D. J. 2018, , 479, 3076

Barbary, K., Asensio Ramos, A., Speagle, J., et al. 2023, GitHub repository

Bayes, M. & Price, M. 1763, Philosophical Transactions of the Royal Society of London Series I, 53, 370

Boselli, A. & Gavazzi, G. 2014, , 22, 74

Bowen, I. S. & Vaughan, A. H., J. 1973, , 12, 1430

Clauwens, B., Schaye, J., Franx, M., & Bower, R. G. 2018, , 478, 3994

Cobb, A. D., Güneş Baydin, A., Markham, A., & Roberts, S. J. 2019, arXiv e-prints, arXiv:1910.06243

Cobb, A. D., Güneş Baydinand, A., Solin, A., & Cranmer, M. 2023, GitHub repository

Conselice, C. J. 2003, , 147, 1

Cooley, J. W. & Tukey, J. W. 1965, Mathematics of computation, 19, 297

Costantin, L., Méndez-Abreu, J., Corsini, E. M., et al. 2020, , 889, L3

de Vaucouleurs, G. 1959, Handbuch der Physik, 53, 275

Djorgovski, S. & Davis, M. 1987, , 313, 59

Domínguez Sánchez, H., Huertas-Company, M., Bernardi, M., Tuccillo, D., & Fischer, J. L. 2018, , 476, 3661

Dressler, A. 1979, , 231, 659

Dressler, A. 1980, , 236, 351

Duane, S., Kennedy, A. D., Pendleton, B. J., & Roweth, D. 1987, Physics Letters B, 195, 216

Einasto, M., Deshev, B., Lietzen, H., et al. 2018, , 610, A82

Erwin, P. 2015, , 799, 226

Faber, S. M. & Jackson, R. E. 1976, , 204, 668

Feroz, F., Hobson, M. P., & Bridges, M. 2009, , 398, 1601

Ferrers, N. M. 1870, Philosophical Transactions of the Royal Society of London Series I, 160, 1

Freeman, K. C. 1970, , 160, 811

Gao, H. & Ho, L. C. 2017, , 845, 114

Graham, A. W. & Driver, S. P. 2005, , 22, 118

Gunn, J. E., Carr, M., Rockosi, C., et al. 1998, , 116, 3040

Gunn, J. E., Siegmund, W. A., Mannery, E. J., et al. 2006, , 131, 2332

Hoffman, M. D. & Gelman, A. 2011, arXiv e-prints, arXiv:1111.4246

Hubble, E. P. 1926, , 64, 321

Hubble, E. P. 1936, Realm of the Nebulae

Huertas-Company, M., Rouan, D., Tasca, L., Soucail, G., & Le Fèvre, O. 2008, , 478, 971

Kopylova, F. G. & Kopylov, A. I. 2022, arXiv e-prints, arXiv:2205.02497

Kravtsov, A. V. & Borgani, S. 2012, , 50, 353

Lange, R., Moffett, A. J., Driver, S. P., et al. 2016, Monthly Notices of the Royal Astronomical Society, 462, 1470

Liu, A., Yu, H., Diaferio, A., et al. 2018, , 863, 102

Méndez-Abreu, J., Aguerri, J. A. L., Corsini, E. M., & Simonneau, E. 2008, , 478, 353

Méndez-Abreu, J., Aguerri, J. A. L., Falcón-Barroso, J., et al. 2018, , 474, 1307

Méndez-Abreu, J., de Lorenzo-Cáceres, A., & Sánchez, S. F. 2021, , 504, 3058

Méndez-Abreu, J., Ruiz-Lara, T., Sánchez-Menguiano, L., et al. 2017, , 598, A32

Moore, B., Ghigna, S., Governato, F., et al. 1999, , 524, L19

Mukherjee, P., Parkinson, D., & Liddle, A. R. 2006, , 638, L51

Murray, I., MacKay, D., Ghahramani, Z., & Skilling, J. 2005, in Advances in Neural Information Processing Systems, ed. Y. Weiss, B. Schölkopf, & J. Platt, Vol. 18 (MIT Press)

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2010, , 139, 2097

Petrosian, V. 1976, , 210, L53

Schwarz, G. 1978

Sérsic, J. L. 1963, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 6, 41

Simard, L., Mendel, J. T., Patton, D. R., Ellison, S. L., & McConnachie, A. W. 2011, , 196, 11

Skilling, J. 2004, in American Institute of Physics Conference Series, Vol. 735, Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, ed. R. Fischer, R. Preuss, & U. V. Toussaint, 395–405

Tacchella, S., Dekel, A., Carollo, C. M., et al. 2016, Monthly Notices of the Royal Astronomical Society, 458, 242

van den Bergh, S. 1976, , 206, 883

Vulcani, B., Poggianti, B. M., Aragón-Salamanca, A., et al. 2011, , 412, 246

Vulcani, B., Poggianti, B. M., Gullieuszik, M., et al. 2023, , 949, 73

WEAVE. 2023

Appendices
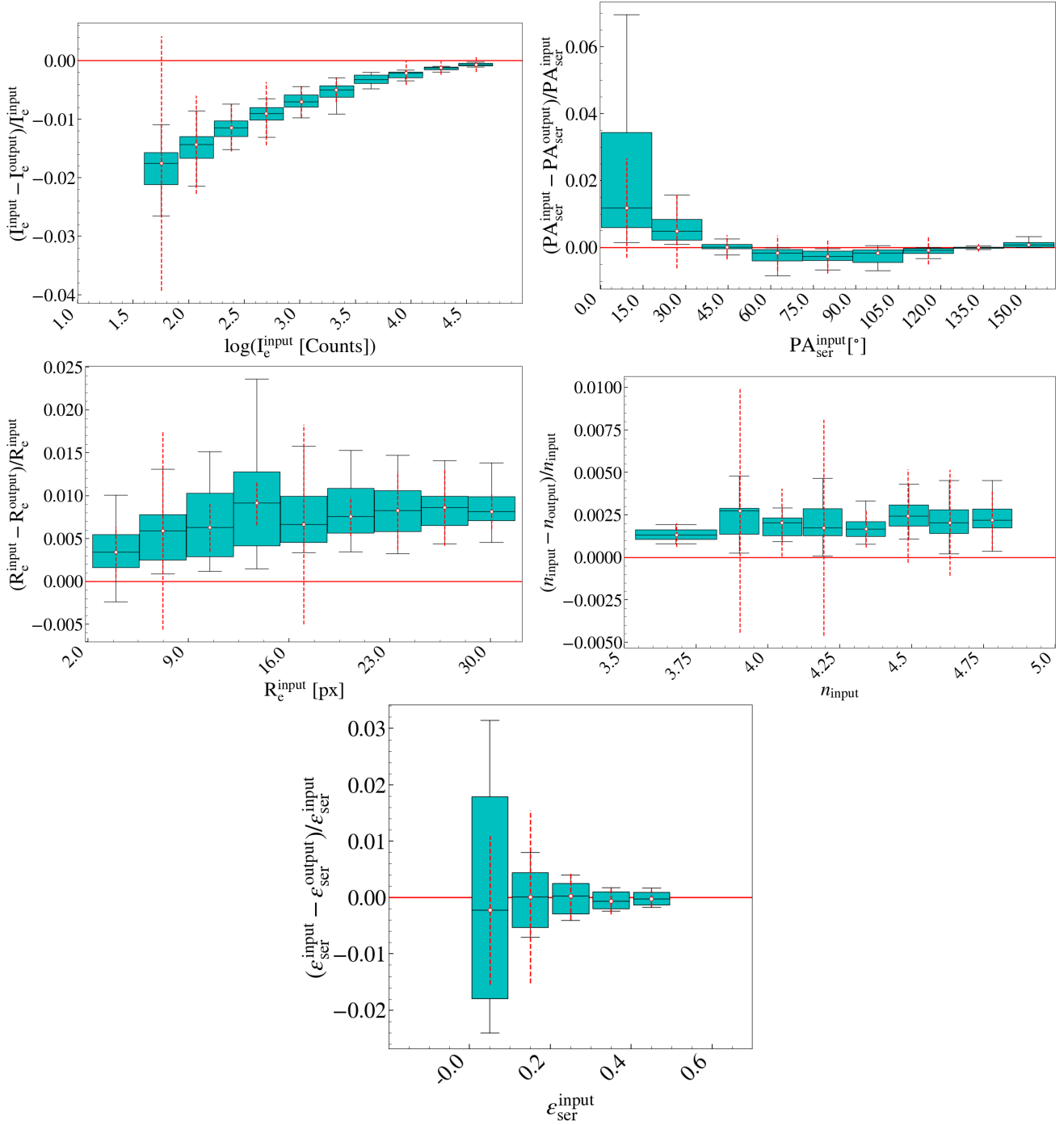
Appendix A: Box plots of mock galaxy samples

Fig. A.1: Different box plots for the parameters resulting from fitting the sample of bulge mocks with a Sérsic+Exponential profile. The box extends from the first quartile (Q1) to the third quartile (Q3) of the data, with a black line at the median. The whiskers extend from the box by a factor of 1.5 the inter-quartile range (IQR, between Q1 and Q3). The red dotted lines represent the mean of the error, given by ANDURYL, of all the points that are inside each box.
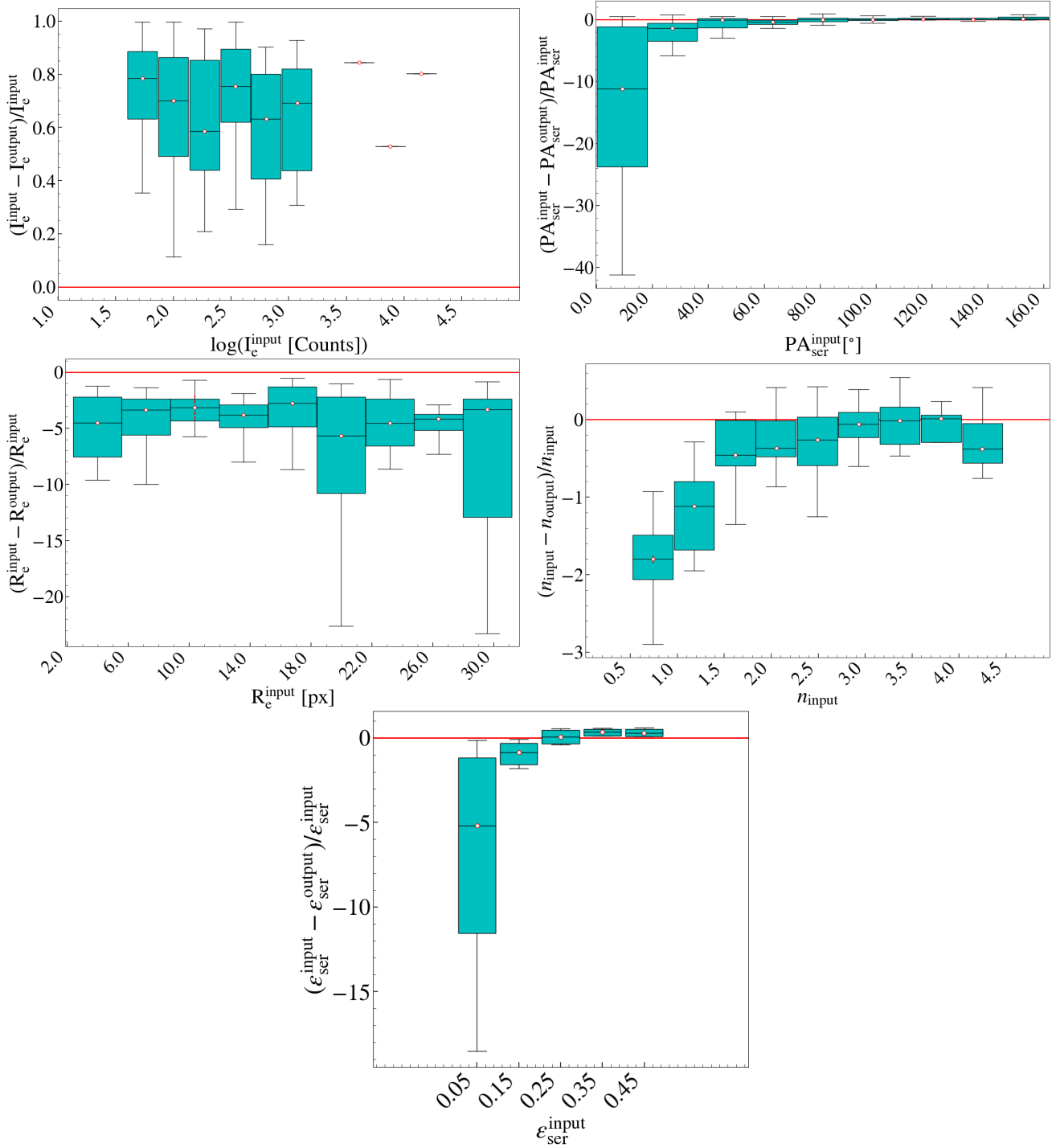
Fig. A.2: Different box plots for the parameters resulting from fitting the sample of bulge-disc mocks with a Sérsic profile. The box extends from the first quartile (Q1) to the third quartile (Q3) of the data, with a black line at the median. The whiskers extend from the box by a factor of 1.5 the inter-quartile range (IQR, between Q1 and Q3). The red dotted lines represent the mean of the error, given by ANDURYL, of all the points that are inside each box.
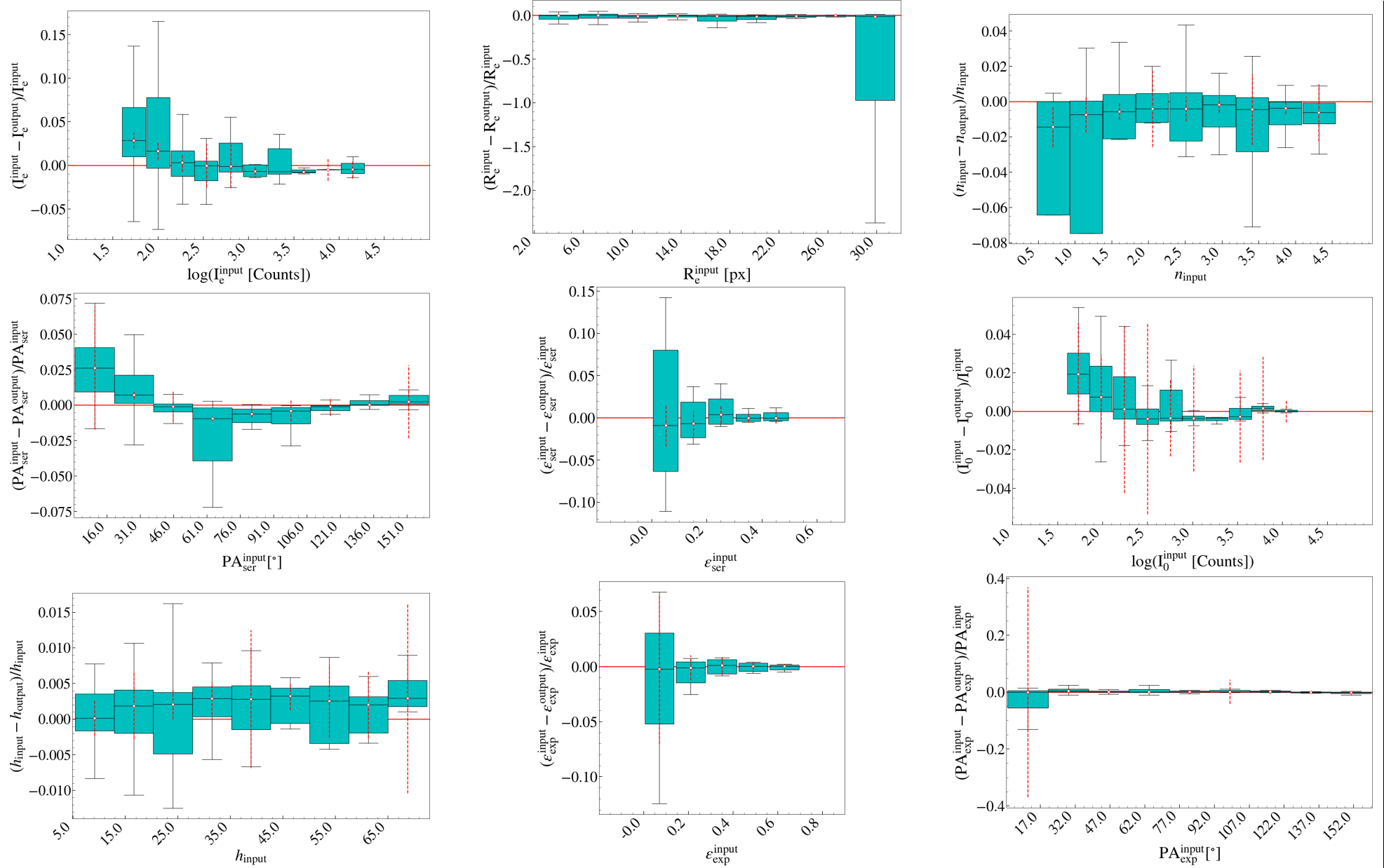
Fig. A.3: Different box plots for the parameters resulting from fitting the sample of bulge-disc mocks with a Sérsic+Exponential profile. The box extends from the first quartile (Q1) to the third quartile (Q3) of the data, with a black line at the median. The whiskers extend from the box by a factor of 1.5 the inter-quartile range (IQR, between Q1 and Q3). The red dotted lines represent the mean of the error, given by ANDURYL, of all the points that are inside each box.
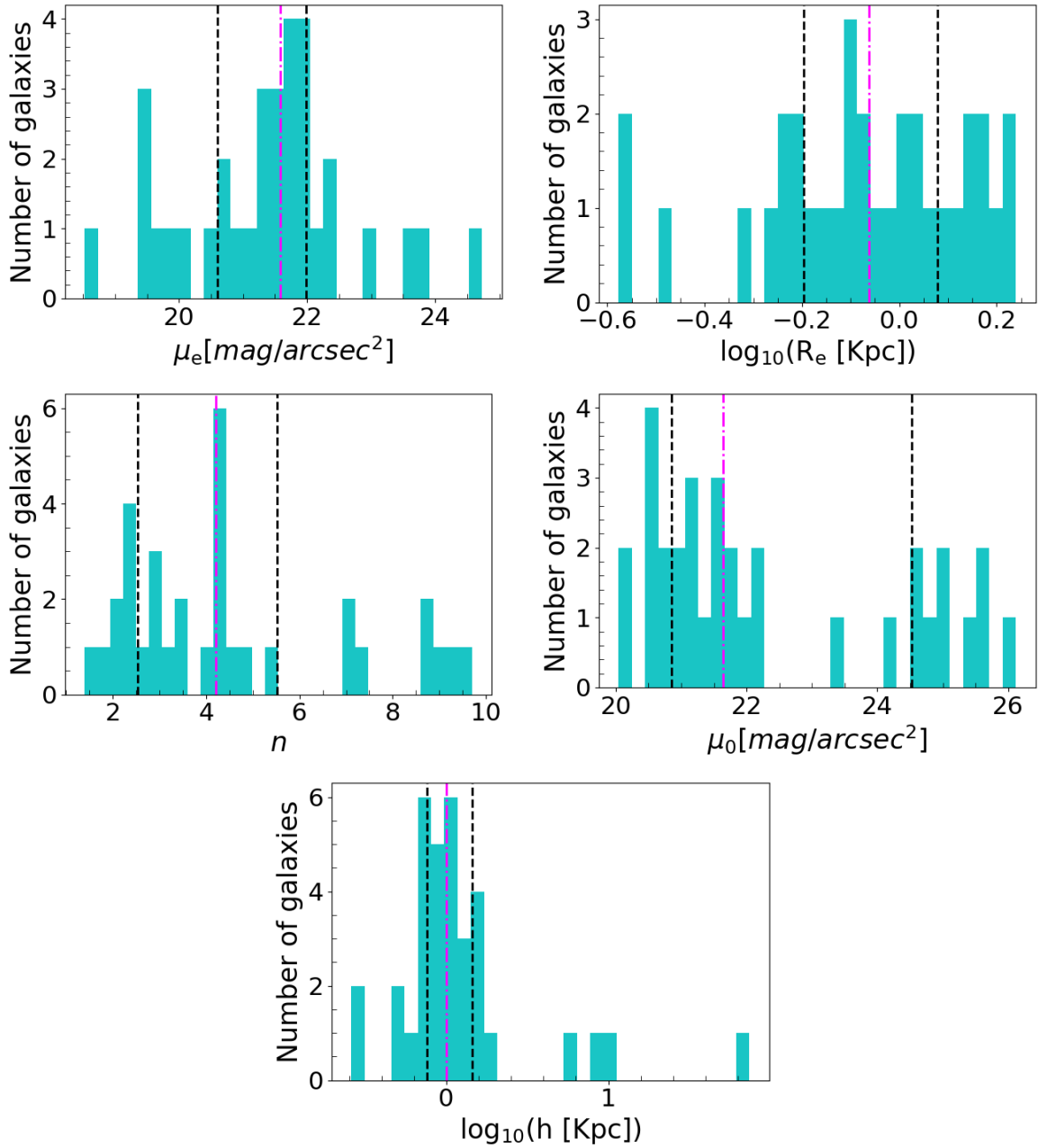
Appendix B: Histograms of the parameters

Fig. B.1: Distribution of the parameters of the galaxies fitted with a Sérsic+exponential profile. In order from left to right, effective surface brightness μ_e, effective radius R_e, and Sérsic index n. The magenta vertical lines represent the mean of the distribution, and the black vertical lines the Q1 and Q3 quartiles.