DETERMINACIÓN DE CÚMULOS DE GALAXIAS MEDIANTE MEDIDAS DE EFECTO DE LENTE GRAVITATORIO DÉBIL EN IMÁGENES DEL TELESCOPIO SUBARU

(Determination of Galaxy Clusters Through Weak Lensing Measurements on Images from Subaru Telescope)

Trabajo de Fin de Grado para acceder al GRADO EN FÍSICA

Autor: Guadalupe Cañas Herrera
Director: Patricio Vielva Martínez
Co-Director: Ian Dell’Antonio

Marzo - 2016
To my parents and Juanchito,
whose love for me
is brighter than any starlight.
ACKNOWLEDGEMENTS

First, I would like to thank my enjoyable supervisor professor Patricio Vielva, for his advices and motivation in the whole process. I extend my gratitude also to professor Ian Dell’Antonio who gave me the possibility of going into the exciting world of Gravitational Lensing. Furthermore, special thanks goes to the Faculty of Science at the University of Cantabria, for making possible my stay at Brown University in the Department of Physics.

Moreover, I have appreciated the scientific atmosphere in both the laboratory at Brown University and at the Observational Cosmology and Instrumentation group from Instituto de Física de Cantabria. Thanks also to Silvia Mateos for listening to me and looking for proper surveys in x-rays to check my candidates for galaxy clusters with the help of Alexis Finoguenov.

I have to say that I feel nothing but gratitude for my partner and soul-mate Juan Claramunt. His love, support and constant dedication in his work has made me grown during this journey as an example of self-improvement.

Para acabar, a mis padres, sólo les puedo decir gracias, con todo el extenso significado que esa palabra conlleva. Gracias por nunca decirme que no a mis sueños, gracias por enfocar todos vuestros esfuerzos a que sea feliz.
# Table of Contents

Abstract 1

Preface 2

1 Theoretical Background 4
   1.1 Cosmological Standard Model 4
   1.2 Dark Matter 6
      1.2.1 Dark Matter Candidates 7
      1.2.2 Alternative theories of Gravitation 9
   1.3 Gravitational Lensing 10
      1.3.1 Lensing Equations 10
      1.3.2 Weak Lensing 14

2 Data Description 17
   2.1 Subaru Telescope 17
   2.2 Characteristics of the Fields 18
   2.3 Deep Lens Survey 18

3 Estimation and Validation 20
   3.1 Estimation of the Catalogues 20
   3.2 Validation of the Catalogues 21
      3.2.1 Cross-Matching of Catalogues 22
      3.2.2 Add-Objects Simulation 23
   3.3 Further cleaning of spurious objects 25

4 Correction of systematics and galaxy identification 26
   4.1 Division in galaxies and stars 26
      4.1.1 Shapes Estimation 30
      4.1.2 Correction of the PSF 31
   4.2 After correction 34

5 Dark Matter Halo Estimation 36
   5.1 Convergence Maps 36
   5.2 Noise Estimation of the Convergence Maps 38
      5.2.1 Importance of PSF anisotropies correction 39

6 Comparison with previous observations 41
   6.1 Galaxy Cluster Candidates 41
   6.2 Looking For Counterparts 42
      6.2.1 X-ray Observations 42
Photographies from Subaru Telescope of Lockman Hole area have been studied, obtaining catalogues of galaxies and stars with the aim of estimating galaxy clusters from the shear associated to shapes of galaxies produced by dark matter through the weak lensing effect. To obtain these catalogues, it was necessary to carry out a correction between images in the optical and in the near infrared. The catalogues have been filtered and corrected with the purpose of eliminating point spread function anisotropies related with the telescope due to difference sources such as the atmosphere. The results are dark matter distribution maps as a function of the galaxies magnitudes. A list of possible galaxy clusters has been found in this field, identifying these candidates with previous observations performed in x-ray and with other works also based on weak lensing effect.

**KEYWORDS:** Cosmology, Dark Matter, Compact Object Detection, Weak Gravitational Lensing, Galaxy Clusters

---

Se han estudiado fotografías tomadas con el Telescopio Subaru del area conocida como el agujero de Lockman, obteniendo catálogos de galaxias y estrellas con el fin de estimar cúmulos de galaxias a partir de la deformación de la apariencia de las mismas producido por materia oscura a través del efecto de lente gravitatorio débil. Para obtener dichos catálogos, ha sido necesario realizar una corrección entre imágenes correspondientes al óptico y al infrarrojo cercano. Los catálogos se han filtrado y se han corregido con el fin de eliminar las anisotropías de la función de respuesta angular del telescopio, debidas entre otras causas a la atmósfera. Los resultados son mapas de distribución de materia oscura en función de las magnitudes de las galaxias. Se ha encontrado una lista de posibles candidatos de cúmulos de galaxias en este campo, identificando dichos candidatos con observaciones previas tanto en rayos X como en trabajos también basados en el efecto de lente gravitatorio débil.

**PALABRAS CLAVE:** Cosmología, Materia Oscura, Detección de Objectos Compactos, Efecto de Lente Gravitatorio Débil, Cúmulos de Galaxias
This Final Degree Project is based on the data analysis of images taken by the telescope Subaru in the optical and near infrared. The images show galaxies placed in the Lockman Hole area, together with some stellar contamination. The main goal is the detection of galaxy clusters, and therefore, of the corresponding Dark Matter (DM) distribution, through the impact Weak Lensing (WL) effect have on the morphology of the galaxies.

Gravitational lensing is the result of the bending of the light emitted by a given source, as it travels toward the observer caused by a distribution of mass (DM halos) placed between that source and the observer. When this mass distribution is concrete and very dense, the effect is known as strong, whose most popular examples are, perhaps, the Einstein rings and crosses. In addition light may also suffer subtle distortions not as extreme as previous cases. This is what we know as WL effect. WL enables to study DM distribution in most parts of the immense universe in a very direct way. It could provide evidence of already-known mass concentrations or over-densities and their matching with galaxy clusters, or even proof new mass concentrations otherwise in darkness. Therefore, WL serves to answer questions such as how galaxy clusters contribute to the mean density of matter in the universe, or how the universe looks like on larger scales, giving an idea about its structure. Furthermore, this effect could be used to constrain alternative theories of gravity. To pursue the already mentioned motivations, essential improvements in astronomical instruments as well as in data analysis techniques are needed. First, new instruments such as CCDs arrays provide images with more resolutions and sensitivity. Second, the increment in data allows to decrease the statistical uncertainties, making indispensable to take care of systematic errors during the analysis workflow.

This project report is divided into seven different chapters. In chapter 1, a brief overview of the currently accepted Cosmological Model is given, making emphasis on DM: its nature and evidences, as well as its distribution in the universe. In addition, a formal introduction to gravitational lensing, its mathematical modelling and the WL limit can be found at the end of the chapter. Furthermore, the experimental framework, meaning the observational instruments and the characteristics of the image fields are discussed in chapter 2. Following, the data acquisition mechanism is explained in chapter 3, together with used validation techniques. In chapter 4, the correction of the fields is discussed as well as the selection process to distinguish between galaxies and stars. Finally, the analysis method to estimate mass over-densities is depicted in chapter 5, providing information about the implemented analysis algorithm. The whole research process is described in chapters 1, 2, 3, 4 and 5 with the help of one example focused in one of the studied fields. The comparisons of the cluster candidates with previous observations are stated in chapter 6, and finally, in chapter 7 the reader may find the conclusions of the project.

In order to keep this thesis to manageable proportions, an appendix have been included. It acts as an independent software documentation including information about the computational programs and
scripts to perform the data analysis as well as the dependencies to libraries and packages. Thus, the purpose of Appendix A is to be useful for future initial investigations in WL so that the main scripts can be reused. Part of the software has been inherited from the Deep Lens Survey (DLS) group, from which the co-director of this project, Professor Dell’Antonio, takes part in. The already-written code by the DLS group that is used in this project is also commented in Appendix A.

This Final Degree Project presents an approach to WL, which is one of the hot topics in astrophysics at the moment. It includes data analysis steps which are performed in this kind of research work, from the measurements of compact objects deformations using astronomical tools to implementations of computer algorithms in different languages. The aim of this report is to give self-contained explanations of the carried-out research work, including necessary concepts, terminology or already developed software.

So much I have learnt from this project since its very beginning. Every step in the research process was a challenge concerning the understanding and refreshing of cosmological concepts and the increment of my computer management skills. I consider this project the crowning moment of my Bachelor’s education, exposing not only the acquired knowledge during the degree but the capacity of accomplishing independent work either in the search of information or in the resolution of problems.

Among the obtained skills during this project, I would like to remark the study of astrophysics and cosmology, explicitly in relation to the understanding of the formalism of gravitational lensing, as well as the learning of useful concepts in statistics. Moreover, I needed to increase my expertise in general computation. I dedicated time to the learning of a new object-oriented programming language, PYTHON, and also to the understanding of scripting languages, for instance BASH and PERL. Besides, I learnt how to compile properly of C written software, and so do I achieved the required knowledge to manage a supercomputer such as ALTAMIRA. Finally, I have to point out the proficiency gain in the search of trustable appropriate bibliographical sources, not only during the study of theoretical concepts but during the programming time too. Additionally, I got hands on experience in the planning and development of a research project, from its conception until the presentation of the results.
1.1 Cosmological Standard Model

During years, cosmologists have tried to narrate the biography of the universe with the purpose of establishing its structure, properties and evolution along time. The results is the Standard Model of Cosmology. The universe as we see it nowadays looks the same along any direction, assuming the cosmological principle. Nevertheless, this is the result of a long physical process. We may have some confidence in the story of its evolution from the time of matter creation to the present, although at earlier times, speculation is dominant.\(^1\)

It is widely believe that the universe was in first instance in a extremely hot and dense state. However, the universe suddenly started to expand around 13.7 billion years ago, decreasing its temperature. This is what we name Big Bang. It seems that approximately in a period from \(10^{-34}\) to \(10^{-32}\) seconds, the universe went through a extremely fast expansion known as inflation. This mechanism\(^2\) explains why the universe is so flat (flatness problems), smooth and regular (smoothness problem) and why it is entirely in thermal equilibrium (horizon problem). This humongous expansion has the dramatic effect of blowing up tiny quantum fluctuations taking place in the primitive universe into macroscopic size inhomogeneities of the energy density. Those quantum fluctuations were roughly equally distributed over all length scales. In other words, the inhomogeneities were produced with equal probability over all wavelengths.

After inflation, the universe behaves like a quantum gas in thermal equilibrium and it is dominated by radiation where all particles in the standard model of fundamental interactions are said to be coupled to the thermal gas. At temperatures higher than \(T \approx 10^{29}\) K everything is expected to be in thermal equilibrium, implying the unification of electromagnetism, weak and strong nuclear forces. Yet, as the universe cooled down due to expansion, it eventually became cold enough to allow the decoupling of diverse particles from the thermal bath. It is supposed that DM was the first type of matter to decouple. Next, at \(T \approx 2 \times 10^{15}\) K, electroweak transition took place when Higgs particle generates mass for the other particles in the Standard Model. Little later, at \(T \approx 10^{15}\) K, it becomes energetically more favourable for quarks to exist in bound states like protons and neutrons instead of being in a quark-gluon plasma. Neutrinos decouple from the cosmic thermal bath around \(T \approx 10^{10}\) K. Below that temperature, gamma photons have no longer enough energy to materialise into electron-positron pair and the population of both electrons and positron starts to decrease steeply by particle-antiparticle annihilation.

During the first 10 minutes the early Universe has reach a temperature cool enough to allow several light elements to be formed allowing a process called primordial nucleosynthesis. This mechanism

\(^1\)For an introduction to the Cosmological Standard Model, the reader may find detailed information in [11].

\(^2\)The introduction of the Inflationary Cosmology was first suggested in 1982 by Alan Guth to solve several problems with the Big Bang model. For further information about this topic, the reader is kindly referred to [21].
is the responsible of the transition between a *radiation dominated universe* and a *matter dominated universe*. At 100000 years old, the universe entered into the *Recombination* epoch, already dominated by matter. During this time, light nuclei began to bound with electrons to form neutral atoms. Due to the decrease in the number density of free electrons because of their bindings to nuclei, the decoupling of photons from matter happened, making the universe transparent to photons everywhere. In fact, we have detected these photons at present with an almost isotropic distribution, estimating that this phenomenon took place when the universe was 380000 years old. This ancient distribution of radiation is what we name *cosmic microwave background* (CMB). Thus, for a temperature $T \approx 10^9$ K, the universe consisted of only four different components meaning photons, dark matter, neutrinos and baryonic plasma containing free electrons, light ions and some neutral atoms. The main consequence of inflation was the enlargement of microscopic oscillation in the youngest universe into macroscopic fluctuations. Due to these oscillations, DM, which decoupled first and had more time to evolve than *baryonic matter* (BM), began to collapse into halos. Subsequently, galaxies and stars began to form at the moment that BM (gas and dust) collapsed to the centre of the pre-existing DM halos.

Currently, it is strongly believed thanks to *Planck Mission*, destined to study the CMB, that the Universe follows an Standard Cosmological Model composed mostly by dark energy $\Lambda$ (a cosmological constant that is hypothesized to be a energy that accelerates the expansion of the universe) and cold dark matter (DM whose particle interacts weakly with the rest of particles and move in a non-relativistic way), receiving the name of *$\Lambda CDM$ Cosmology*. The inferred density quantities for the different components for a flat universe are matter density (including DM and BM) $\Omega_M = 0.315 \pm 0.013$, $\Lambda$ density $\Omega_\Lambda = 0.685 \pm 0.013$ and radiation energy $\Omega_R \approx 0$ [27]. It has been demonstrated that there exist temperature anisotropies within the CMB. These inhomogeneities, of the order of $10^{-5}$ K are the
seeds of the present large scale structure of the Universe, which seems to have an structure with long filaments formed by DM and where clusters of galaxies are located along the DM webs.\footnote{The evolution of the cosmological fluctuations and consequently the creation of the structure we can observe nowadays may be explained thanks to the kinetic theory of scalar perturbations properly described in Chapter 6 in [31].}

![Fig. 1.2 Refined values of the Universe’s ingredients. BM that makes up stars and galaxies contributes just 4.9% of the Universe’s components. DM occupies 26.8%. [6]](image)

### 1.2 Dark Matter

As shown in Fig. 1.2, DM is the second most dominant component in the Universe. It does not interact significantly either with radiation either with the rest of ordinary matter so that it needs to be detected through its indirect gravitational effects over BM and radiation.

Historically, DM was first presented in the 1930s when the astronomer F. Zwicky used the virial theorem to estimate the total mass of the Coma cluster, a very dense cluster of galaxies. He measured the radial velocities of the galaxies belonging to this cluster from their Doppler-shifted spectra, calculating the dispersion of their radial velocities around $\sigma_r = 997 \text{ km s}^{-1}$. From the virial theorem, he established that the Coma cluster mass should be around $M \approx 3.3 \times 10^{15} M_\odot$ (where $M_\odot$ are solar masses) in order to provide enough gravitational attraction to hold the cluster together. Nevertheless, Zwicky compared this value with the one obtained from the mass-luminosity relation, finding that it was smaller than the one obtained from the virial theorem. He concluded that most of the mass of Coma is not luminous matter but dark, as the amount of visible matter observed in Coma is insufficient to explain how the galaxies are bound gravitationally to each other. In general, the virial theorem is applied in the study of galaxy clusters demonstrating the existence of dark matter that made the cluster to be gravitationally together.\footnote{Zwicky compared this value with the one obtained from the mass-luminosity relation, finding that it was smaller than the one obtained from the virial theorem. He concluded that most of the mass of Coma is not luminous matter but dark, as the amount of visible matter observed in Coma is insufficient to explain how the galaxies are bound gravitationally to each other. In general, the virial theorem is applied in the study of galaxy clusters demonstrating the existence of dark matter that made the cluster to be gravitationally together. [11]}

The dark matter problem identified by Zwicky remained on air until the 1970s when rotation velocity curves of our galaxy and from other spiral galaxies were measured by Vera Rubin and her colleagues. From Newton’s Second Law, spiral galaxies are modelled to have a central bulge and a flat circular disk. Astronomers measured experimentally the velocity of stars rotating around the central bulge by Doppler effect. They obtained flat curves, meaning that stars further from the centre of the galaxy are moving faster than expected. In order to explain this phenomena, astronomers proposed the existence of a massive, spherical and invisible halo of non-luminous matter surrounding spiral galaxies.
Rotation curves are in fact one of the main argument that individual galaxies are surrounded by DM. [19] Thanks to these experimental evidences, DM was postulated to exist both in galaxy clusters and in galaxies themselves.

It is widely believe that DM is placed in the form of halos enveloping the galactic disc and extends well beyond the edge of the visible galaxy. The location is usually inferred by observing the motions of stars and gas clouds in the disk as they orbit the center of the galaxy, obtaining the rotation curves. In first instance, it was thought that massive compact halo objects, known as MACHOs, may be responsible of galaxies halos composition. MACHOs can be in the form of brown or black dwarfs or even stellar-mass black holes. Nevertheless, statistical analysis of gravitational effect of those objects suggest that only around a 19% of the mass of our galaxy halo may be explained by MACHOs [25]. In addition, the distribution of DM in galaxy clusters is inferred by analysing the gravitational lensing effect. Researchers used the observed, subtle distortion of the galaxies shapes to reconstruct the dark matter distribution in the cluster (see Fig. 1.3), where the DM is though to be, responsible of assembling galaxies into groups or cluster (see more information in sections below).

Fig. 1.3 Reconstruction of the supercluster Abell 901/902, composed of hundreds of galaxies. The image in the centre shows the entire supercluster. The magenta-tinted clumps represent a map of the dark matter in the cluster. [2]

1.2.1 Dark Matter Candidates

From cosmological and astrophysical observations it is believed that DM is electrically neutral, and made of particles with velocities below the speed of light, in the non-relativistic limit, which is called Cold Dark Matter (CDM). Moreover, detailed studies of the dynamics of galaxy clusters point out that the particles that compose DM must have velocities below the speed of light, in the non-relativistic limit [25], [20].
There are many theories that contemplate particle candidates suitable be CDM components. As DM hardly interacts with baryonic matter, it was originally believed that those particles could be leptons. Yet, three generations of neutrinos have masses at most of the order of 1 MeV, which is not enough to explain DM mass. It was conceived the possibility that particle candidate may be a neutrino belonging to a unknown fourth generation with a larger mass. Developing this idea, there are strong defenders that say *sterile neutrinos* could compose DM. These particles are neutrinos that do not interact electroweakly and may mix with active neutrinos (meaning the lepton neutrinos considered in the Standard Model of Particles). Sterile neutrinos have been proposed in a number of contexts. For instance, they can be a mass-generating mechanism for the active neutrinos, can be the right-handed counterparts to the active species, or just explain some neutrino-experiment anomalies. Because sterile neutrinos mix with active neutrinos, they have a small decay probability to an active neutrino and a photon, and therefore, being detected [26].

Moreover, there are theoretical physicists that dedicate their effort to study *Weakly Interacting Massive Particles* or WIMPs, the perfect candidates to make up DM. These type of particles are deducted from supersymmetry theories expected to be included in the Standard Model in the future. In these theories, a possible candidate suitable to compose DM could be *sneutrinos* (with spin equals 0) or *neutralinos* (with spin equals 1/2). These particles will have masses up to 1 TeV. Furthermore, there are physicists who consider that the Higgs Boson is not a single state but a doublet whose counter part, the *inert Higgs-doublet*, could fulfil the conditions to be a DM candidate. Consequently, the increment of luminosity in the new set of experiments at the LHC greatly encourages the hope of finding these particles and therefore provide an explanation to DM composition.\(^4\) However, we are assuming that DM do not interact with itself but in fact it could, and there are several *Strong Interactive Massive Particles* suitable to compose DM too.

On the other hand, another hypothetical cold dark matter candidate is the *axion*. Axions are light neutral particles with spin equals 0 that were proposed in the ’80 to explain why non-perturbative effect of the strong interactions do not violate CP invariance. In this theory, axions are low interacting particles, although they can convert into photons in the presence of an strong magnetic field, which provides a hope for its future detection. Again, in relation with supersymmetry theories, axions will be matched with a supersymmetric particle named *axino* with spin equals 1/5, which also fulfils the conditions in order to be the lightest particle in these models.

In general, there is a long list of DM candidates coming from extradimensional DM theories, which propose the existence of a fourth spatial dimension not see by us populated by other particle candidates, such as the *Kaliza-Klein* theory. Still, there is also the possibility that DM is composed by more than one type of particles, because we assume one just for simplicity.

\(^4\)Supersymmetry theories also predicts light particles with mass around 1 TeV such as the *gravitino*, although it is not considered to be a dark cold matter candidate because of other reasons. For further information about this topic the reader may focus on chapter 3.4 in [31].
1.2 Dark Matter

1.2.2 Alternative theories of Gravitation

The most fundamental evidence of the existence of DM is the fact that we need to provide with an extra mass distribution in order to explain gravitational phenomena that are in contrast with the mass estimation due to light emission. For instance, when we use rotation curves to estimate the mass, we are really measuring the acceleration of gas moving around the galaxy in circles. Galaxies are very large, so this acceleration is very small, even much than the accelerations that Newton was considering when he established his laws. Therefore, there is always the possibility that Newton’s laws do not apply on galactic scales. The scientist that stated this claim is Mordehai Milgrom. He carried out a research work affirming that at low accelerations Newton’s laws of gravity break down. Milgrom proposed that "when the acceleration is very low, it becomes proportional to the square root of the central mass and inversely proportional to the distance" (not the square of the distance). His alternative theory receives the name of Modified Nonrelativistic Dynamics (MOND), and he has applied it to explain easily flat rotation curves without the need of introducing DM at all [19].

![Image of Bullet Cluster](image-url)

Fig. 1.4 Image of Bullet Cluster taken from Magellan and the Hubble Space Telescope shows galaxies in orange and white in the background. Pink clouds marks the surface density of the hot gas (obtained from Chandra X-ray). Blue clouds points out the surface density of DM. [4]

However, there is a celestial phenomena that provide evidence against MOND theory: the Bullet Cluster [12]. It is a pair of galactic clusters seen shortly after their collision. The collisional intergalactic gas decoupled from the collisionless component. The gas is offset from the two gasless galaxy clusters toward the center of mass. Most mass of the observable matter is in the gas. The analysis of gravitational lensing revealed that the highest concentration of the gravitating matter is located among the galaxies, exactly where DM should be placed. Hence, the Bullet Cluster is traditionally claimed to be a direct
evidence of exotic DM. Despite, there is still other types of modified gravity theories that does not required the existence of DM in Bullet cluster to explain the observations, leaving still some room for these alternative theories [9].

1.3 Gravitational Lensing

In 1979, a group of astrophysicists\(^5\) noticed a pair of quasars at the practically same redshift, separated around 6 arcsec. Moreover, they presented similar spectrum, suggesting that in fact they were the same object but we were able to see two images due to the deflection of light by an strong gravitational field produced by a massive body in the line between the quasar and the Earth. Indeed, this proposal was verified when a galaxy was found in the line of sight to the quasar. Since this moment, several lensing massive object were found. This effect is known as gravitational lensing, and it provides an extremely good mechanism to search for dark objects or even to study the structure of galaxy clusters, adding another example of indirect evidences of dark matter.

Gravitational lensing is produced when light, which tends to follow the straightest possible trajectory (in a curve space, a geodesic), varies its path as it travels through a curve around a massive object. This phenomenon is completely analogous to the normal refraction of light due to a glass lens that happens when light crosses the lens surface passing from a medium with refraction index \(n\), to another where the refraction index is different \(n'\). Exactly as optical lenses, gravitational lenses have the property of magnifying the brightness of an object.

1.3.1 Lensing Equations

In general, it is a difficult problem to determine light rays trajectories in a curved space-time, as it occurs in our universe. It is been stated that the the mass responsible of gravitational lensing can be stars or galaxies, or even galaxy clusters. From now on, we shall show how to deal with the most general case without taking into considerations any particular space geometry, just from an optical point of view.\(^6\) [14].

A wavefront is a surface composed by wave points in the same phase. Spreading from a source, the wavefront is initially spherical until it suffers a delay when passing through a gravitational field. Regarding the position of the observer at the moment they are crosses by the wavefront, they may see the image magnified or demagnified and even they could see double images. For calculation purposes, we use the concept of arrival time, which is relative to wavefront, and the astrophysical approximation of small angles and thin lenses. The arrival time delay is the difference between the light travel time that follows the deflected ray, which arrives at angle \(\theta_I\) and the unlensed one. It has two contributions:

---

\(^5\)D. Walsh, R.F. Carswell and R.J. Weymann observed this quasar named Q0957+561 lensed by a cD galaxy. The picture of the discovery can be found at [14].

\(^6\)Another way of deducting gravitational lensing equations is starting by the definition of a scaled potential \(\Psi(\theta)\). Further information can be found at [31]
1.3 Gravitational Lensing

Fig. 1.5 Illustration of a gravitational lensing system based on geometrical optics. A ray from the source $S$ placed on the Source Plane at redshift $z_S$ is deflected an angle $\theta$ as a consequence of a lens at Lens Plane $z_L$. $\theta_s$ is the unlensed sky position of the source and $\theta_I$ is the apparent position. $\xi$ is the impact parameter. The angular diameter distances $D_L$, $D_S$ and $D_{LS}$ are the distance from the observer $O$ to the lens, from observer to source and from the lens to source respectively. The angular diameter distance $D$ is defined as the ratio of the physical transverse size of an object to its angular size. The redshift $z$ is defined as the change to longer wavelengths of the spectral lines emitted by a celestial object as a consequence of their cosmological movements.
Theoretical Background

a "geometrical" part and a "gravitational" part. From here to go, we will use natural units such as $G = c = 1$. According to geometrical optics, a light ray is normal to the wavefront, so we can calculate the arrival time delay as the difference of the light travel time between the continuous and the dotted lines (see Fig. 2.5) [28]. This difference can be written as,

$$\Delta t_{geo} = \frac{\xi \alpha}{2}.$$  (1.1)

Regarding Fig 2.5, the following relations remain as true if $\theta_S, \theta_I$ and $\alpha \ll 1$,

$$\beta = \theta_I - \theta_S,$$  (1.2)
$$\theta_I D_S = \theta_S D_S + \alpha D_{LS}.$$  (1.3)

Geometrically, we can obtain the relationship between $\xi$ and the angles $\theta_S$ and $\theta_I$ such as,

$$a = D_L \cos \theta_I; c = 270 - \theta_I; \delta = 90 + \theta_S.$$  (1.4)

Applying the sine theorem,

$$\frac{\xi}{\sin(\theta_I - \theta_S)} = \frac{a}{\sin \delta} \rightarrow \xi = \frac{D_L \sin(\theta_I - \theta_S)}{\cos \theta_I \sin(90 + \theta_S)} = \frac{D_L \sin(\theta_I - \theta_S)}{\cos \theta_I \cos \theta_S} = \frac{D_L \cos \theta_S \sin \theta_I - \cos \theta_I \sin \theta_S}{\cos \theta_S \cos \theta_I} = D_L (\tan \theta_I - \tan \theta_S) \approx D_L (\theta_I - \theta_S),$$  (1.5)

Substituting this result and the one shown in Eq. (1.3) in (Eq. 1.1), together with the redshift factor $(1 + z_L)$ we obtain,

$$\Delta t_{geo} = \frac{1}{2}(1 + z_L)(\theta_I - \theta_S)^2 \frac{D_L D_S}{D_{LS}}.$$  (1.6)

On the other hand, the gravitational delay is the Shapiro time delay in a gravitational field deduced from general relativity which has the form$^7$,

$$\Delta t_{grav}(\theta_I) = (1 + z_L)8\pi \nabla^{-2} \Sigma(\theta_I),$$  (1.7)

where $\nabla^{-2}$ is the inverse of a two-dimensional Laplacian with respect $\theta_I$ and $\Sigma(\theta_I)$ is the surface mass density of the lens. Therefore, the total time delay is,

$$\Delta t = \frac{1}{2}(1 + z_L)(\theta_I - \theta_S)^2 \frac{D_L D_S}{D_{LS}} - (1 + z_L)8\pi \nabla^{-2} \Sigma(\theta_I),$$  (1.8)

---

$^7$There are several ways of deducting Eq. 1.7. Nevertheless, they are out of scope of this project. For further information the reader may look at [10].
From Fermat’s principle, it is stated that light takes the path that make the arrival time minimum. Therefore,

$$\nabla (\Delta t) = 0. \tag{1.9}$$

However, working with Eq. (1.8) may be tedious in order to calculate $$\nabla (\Delta t)$$. For this reason, we introduce some scales. First at all, we define a new scaled-time $$T_0$$ (of the order of the light travel time or Hubble time in cosmological situations) such as:

$$T_0 = (1 + z_L) \frac{D_L D_S}{D_{LS}}, \tag{1.10}$$

so that $$\Delta t = \tau \times T_0$$, and we also define the critical density such as:

$$\Sigma_c = \frac{1}{4\pi} \frac{D_L D_S}{D_{LS}}, \tag{1.11}$$

Consequently, it is possible to express Eq. (1.8) as a function of these scales such as:

$$\tau (\theta_I, \theta_S) = \frac{1}{2} (\theta_I - \theta_S)^2 - 2v^{-2} \kappa, \tag{1.12}$$

where $$\kappa = \Sigma / \Sigma_c$$ defines a new magnitude known as convergence or scaled surface density. It fulfils Poisson Equations as,

$$\nabla^2 \Psi = 2 \frac{\Sigma}{\Sigma_c}, \tag{1.13}$$

where $$\Psi$$ is the gravitational potential. Applying Fermat’s Principle established at Eq. (1.9) to Eq. (1.12) we obtain

$$\nabla^2 \tau (\theta_I) = 0 \rightarrow \theta_S = \theta_I - \nabla \Psi (\theta_I), \tag{1.14}$$

which is known as the lensing equation. This means that the observer would see an image wherever the $$\tau (\theta_S)$$ has a minimum, maximum or saddle point. We can study the curvature of $$\tau (\theta_S)$$ by

$$\nabla^2 \tau (\theta_I) = \nabla \tau (\theta_I), \tag{1.15}$$

which is analogous to

$$\nabla \theta_S = \nabla \tau (\theta_I). \tag{1.16}$$

where $$\nabla$$ denotes the identity matrix. Establishing a relation between Eq. (1.15) and Eq. (1.16) we obtain that $$\nabla \theta_S = \nabla^2 \tau (\theta_I)$$, and we name this quantity the inverse of the magnification matrix $$M^{-1}$$ due to the fact that $$\nabla \theta_S$$ expresses how much source-plane displacement is needed in order to produce a given small image. $$M^{-1}$$ is a two dimension tensor that depends only on $$\theta_I$$. $$M^{-1}$$ must be a symmetric tensor such as,

$$M^{-1} = \nabla^2 \tau (\theta_I) = \begin{pmatrix} 1 - \frac{\partial^2 \Psi}{\partial \theta_I x^2} & -\frac{\partial^2 \Psi}{\partial \theta_I x \partial \theta_I y} \\ -\frac{\partial^2 \Psi}{\partial \theta_I x \partial \theta_I y} & 1 - \frac{\partial^2 \Psi}{\partial \theta_I y^2} \end{pmatrix}, \tag{1.17}$$
Theoretical Background

Fig. 1.6 Schema illustrating the wavefront coming from a source and the different regimes of gravitational lensing regarding the position of the observer. Lensed quasars, Einstein cross and rings fall in the strong lensing limit. The rest of regimes are important in lensing by galaxy clusters. [14]

By comparing Eq. (1.12) and Eq. (1.17) we have:

\[
M^{-1} = (1 - \kappa) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \gamma \begin{pmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{pmatrix}, \tag{1.18}
\]

where \( \kappa \), defined as the convergence and related with the trace of the matrix, has the interpretation of an isotropic magnification and the traceless part of the matrix is defined as the shear \( \gamma \), which changes the shape of a image but not its size.

1.3.2 Weak Lensing

An interesting aspect about gravitational lensing is to consider the transition between strong and weak effect. While strong lensing can be seen at glance as it forms numerous images and arcs, weak lensing is so subtle that is difficult to establish if a galaxy is lensed or not as it is only produce shear. In order to find this turning point, we begin with inverse magnification matrix \( M^{-1} \), which describes the change in source coordinates for an infinitesimal change in image coordinates and is defined in Eq. (1.18). In general, sources are approximated by ellipses described by the the scalar ellipticity \( \varepsilon \) is defined as,

\[
\varepsilon = 1 - \frac{B}{A}, \tag{1.19}
\]

where \( A \) and \( B \) are the semimajor and semiminor axes respectively. We can also define a vector ellipticity \( e_i \) such as,

\[
e_i = (\varepsilon \cos 2\theta, \varepsilon \sin 2\theta) = (e_1, e_2), \tag{1.20}
\]

where \( \theta \) is the angle between the mayoral axis and the horizontal axis, \( e_1 \) is also called the tangential ellipticity \( \varepsilon_t \) and \( e_2 \), the cross ellipticity \( \varepsilon \times \). The vector ellipticity encodes the position angle \( \theta \) and the
1.3 Gravitational Lensing

scalar ellipticity into two quantities that are suitable to be compared between them. The dependence of $2\theta$ indicates the invariance under rotation by 180 degrees. The distribution of $e_i$ is approximately Gaussian with mean value equals to zero (see Fig. 1.7).

Therefore, it is usually assumed that in the absence of lensing, celestial objects are randomly oriented. The effect of the magnification matrix on $e_i$ can be estimated if this matrix is constant over a source (which is mostly true for the vast majority of typical source whose sizes are around few arcseconds). The result is

$$
\epsilon^I = \epsilon^S + \gamma \frac{1}{1 - \kappa},
$$

(1.21)

where the super-indexes point out Image (I) and Source (S). We are not able to compute either $\epsilon^I$ or $\epsilon^S$ for just one source, but we can obtain mean values averaging over several sources. Since, as mentioned above, it is assumed that $\langle \epsilon^S \rangle = 0$, then,

$$
\langle \epsilon^I \rangle = \left\langle \frac{\gamma}{1 - \kappa} \right\rangle,
$$

(1.22)

where the expression on the right side is often named reduced shear $g$. Finally, we are in the weak lensing limit when the convergence $\kappa \ll 1$ so that $\langle \epsilon^I \rangle = \langle \gamma \rangle$. On the other hand, if $\kappa \leq 1$, we are in the strong lensing limit. In Fig. 1.8 we can observe a representation of how circular compact objects are deformed under the weak lensing effect as a consequence of a central mass as a gravitational lens.

Hence, WL induces convergence $\kappa$ and mostly shear $\gamma$ in compact objects. Both features are related with the gravitational potential $\Psi$ [13], according to Eq. (1.17). The convergence $\kappa$ can be derived directly from Eq. (1.13) such as:

$$
\kappa = \frac{1}{2} \nabla^2 \Psi = \frac{1}{2} \left[ \frac{\partial^2 \Psi}{\partial^2 \theta_x} + \frac{\partial^2 \Psi}{\partial^2 \theta_y} \right] = \frac{1}{2} \left[ \Psi_{11} + \Psi_{22} \right].
$$

(1.23)
Theoretical Background

In addition, the shear is taken from the second order derivative of the potential as:

\[
\gamma = \frac{1}{2} \left[ \frac{\partial^2 \Psi}{\partial^2 \theta_x} - \frac{\partial^2 \Psi}{\partial^2 \theta_y} \right] + \frac{1}{2} \left[ \frac{\partial^2 \Psi}{\partial \theta_x \partial \theta_y} + \frac{\partial^2 \Psi}{\partial \theta_y \partial \theta_x} \right] = \frac{1}{2} [\Psi_{11} - \Psi_{22}] + \Psi_{12},
\] (1.24)

Therefore, it is possible to estimate the mass of a cluster from the shear of the galaxies, and taking into account Eq. (1.22), from the ellipticities of the galaxies themselves.

There is a side-effect of WL effect analysis as a consequence of the uncertainty in determining the convergence \( \kappa \) of a lens, known as the mass-sheet degeneracy. We cannot determine by any method the sizes and shapes of the unlensed galaxies as most of the compact objects are affected by this type of distortions. Therefore, we would require an alternative method so that we can determine the absolute value of \( \kappa \) in order to break the degeneracy. However, the proper measurement of \( \kappa \) is only intended for obtaining a value of the mass the cluster, and the mass-sheet degeneracy is not an issue for just detecting galaxy clusters.
We have studied two images that were taken from the Subaru Telescope. Both images correspond to the same sky region but were obtained using different filters. The studied sky region is a subfield of the Lockman Hole area.

2.1 Subaru Telescope

The Subaru Telescope is an optical-infrared telescope placed at the top of Manua Kea volcano in Hawaii (USA), at 4200 meters. Operated by the National Astronomical Observatory of Japan (NAOJ), the telescope has a primary mirror whose diameter is 8.2 meters which collects the light and converges it into four focus, and an advanced system of adaptive optics to get over atmospheric turbulence installed in the infrared camera. It is also provided with a tracking mechanism using a magnetic driving system and a system that controls the quality and performance of the mirror among some other features. In order to reduce any air turbulence inside the enclosure, Subaru adopted a special cylindrical enclosure design [7].

![Fig. 2.1 Picture of Northern Stars over the Subaru Telescope at night [3].](image)

Subaru has a set of astronomical instruments available. These instruments perform several functions with the light it collects. Among the tools, the cameras are the ones destined to capture images of the objects that show their structure and brightness. In our study, images were taken by the Subaru Prime Focus Camera (Suprime-Cam). This camera is an 80 megapixel, optical camera mounted on Subaru’s prime focus. It has a mosaic of ten 2048 x 4096 CCDs and has the capability not only of efficiently imaging a wide field of view but also of capturing images of very faint objects with high levels of detail and contrast. It is especially indicated for surveying large areas of the distant Universe, detecting
small bodies on the outskirts of the Solar System, and map the distribution of DM in the Universe. It has also a set of astronomical filters disposed for deep captures [29].

2.2 Characteristics of the Fields

The images represent a subfield of the area named Lockman Hole. Lockman Hole is a region in the sky placed at the constellation of Ursa Major, centered at around RA 10h45m and DEC 57°. This area is almost free of objects from the Milky Way, giving astronomers the possibility of study far distant galaxies. Moreover, minimal amounts of neutral hydrogen gas have been observed so that extreme ultraviolet and soft x-ray radiation from extragalactic objects are well detected due to the fact that they are not absorbed by the neutral hydrogen gas. Some important objects have been found in this place such as supermassive black holes and active galactic nuclei [25].

We have studied the subfield named lhm1n1n, whose coordinates are from RA 10h55m39s DEC 57°32’58” to RA 10h51m01s DEC 58°03’14”. Subaru took images of this subfield using the camera mentioned above. Therefore, we have images from this field in the optical with the use of a R filter ($\lambda_R = 658$ nm), and in the near infrared with the help of a Z filter ($\lambda_Z = 900$ nm). The complete area has a dimension of $11024 \times 8990$ pixels, where the pixel is $10^{-6}$ radians equivalent to 0.2 arcsec [29].

The analysed images for each filter were created by combining the multiple exposures. This procedure is called stacking.

2.3 Deep Lens Survey

The Deep Lens Survey (DLS) is a 20 arcsec$^2$ ultra-deep multi-band sky survey. The images were taken over about 100 nights on NOAO’s Blanco and Mayall 4-meter telescopes. The main science goal of this survey is mapping DM distribution through the weak gravitational lensing. In addition, they use the multi-wavelength photometry to derive photometric redshifts, and therefore approximate distances, to the galaxies. Because distance corresponds to look back in time, combining these two techniques reveals how massive structures over cosmic time grew. The DLS group has almost finished processing the data, providing several results and publications comparing detected galaxy cluster with other x-ray, optical and spectroscopic observations. Therefore, DLS represents the state-of-the-art in weak lensing techniques [13].

The Deep Lens Survey group developed a data processing pipeline in order to measure WL from the distortion introduced in distant galaxies. The Observational Cosmology and Weak Gravitational Lensing laboratory at Brown University took part in this project [32]. The DLS software was optimized to measure WL effect in fields observed in the R-band, taking into account only images taken on nights with excellent atmospheric conditions. The processing pipeline is able to produce stacked images, calculate convergence maps and make use of PSF anisotropies to reduce the effect of optical and atmospheric distortions within the field (see Chapter 4 for more information about this topic).
Fig. 2.2 Lockman Hole subfield $lhn1n1$ of study in the R-band (upper) and Z-band (bottom). Both show galaxies and stars, and even saturated compact objects.
The study of the sample field \textit{ln1n1} starts by obtaining catalogues that contain lists of detected compact objects. Some of these are real objects (galaxies and stars) whereas others are spurious (artefacts or noise from the background). The workflow of the analysis is explained in the diagram below.

![Diagram of the Estimation and Validation workflow](image)

The first step towards analysing clusters of galaxies through WL is obtaining a catalogue from the image containing a list of the detected compact objects together with several of their characteristics. The astronomical problem of detecting compact objects from a given image has been arduous to work out for years. There are several possibilities to extract the compact objects and most of them require the use of detection programs. Here, \textsc{Source Extractor} has been used [22]. This software is indicated for its application in automated detection and photometry measurements in astronomical images. Among its main pros, its speed and its capacity of handling large files are the most remarkable ones. On the other hand, \textsc{Source Extractor} lacks in proper accuracy regarding the measurement of photometric characteristics.

Basically, \textsc{Source Extractor} works in three separate steps. First, it determines the background locally around the candidate object. This estimation is done by subtracting the mean value around the source computed on a compact region. Second, it applies a filter to smooth the image. In our case, we have selected a Mexican Hat filter, which produces a bandpass-filtering removing both, large and small scale fluctuations. The major advantage of applying a filter before detection is increasing the signal-to-noise of the detection. Third, to consider an intensity peak a compact object, \textsc{Source Extractor} fulfils some conditions at the same time: all the pixels must be above the background threshold, must be shared either corners or sides in common, and they form a set of pixels, at least,
equal or larger than a set formed by the minimum number of pixels determined by a given input parameter.

Once that the object is detected, SOURCE EXTRACTOR approximates its shape by an ellipse. The position of the center of these objects as well as the basic shape parameters $A$ and $B$, the semi-major and semi-minor axis lengths respectively, are computed from the first and second order moments of the elliptical object’s profile. Another characteristic computed from the first and second order moments are the integrated per area flux $F_{iso}$ and magnitude $m_{iso}$ (denominated by the term isophotal), their positions along the axis $x$ and $y$ ($x, y$), the ellipticity $\varepsilon$ and the position angle between the axis $x$ and the mayor axis of the ellipse $\theta$.

SOURCE EXTRACTOR requires some input data values, for instance the detection threshold, the number of pixels that belongs to a compact object, and the detection type. After testing the results for several input parameters required by SOURCE EXTRACTOR, we have maintained mostly those provided by SOURCE EXTRACTOR by default, incrementing the number of deblending sub-thresholds and decreasing the minimum contrast parameter for deblending. The catalogues resulting from the extraction are shown in Fig. 3.2. Hereafter, the catalogue corresponding from the optical image is named $R$-catalogue; analogously, the catalogue from the near infrared one is called $Z$-catalogue.

![Fig. 3.2 Detection of compact objects using SOURCE EXTRACTOR in a region of the lhm1n1 field. (Left) for the optical R-band image. (Right) for the near infrared Z-band image.](image)

### 3.2 Validation of the Catalogues

As mentioned above, SOURCE EXTRACTOR creates catalogues that may include a high number of spurious detections corresponding to compact objects that are not real. Obviously, preserving those spurious objects in the catalogues is a source of systematic errors due to the fact that those spurious detections may induce fictitious galaxy clusters when studying mass distributions. Thus, it is necessary to carry out different tests in order to side apart those spurious detections from the R-catalogue and Z-catalogue. For this aim, we have developed two independent validation assessments: one based on cross-matching methods and another in simulation techniques.
3.2.1 Cross-Matching of Catalogues

This checking process consists on comparing both the R-catalogue and the Z-catalogue. The premise can be summarized such as "spurious objects that do not really exist in the field and consequently should not be detected, will not appear in both catalogues after extraction". This is specially true for detections coming from the background and near saturated stars.

The comparison is performed by a method denominated cross-matching: we look for the presence of objects from the R-catalogue in the Z-catalogue and vice-versa. The search is performed using the KD-Tree algorithm [1] with a confidence radius \( r = k \times r_p \), being \( k \) an integer and \( r_p \) the size of the pixel in radians. Firstly, the objects of the Z-catalogue are referred as keys of the search. We count the number of object that are not found \( n_{lost} \) in the R-catalogue and we write a new catalogue, called \( R-Z\text{-catalogue} \) containing the objects found in the R-catalogue according to Z-catalogue keys. Secondly, we repeat these steps looking for the objects of the Z-catalogue in the R-catalogue. We write a new catalogue, named \( Z-R\text{-catalogue} \), that contains the objects found in Z-catalogue according to R-catalogue keys.

The cross-matching algorithm was successfully applied to our field. The features of the R-Z and Z-R catalogues are recorded in Table 3.1. The length of the R-catalogue is 118135, meaning that it contains information of that number of objects. On the other hand, the length of the Z-catalogue is 99218. Therefore, using the objects from the Z-catalogue as keys (which is the process of the cross-matching R-Z), we seek for 99218 objects in the catalogue that hosts 118135 objects. From these 99218 objects, only 79815 were found, losing in the process 19404 objects. In the same way, from the process of cross-matching Z-R, we seek for 118135 objects in a catalogue containing only 99218 objects. The result is a catalogue containing 79815 objects, after missing 38331 objects in the cross-matching process.

<table>
<thead>
<tr>
<th>cross-matching</th>
<th>( n_{lost} )</th>
<th>( l_{oc} )</th>
<th>( l_{cmc} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-Z</td>
<td>19404</td>
<td>118135</td>
<td>79815</td>
</tr>
<tr>
<td>Z-R</td>
<td>38331</td>
<td>99218</td>
<td>79815</td>
</tr>
</tbody>
</table>

Table 3.1 Summary of the characteristics of the obtained catalogues after the cross-matching. \( n_{lost} \) refers to the number of objects from the key catalogue not found in the other catalogue. \( l_{oc} \) is the length of the original catalogue. \( l_{cmc} \) is the length of the catalogue got after the cross-matching process. The confidence radius for the matching is \( r = k \times r_p = 3 \times 10^{-6} \text{ rad} \). The R-Z catalogue contains the magnitudes associated to the R-band image. The Z-R catalogue contains the magnitudes associated to the Z-band image.

We have also studied the relation between the radius \( r \) for which the cross-matching process is performed. The results are recorded in Table 3.2. It is observed that for confidence radii \( k > 3 \), the cross-matching process does not find the same amount of objects for both R-Z and Z-R catalogues. Moreover, it is checked that for values \( k > 100 \) the cross-matching process for R-Z and Z-R recovers almost the totality of the 99218 objects from the most restrictive catalogue. Therefore, the selection of \( k = 3 \) for the confidence radius \( r \) implies that we are using radius \( r \) that offers the minimum restriction during the search but assures that both R-Z and Z-R catalogues have the same dimension.
3.2 Validation of the Catalogues

<table>
<thead>
<tr>
<th>k</th>
<th>cross-matching</th>
<th>n_{lost}</th>
<th>l_{oc}</th>
<th>l_{cmc}</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>R-Z</td>
<td>29187</td>
<td>118135</td>
<td>70032</td>
</tr>
<tr>
<td>1</td>
<td>Z-R</td>
<td>48104</td>
<td>99218</td>
<td>70032</td>
</tr>
<tr>
<td>2</td>
<td>R-Z</td>
<td>20965</td>
<td>118135</td>
<td>78254</td>
</tr>
<tr>
<td>2</td>
<td>Z-R</td>
<td>39882</td>
<td>99218</td>
<td>78254</td>
</tr>
<tr>
<td>3</td>
<td>R-Z</td>
<td>19404</td>
<td>118135</td>
<td>79815</td>
</tr>
<tr>
<td>3</td>
<td>Z-R</td>
<td>38311</td>
<td>99218</td>
<td>79815</td>
</tr>
<tr>
<td>4</td>
<td>R-Z</td>
<td>18607</td>
<td>118135</td>
<td>80612</td>
</tr>
<tr>
<td>4</td>
<td>Z-R</td>
<td>37509</td>
<td>99218</td>
<td>80627</td>
</tr>
<tr>
<td>5</td>
<td>R-Z</td>
<td>18072</td>
<td>118135</td>
<td>81147</td>
</tr>
<tr>
<td>5</td>
<td>Z-R</td>
<td>36945</td>
<td>99218</td>
<td>81191</td>
</tr>
<tr>
<td>10</td>
<td>R-Z</td>
<td>14332</td>
<td>118135</td>
<td>84887</td>
</tr>
<tr>
<td>10</td>
<td>Z-R</td>
<td>31776</td>
<td>99218</td>
<td>86360</td>
</tr>
<tr>
<td>100</td>
<td>R-Z</td>
<td>4</td>
<td>118135</td>
<td>99214</td>
</tr>
<tr>
<td>100</td>
<td>Z-R</td>
<td>18917</td>
<td>99218</td>
<td>99218</td>
</tr>
</tbody>
</table>

Table 3.2 Summary of the characteristics of the obtained catalogues after the cross-matching process for different confidence radius \( r = k \times r_p = k \times 10^{-6} \) rad. \( n_{lost} \) refers to the number of objects from the key catalogue not found in the other catalogue. \( l_{oc} \) is the length of the original catalogue. \( l_{cmc} \) is the length of the catalogue got after the cross-matching process.

Studying the catalogues R-Z and Z-R obtained from the cross-matching process for \( r = 3 \times 10^{-6} \) rad, it is checked that there is no cross-matched object whose isophotal magnitude \( m_{iso} \) is higher than 25. This statement attempts to demonstrate that objects detected with \( m_{iso} > 25 \) are spurious objects that will be likely not detected at different filters. In order to verify this attestation, another validation test is performed.

### 3.2.2 Add-Objects Simulation

This simulation consists on throwing \( n \) number of mock galaxies in the field \( lhn1n1 \) at known positions \((x,y)\) following a linear distribution in magnitude, and then analysing those galaxies by SOURCE EXTRACTOR with the purpose of establishing a limit in the detection of faintest galaxies through the value of the maximum magnitude \( m_{max} \).

We have calculated the mean values of the detected objects isophotal magnitudes \( \langle m_{iso} \rangle \), ellipticities \( \langle \epsilon \rangle \) and minor axis \( \langle B \rangle \) directly from the data contained in the R-catalogue, giving the mean values of the major and minor axis of the detected objects to mock galaxies. Besides, we have established a relation between the flux \( F \) and the magnitude \( m \) data from the R-catalogue according to the general expression in [11], so that translation between fluxes and magnitudes are possible. Then, we have calculate \( n \) pairs of random numbers for values of \( x \) and \( y \) belonging to the R-band image size. We check whether a compact object was already detected at the position of those random pairs \((x,y)\). If
not, we create elliptical mock galaxies for which their magnitude $m$ are introduced as input data. The distribution of the intensity flux $F$ along the ellipse is defined by an elliptical gaussian function.

Once that the mock galaxies are created in the R-band image, we apply SOURCE EXTRACTOR again to extract a new R’-catalogue including photometric information of those $n$ galaxies. We look for their features information corresponding to the new $n$ mock galaxies in the new catalogue, comparing the mean value of the mock galaxies output isophotal magnitude $m_{iso}^{i}$ provided by SOURCE EXTRACTOR regarding the input value of the isophotal magnitude $m_{iso}^{i}$ that was given to the simulated galaxies. We have also counted the number of mock galaxies not found ($n_{lost}$) by SOURCE EXTRACTOR.

We have repeated this process for mock galaxies whose input maximum magnitude $m_{max}^{i}$ goes from 1 to 30, launching each time a number of mock galaxies equals to $n = 4000$. Fig. 3.3 shows the detection through SOURCE EXTRACTOR of the mock galaxies together with diverse real objects.

![Fig. 3.3 Detail of the field lhn1n1. (Right) for own-made galaxies at $m_{max}^{i} = 15$. (Left) for own-made galaxies at $m_{max}^{i} = 22$. Note that the red circular apertures remarks just a pair of those thrown galaxies that do not belong to the original field.](image)

The aim of this simulation is to study whether a cut in faint magnitudes can be applied in the field lhn1n1, and if this cut is compatible with the previous results obtained from the cross-matching process. A suitable cut in faint magnitudes has been found, according to Fig. 3.4, for $m_{iso}^{i} \approx 25$, due to the fact that the linear dependency observed in the left part of Fig. 3.4 starts to fail because SOURCE EXTRACTOR has issues to attach the proper value to the isophotal flux and therefore to the isophotal magnitude. This effect is known as Eddington bias. The suitable cut is verified in the right part of Fig. 3.4 where it is observed an increment of the number of mock galaxies lost during the detection process.
around $m_{iso}^i \approx 25$. This cut of $m_{iso}^i \approx 25$ is in agreement with the photometric characteristics of objects belonging to the R-Z and Z-R catalogues, where no object was found with a $m_{iso} > 25$.

Fig. 3.4 (Left) graphical representation of the percentage of the number of lost thrown galaxies $n_{lost}$ after the second detection process by SOURCE EXTRACTOR as a function of the isophotal input magnitude $m_{iso}^i$. (Right) output isophotal magnitude $m_{iso}^o$, obtained from the second detection of the thrown galaxies as a function of the input isophotal magnitude $m_{iso}^i$, in semi-logarithmic scale. Note that the values of the input and output magnitudes are the mean values for each input data of the total number of thrown galaxies $n = 4000$.

### 3.3 Further cleaning of spurious objects

In spite of the cross-matching process and the simulation technique, it is required to rid of still remaining spurious objects related with saturated pixels before starting further steps.

We proceed firstly by rejecting pixels on the edge of the image since SOURCE EXTRACTOR fails to perform object detection on this edges. Objects with coordinates $x < 400$ and $y < 300$ as well as $x > 10500$ and $y > 8500$ are discarded (see Fig. 2.2 for the details of these saturated pixels). The length of the R-Z catalogue after rejecting compact objects detected at edges is $l_{edges}^{R-Z} = 69332$, and the length of the Z-R one is $l_{edges}^{R-Z} = 69385$.

In addition, we eliminate objects whose isophotal flux $F_{iso}$ is larger than $F_{iso} > 300000$. This value is associated to saturated stars showed in the R-band and Z-band images. Therefore, this cut is performed in order to avoid selecting spurious objects detected by SOURCE EXTRACTOR close to saturated stars (see Figs. 2.2 and 3.2). The length of the R-Z catalogue after rejecting compact objects detected close to saturated stars is $l_{saturated}^{R-Z} = 69015$, and the length of the Z-R one is $l_{saturated}^{R-Z} = 68908$. 
Once that the R-Z and Z-R catalogues have been estimated and validated, the next process consists on the study of those catalogues in order to establish a classification of the compact objects in stars or galaxies and the estimation of their sizes given. Furthermore, it is necessary to correct the distortion introduced in the shear of the objects by different phenomena such as the atmosphere or the optical instruments, given in terms of the point spread function (PSF). This is a critical point since, as the shear introduced by WL effect is considerable small, PSF anisotropies will complicate the interpretation of the data if left uncorrected. This process is performed by a script (see Appendix A, section A.5.3.) and outlined as:

Fig. 4.1 Diagram of the division and PSF correction workflow. Both R-Z and Z-R catalogues are analysed to classify compact objects in galaxies or stars. Secondly, the shapes of those objects are estimated by alternative software, and compared to the ones provided by SOURCE EXTRACTOR to gain in precision and reliability. Thirdly, the PSF anisotropies are corrected by studying the profile of the stars. Finally, the classification and shape estimation process is repeated with the corrected image.

4.1 Division in galaxies and stars

It is necessary to establish a classification between galaxies and stars of the compact objects for two main reasons. First, obviously, galaxies are the required objects to study DM distribution; second, stars play a key role to correct the data for systematics, in particular, PSF anisotropy correction. To perform this classification, we have followed two different methods. The first method is based in the study of the size of compact objects as a function of isophotal magnitudes. The second one is based on a neural network technique implemented in SOURCE EXTRACTOR, returning a real parameter, denominated class\textsubscript{star}, which is equals to 1 for objects classified as stars, and equals to 0 for galaxies.

To check the robustness of both methods, we have study the number of objects classified from the R-Z and Z-R catalogues produced by each of the two criteria mentioned above studying which is the most restrictive one.
4.1 Division in galaxies and stars

Fig. 4.2 R-Z Catalogue. Distribution of the size of the objects given in terms of their $FWHM$ in arcsec in as a function of their isophotal magnitude $m_{iso}$. See text for a description of the colours.

First, we represent of the size of a given objects, studied by SOURCE EXTRACTOR from the Full Width at Half Maximum $FWHM$ of the elliptical distribution, as a function of its isophotal magnitude of the objects $m_{iso}$ (see Fig. 4.2 and 4.4). From these plots, it is possible to establish an hypothesis to divide celestial objects into stars and galaxies (see e.g. [13]). We consider as stars those objects whose size $FWHM$ remains almost constant and low for bright magnitudes $m_{iso}$. Looking at Fig. 4.2, and equivalently 4.4, stars would be placed at bottom-left part of the graph showed in dark blue stars, which almost practically are covered by the green stars (see below for details). On the other hand, we consider galaxies to those objects above the imaginary straight line (see black solid line in Fig. 4.2) that fits the scattering dots in the left side of the dashed black line. In this manner we compel our galaxies to be objects whose size is bigger than the value associated to stars.

In particular, the parameters that define this classification can be summarized as: for the R-Z catalogue, we have used values for $FWHM$ and $m_{iso}$ for the stars filtering such as $(m_{iso}^{min}, m_{iso}^{max}) = (15.5, 17.5)$ and $(FWHM^{min}, FWHM^{max}) = (3.4, 3.7)$, and for the galaxies filtering $(m_{iso}^{min}, m_{iso}^{max}) = (8.0, 17.5)$ and $FWHM^{min} = 3.7$. For the Z-R catalogue, the values for $FWHM$ and $m_{iso}$ for stars filtering were $(m_{iso}^{min}, m_{iso}^{max}) = (16.0, 17.0)$ and $(FWHM^{min}, FWHM^{max}) = (3.0, 4.0)$, and for the galaxies filtering $(m_{iso}^{min}, m_{iso}^{max}) = (8.0, 17.0)$ and $FWHM^{min} = 4.0$. The length of the resulting stars and galaxies catalogues for the R-Z catalogue are $l_{stars}^{R-Z} = 247$ and $l_{galaxies}^{R-Z} = 61769$. Regarding the Z-R catalogue, the length of the resulting stars and galaxies catalogues are $l_{stars}^{Z-R} = 187$ and $l_{galaxies}^{Z-R} = 61363$. In Fig. 4.2 and 4.4), the objects resulting from this filtering are plotted in dark blue and red, respectively.
We have compared the results obtained in the above classification with those obtained directly from filtering R-Z and Z-R catalogues attending to the \( \text{class}_{\text{star}} \) parameter provided by \textsc{source extractor} neural network technique (hereinafter referred to as \( \text{class}_{\text{before}} \)). The length of the resulting stars and galaxies catalogues for the R-Z catalogue are \( l_{\text{stars}}^{R-Z} = 1194 \) and \( l_{\text{galaxies}}^{R-Z} = 59931 \), respectively. On the other hand, the length of the these catalogues for the Z-R catalogue are \( l_{\text{stars}}^{Z-R} = 1528 \) and \( l_{\text{galaxies}}^{Z-R} = 60898 \). In Figs. 4.2 and 4.4, the objects resulting from this filtering are plotted in yellow and purple, respectively. As mentioned before, \textsc{source extractor} gives the parameter \( \text{class}_{\text{star}} \) a value 1 for those objects classified as stars, and a value 0 to galaxies. However, \( \text{class}_{\text{star}} \) is not an integer number but a real one, so that to carry out this classification according to this parameter with have applied a cut to consider stars as objects with have \( 1 > \text{class}_{\text{star}} > 0.9 \) and galaxies \( 0 < \text{class}_{\text{star}} < 0.1 \), allowing the parameter to have an uncertainty of 10%. In Fig. 4.3, we can observe how the population of galaxies and stars found in the R-Z catalogue for the cut values mentioned previously are mostly focused on those intervals (similar results are found for the Z-R catalogue).

![Normalized histogram](image)

Fig. 4.3 Normalized histogram based on the parameter \( \text{class}_{\text{star}} \) from the stars catalogue (blue) and the galaxies (green) after the \( FWHM \) vs. \( m_{\text{iso}} \) and neural network filtering.

From the use of this criteria, we find that the number of galaxies obtained is similar to those classified by the \( FWHM \) vs. magnitude method, but the discrepancy in the number of stars is significant, finding a larger number of this type of objects when applying \( \text{class}_{\text{before}} \) criteria. This indicates that the \( FWHM \) vs. magnitude method is more restrictive regarding the classification of stars and, therefore, safer for the final purpose.
Finally, we have filtered R-Z and Z-R catalogues according to the $class_{star}$ parameter after the application of the key $FWHM$ vs. $m_{iso}$ (hereinafter referred to as $class_{star}^{after}$). This permutation is the most restrictive one for the galaxies catalogue. The length of the resulting stars and galaxies catalogues for the R-Z catalogue are $l_{stars}^{R-Z} = 246$ and $l_{galaxies}^{R-Z} = 53560$. For the Z-R catalogue, the length of the resulting stars and galaxies catalogues are $l_{stars}^{Z-R} = 187$ and $l_{galaxies}^{Z-R} = 38248$. In Fig. 4.2 and 4.4, the objects resulting from this filtering are plotted in green and light blue, respectively.

![Fig. 4.4 Z-R Catalogue. Distribution of the size of the objects given in terms of their FWHM in arcsec in as a function of their isophotal magnitude $m_{iso}$. See text for a description of the colours.](image)

At the end, we conclude that using a joint classification (first $FWHM$ vs. magnitude plus the neural network) most of objects containing in the stars catalogue are in fact classified as stars by SOURCE EXTRACTOR, missing one object regarding the R-Z catalogue and none from the Z-R catalogue. In the same way, the galaxies catalogue contain most objects identified by SOURCE EXTRACTOR as galaxies. In spite of the fact that we are missing several objects classified as stars by SOURCE EXTRACTOR by the key $class_{star}^{before}$, the joint of both methods (neural network and $FWHM$ vs. $mag_{iso}$ study) offers an extra robustness. Stars will be used in future steps in order to correct PSF anisotropies. This also applies to the galaxies catalogue. Combining both methods offers the certainty that our galaxy catalogue contains objects that are in fact galaxies, which is crucial for obtaining DM distribution maps in next steps. The filtering of the catalogues into several classes has been performed with FIATFILTER (see Appendix A for further information).
4.1.1 Shapes Estimation

Although SOURCE EXTRACTOR measures the ellipticities of the objects, we have used another software, ELLIPTO [8], to recalculate the ellipticities of compact objects in order to gain in precision. We perform this step because at later stages of it is very important to have an accurate measurement of the shapes, helping to reduce systematic errors. ELLIPTO measures the intensity moments by fitting an elliptically Gaussian to each compact objects using as input parameters SOURCE EXTRACTOR ellipticities as a first guess. Then, it maximizes signal-to-noise when the size and shape of the Gaussian matches those of the object. Finally, it adjusts the Gaussian function iteratively until the fit converges to a single best value. If the fit fails to converge, or if it converges to a value that is far from the original SOURCE EXTRACTOR input, an error code is generated.

Both the galaxies and stars catalogues are filtered erasing those objects that ELLIPTO reports with an error code larger than 2, meaning that the shape calculated by ELLIPTO differs considerably from that one calculated by SOURCE EXTRACTOR. The results of the new shape estimation can be shown in Fig. 4.5 and 4.6 for both stars and galaxies of the R-Z catalogue (similar results are obtained for the Z-R catalogue). It can be observed how ELLIPTO circularizes the shape of both types of compact objects. This change of the ellipticity components towards 0 can be observed through a parameter $d$ that indicates the distance of each pair of the ellipticity components $e_1$ and $e_2$ with respect the centroid of the distribution, and it is defined such as:

$$d(i) = [(e_1(i) - \langle e_1 \rangle)^2 + (e_2(i) - \langle e_2 \rangle)^2]^{1/2} \rightarrow d = \langle d(i) \rangle,$$

where $i$ indicates the elements of the ellipticity vectors components $e_1$ and $e_2$, $\langle e_1 \rangle$ and $\langle e_2 \rangle$ are the mean values of $e_1$ and $e_2$ which identify the centroid of the distribution, and $\langle d(i) \rangle$ is the mean value of all distances with respect the centroid. This parameter decreases for each correction of stars shapes (ELLIPTO and further correction in PSF anisotropies).

![Fig. 4.5 R-Z Catalogue. Graphical representation of the two components of the vector ellipticity $(e_1, e_2)$ for the stars catalogue. (Left) before correcting with ELLIPTO ($d = 0.0177$). (Middle) after correcting with ELLIPTO ($d = 0.0148$). (Right) after correcting PSF anisotropies with DLSCOMBINE ($d = 0.0095$)](image-url)
4.1 Division in galaxies and stars

Fig. 4.6 Graphical representation of the two components of the vector ellipticity \((e_1, e_2)\) for the galaxies catalogue. (Left) before correcting with ELLIPTO. (Right) after correcting with ELLIPTO.

With respect Fig. 4.6, we observe how there are some galaxies that tend to have higher values of \(e_1\), whereas \(e_2\) is almost 0. Regarding Fig. 1.7, these values correspond with galaxies that are almost vertical and are considerably stretched. Probably, these objects are not galaxies but spurious objects still detected close to saturated stars. This problem is eliminated during the second cross-matching process after correction of PSF anisotropies (see Fig. 5.3).

4.1.2 Correction of the PSF

The shear due to WL effect would be covered by other effects as it is quite subtle so that other distortions must be removed. Therefore, once that the shapes have been properly estimated using ELLIPTO, it is time to correct PSF from systematics. The PSF describes the widening and blurring of light from a point-like source as it passes through the atmosphere and the telescope optics before reaching the detector [23]. This function determines the resolution that system is able to achieve. Usually, the PSF anisotropy is on the order of a few percent on one arcminute scales.

As it has been already mentioned, the apparent size and shape of the stars in an image can be used to measure the PSF across the image. The profile of the stars are fitted to a model of the PSF, which is then used to deconvolve the image.

The goal of this correction is eliminate PSF anisotropies in the image, and we use stars as tracers of this anisotropies. In particular, we connect any departure from isotropy on the shape of the stars, since such departures will be produced by systematics, but not, obviously, by the WL effect. In this case, we are not considering a particular source of anisotropies, but rather try to correct all of them effectively. In fact, the smearing observed in PSF anisotropies can be due to the atmosphere, for instance the atmospheric refraction and dispersion, but even due to optical aberrations or guiding errors. The solution to this problem is resolved by DLSCOMBINE, a software contained in the DLS pipeline that corrects PSF anisotropies by devising a position-dependent convolution kernel which circularizes these PSF [17].
DLSCOMBINE studies alignments in the stars ellipticity due to PSF anisotropies in order to apply a global correction to the image. Consequently, DLSCOMBINE requires a guide to apply the kernel to correct PSF in the image. This guide is perform by a fitting of the vector ellipticity components of the stars, \( e_1 \) and \( e_2 \), as a function of their coordinates \((x, y)\). In our case, we have developed a software to provide to DLSCOMBINE a model of a polynomial fit whose base are Legendre polynomials \( L(x, y) \), \(^1\). The fitting is performed by minimizing the sum \( \chi \) of squares of the difference between the experimental value of the ellipticity component \( e_i \) and the model such as:

\[
\chi_i^2 = \sum_x \sum_y \left( \frac{e_i(x, y) - \sum_n \sum_m c_{nm}(x, y) L_n(x) L_m(y)}{\sigma_e(x, y)} \right)^2 ,
\]

where \( L_n(x) \) and \( L_m(y) \) are Legendre polynomials, of order \( n \) and \( m \), respectively, \( \sigma_e(x, y) \) is the relative error associated to the ellipticity, and \( c_{nm}(x, y) \) are the coefficients of the fitting. The grade of the fitting is so that \( n + m \leq 4 \), therefore \( n, m = \{0, 1, 2, 3, 4\} \). Ellipticity components have been fitted according Eq. (4.2) following a non-linear fit through the method of least square fitting, described in Chapter 5 of [30]. This least square fitting is performed by the iterative Levenberg–Marquardt algorithm (see Appendix A for more information).

Furthermore, the fitting is retroactive so that it is performed in two steps. First, we perform the fitting using all stars following Eq. (4.2). We fit both components of the vector ellipticity \( e_i \) to the polynomial model, evaluating the residuals as the difference of the experimental values of the ellipticities and the theoretical values according to the obtained fitting. Second, we reject objects whose ellipticity components provide values for the residuals bigger than \( 3\sigma \), being \( \sigma \) the standard deviation of the obtained residuals, and then we repeat the fitting. Performing this retroactive step, we assure that the objects contained in the stars catalogue are, in fact, stars because they follow a similar distribution regarding their coordinates \((x, y)\).

With the coefficients \( c_{nm} \) of the fitting, DLSCOMBINE constructs a kernel that is used to deconvolve the PSF anisotropies in \( e_1 \) and \( e_2 \). Finally DLSCOMBINE returns a new image containing information of the original image but with the PSF anisotropies already corrected globally. The result of the fitting according to Eq. (4.2) for stars belonging to both R-Z and Z-R catalogue are shown in Fig. 4.7 and 4.8.

In general, we have observed how the uncertainty associated to the fitting coefficient increases as the order of the fitting also increases. When correcting PSF anisotropies from images, it is expected that anisotropies are larger at the edges of the images, having a clear region in the center of the image. This expected behaviour is observed in Fig. 4.7 and 4.8, where we represent by colour-maps the fitting of \((e_1, e_2)\) for R-Z and Z-R catalogues.

---

\(^1\)A simple polynomial fitting instead of using Legendre polynomials can be performed. However, the use of Legendre Polynomials offers more stability during the correction process as they are orthogonal.
4.1 Division in galaxies and stars

Fig. 4.7 Colour representation of the two components of the vector ellipticity \((e_1, e_2)\) for the stars catalogue belonging to the R-Z catalogue as a function of the coordinates \((x, y)\). (Left) fitting performed with the totality of the stars. (Right) after the rejection according to a cut of \(3\sigma\) (see text).

Fig. 4.8 Colour representation of the two components of the vector ellipticity \((e_1, e_2)\) for the stars catalogue belonging to the Z-R catalogue as a function of the coordinates \((x, y)\). (Left) fitting performed with the totality of the stars. (Right) after the rejection according to a cut of \(3\sigma\) (see text).

To check that the PSF anisotropies correction was applied globally to the R-band and Z-band images, we have subtracted the originals images from the corrected ones. The results are images showing an excess of intensity flux, corresponding to those pixels where \textsc{dlscombine} played a more important role to rid of PSF anisotropies. Consequently, images observed in Fig. 4.9 look pretty similar to above-placed colour-maps.
Fig. 4.9 Images corresponding to the colour-maps shown in (at left) Fig. 4.7 after subtraction original R-band picture from the r corrected picture, and (at right) Fig 4.8, after subtraction original Z-band image from the Z-band corrected one. The similarities between theses images and cited fitting colour-maps are relevant.

4.2 After correction

Once that the PSF anisotropies have been corrected in both fields obtaining two new images, we repeat the main steps of our analysing pipeline. SOURCE EXTRACTOR is applied again to each field in order to get two corrected full-catalogues. In order to avoid spurious objects, the process of cross-matching explained in section 3.2.1 is repeated with the new corrected catalogues. The confidence radius is set at $r = 3 \times 10^{-6}$ rad. The length of the resulting corrected catalogue is $l_{cc} = 70201$ objects, losing no objects in the final step comparing $R - Z$ and $Z - R$, and obtaining a final catalogue named (R-Z)'-catalogue.

Precisely because the software used in the DLS pipeline is optimized for obtaining DM distribution maps from R-band observations, the (R-Z)'-catalogue contains the features associated to the R-band image. From this moment and on, we do not keep working with properties associated to the Z-band.

From this corrected and cross-matched catalogue we repeat sections 4.1 and 4.2. The cuts used in section 3.3 are maintained also in this step. For the division of galaxies and stars, the cuts performed to the R-Z catalogue are still applied in the classification process. The resulting final galaxies catalogue (see Fig.4.10 bottom) has a length of $l_{galaxies} = 38324$ objects.

In Fig. 4.10, it is plotted a central detail of the studied field overlapped with two different catalogues: at the upper panel, the original R-catalogue, and at the bottom, the final galaxies catalogue. We can observe how several spurious objects detected by SOURCE EXTRACTOR inside massive compact objects have been cleaned up from the final catalogue. In the same way, obvious objects identified visually as stars are not included in this final galaxy catalogue.
Fig. 4.10 Small central region of the image in the optical of the Lockman Hole area showing the cross-matching catalogue obtained at the end of chapter 4 before filtering (above) and the final galaxies catalogue after PSF correction, cross-matching, and filtering (bottom).
Once we have obtained the final catalogue of galaxies, corrected in shape and without PSF anisotropies, it is necessary to apply a reconstruction method in order to create a map of the mass distribution in the field of observation. Two main groups of reconstruction methods can be defined. First, direct methods derive an estimate of the shear field just using the observed distortion in the galaxies ellipticities. On the other hand, inverse methods transform the field to get a two-dimensional projected surface mass density in the studied area [13]. In this project, a direct method has been used, provided as a way to detect galaxy clusters. This approach does not provide information about the mass of the cluster itself.

![Diagram of the analysis method workflow to obtain a mass-density map]  
Fig. 5.1 Diagram of the analysis method workflow to obtain a mass-density map. First, the final galaxies catalogue is evaluated by a direct method implemented in the software FIATMAP to obtain a convergence map. Secondly, a Monte Carlo method is used to produce a list of randomized convergence maps of the field. The latter are used to estimate the noise of the reconstruction process so that signal-to-noise S/N maps are produced.

### 5.1 Convergence Maps

The convergence $\kappa$ is calculated from the tangential ellipticities $\varepsilon_t$ representing in the final mass-map over-densities. Convergence maps are obtained thanks to a direct method implemented in a program called FIATMAP [17], which is based on the earlier software developed by J. Anthony Tyson and Francisco Valdes invlens [18] and also in the pioneered work of Kaiser and Squires [23]. This code needs as input parameters the dimensions of the original image, an inner radius $r_{in}$, an outer radius $r_{out}$, a block size $b$, and a catalogue of objects and their associated ellipticities as input parameters, giving as output an image with the mass-map.

FIATMAP computes a sum of the tangential ellipticities $\varepsilon_t$ of the galaxies inside a block $b$. The tangential ellipticities $\varepsilon_t$ are weighted by an exponential function which gives more weight to closer galaxies within the block. The convergence $\kappa$ for every object in the block is identified by this sum as

$$\kappa(p) = \frac{1}{N} \sum_{q \in b} \frac{\varepsilon_t}{r^2} W,$$

(5.1)
where \( q \) is the pixel at which the galaxy belonging to the block \( b \) is placed, \( \epsilon \) is the tangential ellipticity of the galaxies whose centroid is a distance \( r = \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2} \) from the point of evaluation \( p \). Eq. (5.1) is a phenomenological expression obtained empirically in [17]. Every block \( b \) will form up a pixel in the final mass-map. The sum is repeated for the number \( N \) of galaxies in the block placed at \( q \) if any source exists. The weight function \( W \) is defined as,

\[
W = e^{-r/r_{\text{out}}} \left( 1 - e^{-r/r_{\text{in}}} \right).
\] (5.2)

In addition, FIATMAP uses a weight file to decrease the contributions of objects near the edges of the original image. The formal description of the complementary software to run FIATMAP is included in Appendix A.

Fig. 5.2 Convergence maps corresponding for sub-catalogues whose isophotal magnitude \( m_{\text{iso}} \) goes from 17 to 20 (left), 20 to 22 (middle), and 22 to 25 (right).

Fig. 5.3 Graphical representation of two components of the vector ellipticity \((e_1, e_2)\) corresponding for sub-catalogues whose isophotal magnitude \( m_{\text{iso}} \) goes from 17 to 20 (left), 20 to 22 (middle), and 22 to 25 (right).

The final galaxy catalogue mentioned at the end of Chapter 4 is split into three different sub-catalogues attending to a division in magnitude. The first sub-catalogue contains objects whose magnitude is smaller than 20, the second sub-catalogue contains object whose magnitude goes from 20 to 22, and the final sub-catalogue records objects whose magnitude goes from 22 until 25. The representation of the ellipticity vector components \( e_i \) corresponding to the galaxies belonging to each sub-catalogue are shown in Fig. 5.3. From this figure, it is possible to observe that the number
of galaxies in the sub-catalogue whose magnitudes are smaller than 20 is smaller than the other sub-catalogues.

The convergence map corresponding to each sub-catalogue is determined using FIATMAP. The convergence maps, shown in Fig. 5.2, are given in a grey scale where the white means over-densities in mass and black means empty spaces. The values of the block $b$, inner radius $r_{in}$, outer radius $r_{out}$ are 100, 300 and 3000 respectively. Note that the dimension of the images obtained by FIATMAP are $101 \times 82$ pixels since we have run FIATMAP for blocks of $b = 100 \times 100$ pixels at the original picture, after the edge removal has a dimension $(10100, 8200)$ pixels.

### 5.2 Noise Estimation of the Convergence Maps

![Example of Randomized Maps](image)

Fig. 5.4 Example of Randomized Maps corresponding for sub-catalogues whose isophotal magnitude $m_{iso}$ goes from 17 to 20 (left), 20 to 22 (middle), and 22 to 25 (right).

In addition to the DM density map associated to each sub-catalogue, we generated random similar maps statistically coherent with the real image. FIATMAP offers the possibility of creating a “randomized” map together with the real convergence map by implementing a Monte Carlo method. The algorithm assigns to a given galaxy the ellipticity and size of another one, obtaining so a random distribution of galaxies, where the positions are kept but the properties are changed randomly. Therefore, the alignment between adjacent galaxies contained in the catalogue is destroyed. The set of randomized maps gives an idea of the behaviour of the spurious alignments occurring. For each galaxy catalogue, we create 1000 randomized maps from the original convergence maps.

Comparison of the real convergence map obtained from the galaxy catalogue with the 1000 randomized maps provide a measurement of the probability of finding any mass over-densities in the field of study. We have checked that the distribution of the of the value corresponding to random convergence maps follows a Gaussian distribution, and the value is almost the same for all the pixel of the average $\sigma_{<20} = 0.009418 \pm 0.0000005$, $\sigma_{20-22} = 0.0051154 \pm 0.0000004$ and $\sigma_{>22} = 0.0143900 \pm 0.000002$. The mean dispersion of these convergence maps, at each pixel, remarks the average noise per pixel which provides a stable estimate of the noise. A sample of randomized maps are shown in Fig. 5.4
Signal-to-noise $S/N$ maps are just computed by dividing the convergence map of the data by the calculated standard deviation $\sigma$ that characterizes the noise for each pixel intensity. Signal-to-noise maps are shown in Fig. 5.5. The higher number of galaxy density present in the Lockman Hole decreases the noise in the convergence maps and therefore increases the signal-to-noise of peaks due to massive clusters.

Fig. 5.5 $S/N$ maps corresponding for sub-catalogues whose isophotal magnitude $m_{\text{iso}}$ goes from 17 to 20 (left), 20 to 22 (middle), and 22 to 25 (right).

5.2.1 Importance of PSF anisotropies correction

Fig. 5.6 Output $S/N > 3.5$ maps corresponding to the galaxies catalogue from the image taken in the R-band whose objects have not been corrected from PSF anisotropies for the sub-catalogue containing objects whose magnitude goes from 20 to 22 (at left) and from 22 to 25 (at right). These images are analogous to middle and right images of Fig. 5.5.

If the PSF anisotropies remain uncorrected, final signal-to-noise maps would show galaxy cluster candidates which are in fact spurious objects due to a systematic error which has not been treated. This is demonstrated in Fig. 5.6, where the analysis method has been applied to the galaxies catalogue
obtained from the image taken in the R-band, but whose PSF anisotropies have not been corrected according to the method described in section 4.1.2. We detect possible intensity peaks, and therefore galaxy clusters in both maps. For the signal-to-noise map of the sub-catalogue containing galaxies whose magnitude goes from 10 to 22 there is an intensity peak in agreement with those shown in Fig. 5.5. Nevertheless, the rest of candidates are shown at the edges of the subfield, where PSF anisotropies are much considerable according colour-maps plotted in Fig. 4.7 and 4.8. For this reason, these galaxy cluster candidates are not trustable because there are likely due to the effect that PSF anisotropies have of galaxies shapes.
In the previous chapter we have presented the obtained signal-to-noise maps after the analysis method. In this chapter the study of these maps are presented, and they are compared to previous x-ray observations in the Lockman Hole area and to other WL results.

### 6.1 Galaxy Cluster Candidates

![Convergence maps](image)

The highlighted pixels shown in Fig. 6.1 may be understood as intensity peaks. Each intensity peak is a candidate for cluster of galaxies. No peak was found for the first sub-catalogue, and the one of the
Comparison with previous observations

main reasons why is as shown in Fig. 5.3, this sub-catalogue contains probably few galaxies in order to compute the convergence $\kappa$.

The signal-to-noise maps corresponding to the $20 < m_{iso} < 22$ have 5 peaks with $S/N > 3.5$ (is shown in the top-right panel of Fig. 6.1). For the signal-to-noise maps corresponding to the $22 < m_{iso} < 25$ sub-catalogue we find 1 peaks with $S/N > 3.5$. Let us remarks that the peaks are identified as possible galaxy clusters, and they are listed in the table 6.1. From this list, the largest WL effect comes from the candidates C2 and C3 in Fig. 6.1, as they show a higher signal-to-noise ratio and a larger size (taking into account the number of blocks of $100 \times 100$ pixels).

<table>
<thead>
<tr>
<th>ID</th>
<th>Colour</th>
<th>$S/N$ value</th>
<th>RA/hour angles</th>
<th>DEC/deg</th>
<th>$m_{iso}$ bin</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>Red</td>
<td>4.32</td>
<td>10 : 53 : 55</td>
<td>+57 : 34 : 20</td>
<td>20-22</td>
</tr>
<tr>
<td>C3</td>
<td>Yellow</td>
<td>4.51</td>
<td>10 : 52 : 10</td>
<td>+57 : 50 : 32</td>
<td>20-22</td>
</tr>
<tr>
<td>C5</td>
<td>Green</td>
<td>3.51</td>
<td>10 : 52 : 44</td>
<td>+57 : 35 : 34</td>
<td>20-22</td>
</tr>
</tbody>
</table>

Table 6.1 Galaxy clusters candidates in the studied Lockman Hole area. ID identifies the candidate in Fig. 6.1 through the Colour, $S/N$ specifies the signal-to-noise value for which they were detected, RA is the right ascension coordinate in hour angles and DEC is the declination in degrees and $m_{iso}$ bin remarks the interval in magnitudes corresponding to each clusters.

6.2 Looking For Counterparts

The cluster candidates included in Table 6.1 can be matched with already observed galaxy clusters detected on previous studies carried out in the Lockman Hole area. For this aim, we obtain information regarding this particular area about x-ray emission. Additionally, we use the NASA/IPAC Extragalactic Database (NED) in order to match the highest signal-to-noise peaks with galaxy clusters.

6.2.1 X-ray Observations

We have checked our galaxy cluster candidates with those candidates proposed by Finoguenov et al. [16] after searching for diffuse source in the $0.4 \, \text{deg}^2$ area of the XMM-Newton 700 ks survey of the Lockman Hole central area (see Fig 6.2). They checked their x-ray candidates with optical images, stating that they correspond to a concentration of red galaxies at redshift near $z = 1$ creating low-mass galaxy clusters. They derive a catalogue of 14 cluster candidates within an area $0.5 \times 0.7 \, \text{deg}^2$. The sub-field that Finoguenov et al. have analysed overlaps partially with the Lockman Hole sub-field that we have studied in this work. In the area of overlapping, which is around 108000 pixels$^2$, only 3 out of these 6 candidates are included. They are labelled in Table 6.2 and shown in Fig. 6.2.
6.2 Looking For Counterparts

Fig. 6.2 X-Ray image for the Lockman Hole deep field study by [16]. The 14 new clusters are marked in yellow and white.

<table>
<thead>
<tr>
<th>ID</th>
<th>RA/hour angles</th>
<th>DEC/deg</th>
<th>R/arcsec</th>
<th>M/10^{13}M_{\odot}</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td>A08</td>
<td>10:53:59.15</td>
<td>+57:34:20.6</td>
<td>36</td>
<td>4.0839</td>
<td>0.440</td>
</tr>
<tr>
<td>A10</td>
<td>10:53:46.69</td>
<td>+57:42:39.0</td>
<td>28</td>
<td>1.4815</td>
<td>0.080</td>
</tr>
<tr>
<td>A12</td>
<td>10:53:33.72</td>
<td>+57:43:01.36</td>
<td>26</td>
<td>1.5393</td>
<td>0.560</td>
</tr>
</tbody>
</table>

Table 6.2 Galaxy cluster candidates present in the catalogue obtained in [16] that are present in our analysed field. ID identifies the candidate according to the label used in [16]. RA is the right ascension coordinate in hour angles, DEC is the declination in degrees, R is the proposed radius for the cluster candidates in arcsec, M is the estimate mass of the cluster in solar masses and z is the redshift estimate for the cluster.

Fig. 6.3 X-Ray Galaxy clusters candidates studied in [16] present in the overlapping area, from left to right: A08, A10 and A12. The images correspond to the optical image of the sub-field of study observed in the R-band. The white circles indicate the position of the cluster candidates according to Table 6.2 as well as their proposed radius R. The red circles point out the galaxies included in the final galaxies catalogue from which convergence maps were obtained.
Comparison with previous observations

We can identify our cluster candidates C1 with A08 and C4 with A12. As it will be expected, we detected the bigger cluster candidate A08 for a signal-to-noise larger than 4, which is larger than the signal-to-noise value associated for the cluster candidate A12, which was only observed for a signal-to-noise value of 3.59. It is not possible to identify the A10 cluster candidate as we do not obtain any signal-to-noise peak in our convergence maps for that position, for a signal-to-noise value larger than 3.5. Probably, this is due to the fact that in our final galaxies catalogue we did not include most of compact objects within the white circle shown in the middle picture of Fig. 6.4 as a saturated star blocks part of the further celestial objects and we avoid to take any saturated objects in the final galaxies catalogues by applying a cut in the isophotal flux. Finoguenov et al. use optical images taken from the Subaru Telescope in 2001 to check their x-ray cluster candidates. In Fig. , it is possible to observe how the same saturated star blocks part of the further galaxies. Thus, this cluster of galaxies could not be found by WL effect from observations in the optical without removing properly the saturation associated to this kind of stars and trying to avoid missing essential information.

![A10](image)

Fig. 6.4 $R$-band image with the X-Ray contour overlays for the cluster candidate A10 obtained from [16].

From this comparison, it is significant the relation between the signal-to-noise value and the mass associated to the galaxy cluster. The candidate C1 is one of our cluster candidates that shows a higher signal-to-noise value, and it is also the most massive and larger galaxy cluster observed in x-ray in our overlapping area. Hence, it is possible to expect a phenomenological relation between the signal-to-noise value and the mass of the cluster, in spite of the fact that the empirical relation implemented in FIATMAP is not prepared for obtaining a correct value of the mass of the cluster.

6.2.2 Comparisons With Previous Weak Lensing Observations

Apart from the x-ray matched galaxy clusters, we can identify the two galaxy clusters that have the largest signal-to-noise ratio detected in our convergence maps concerning the catalogue with cuts in magnitude from 20 to 22, which are C2 and C3, to previous galaxy clusters detected by WL studies. In
the study carried out by Miyazaki et al. [24], they analysed a fraction of the Lockman Hole area using the inverse method of Kaiser and Squires [23] to obtain the convergence $\kappa(\theta)$. They have also detected two relatively massive galaxy cluster candidates compatible with our cluster candidates. Moreover, the cluster candidate detected in the convergence map corresponding to the catalogue that contains object whose magnitude goes from 22 to 25, C6, is also found in Miyazaki et al. studies.

### 6.3 Probabilistic Study

Once that we have matched suitable cluster galaxies, we wonder whether these results are consistent. We can calculate the ratio of the area occupied by candidate galaxy clusters, supposing that the mass is distributed in a circular area, such as:

$$F = \frac{A_{gc}}{A_t} = \frac{\sum_i \pi R_i^2}{A_t}, \quad (6.1)$$

where $R_i$ are the radii of the $N = 14$ galaxy clusters detected by Finoguenov et al., giving a total area occupied by the candidate galaxy clusters $A_{gc} \approx 1431313 \text{ pixels}^2$, and $A_t$ is the area of the studied sub-field of $0.5 \times 0.7 \text{ deg}^2$. We obtain a value for this fraction $F = 1.26\%$. If we consider that the sub-field studied by Finoguenov et al. is characteristic and can be extrapolated to the rest of the sky, we can deduce the number $N(lhn1n1)$ of galaxy cluster that we may find in our subfield. This number of galaxy clusters suitable to be found in our field $lhn1n1$ is defined such as:

$$N(lhn1n1) = \frac{A_{ge}(lhn1n1)}{\langle A \rangle_{gc}} = \frac{A_t(lhn1n1) \times F \times N}{A_{gc}}, \quad (6.2)$$

where $A_{gc}(lhn1n1)$ is the area occupied by our hypothetical galaxy clusters to be found in our field, $\langle A \rangle_{gc}$ is the mean value associated to a characteristic galaxy cluster found by Finoguenov et al. defined as $\langle A \rangle_{gc} = \frac{A_{gc}}{N}$, and $A_t(lhn1n1)$ is the effective area of the field $lhn1n1$ taking part the area computed to edges and saturate compact bodies. The result is an area of $A_t(lhn1n1) = 63707692 \text{ pixels}^2$. Therefore, the number of galaxy clusters $N(lhn1n1)$ that we should expect to find in our field is $7.9 \pm 2.8$ assuming that is Poissonly distributed. This result is in agreement with the obtained number of galaxy cluster at 1$\sigma$, which is equals to 6 (see Table 6.1).
An analysis technique to detect galaxy clusters based on WL effect has been successfully developed and applied to a Lockman Hole subfield. We have performed the extraction of compact objects from a R-band and a Z-band images, ridding of spurious objects belonging to those catalogues through a pipeline that included a cross-matching process, simulations and correction of systematics among others. In addition, we have classified these compact objects into stars and galaxies, re-calculating their shapes and correcting the PSF anisotropies, obtaining a final galaxies catalogue suitable to be analysed for estimating DM halos. We have found six galaxy clusters candidates, where five of them have been matched with previous galaxy cluster detected in x-ray and WL observations.

Systematic errors have been identified and avoided during the analysis process. In order to evaluate the quality of the catalogues generated by SOURCE EXTRACTOR that are the basis for WL measurements, the full analysis include the validation of those catalogues by means of simulation of mocked galaxies and cross-matching techniques so that spurious objects are not included in the study. The catalogues were cleaned from saturated celestial objects by applying a cut in the isophotal flux as well as those coming from detection on the edges. The objects they content were classified according its star or galaxy nature. In addition, the shape of these celestial objects associated by SOURCE EXTRACTOR were improved by performing an alternative shape estimation using an specific software. Finally, the PSF anisotropies associated to different sources were corrected both from the R-band image and the Z-band one. This is a key aspect since WL is extremely sensitive to uncorrected anisotropies, as it is has been demonstrated.

From the carried-out work, we have checked how the cross-matching process is an useful tool to rid of spurious objects, not only those coming from the background but compact object detected close to saturated stars. Saturated stars do not look the same observed in different bands, and therefore SOURCE EXTRACTOR detects objects close to saturated stars differently. As the cross-matching process is based on geometrical construction of distances, a variation in the position of spurious compact objects close to saturated stars causes that those objects are not matched.

During the study we have accounted the above-mentioned systematic errors. However, for future studies it is possible to deal with saturated pixels by creating a mask to hide those pixels so that SOURCE EXTRACTOR can ignore those contributions. Furthermore, we can apply the developed analysis pipeline to several fields observed in different bands in order to detect already-observed galaxy clusters or even to discover new ones together with their corresponding DM.

DM halos are considered to be the building-blocks of cosmic structure. The detection of galaxy clusters can provide an idea of the distribution of DM in the universe so that WL studies may constraint cosmological models and the parameters that define them. WL measurements are methodologically arduous but simple, offering a robust cosmological application to study the larger structure of the universe.
REFERENCES


\textbf{A.1 General Information}

\textsc{Weak-Lensing software} is a set of different programs designed to work as a Weak Lensing analysis pipeline. It is placed on a Github repository at \url{https://github.com/gcanasherrera/Weak-Lensing.git}, containing classes and scripts required to detect clusters of galaxies provided an image of a sky region. The detection of these cluster of galaxies is performed thanks to the concept of gravitational lensing: light coming from distant objects suffers a deviation due to the mass present in the universe. When this effect is subtle (meaning weak), this deviation produces a shear in the size of distant objects susceptible to be measured.

\textsc{Weak-Lensing} software is fully written in \textsc{Python 2.7} and it basically consists on scripts assembling the DLS pipeline. It works together with the programs contained in this pipeline written mainly by Professor David Wittman\textsuperscript{1}. This package is written in \textsc{C}, \textsc{Fortran} and \textsc{Perl}, and it is divided into many tools to work with DLS catalogues and shear-measurement programs. DLS catalogues are in FIAT format designed by D. Wittman himself and whose tools are open to anybody at \url{http://dls.physics.ucdavis.edu/fiat/fiat.html}. \textsc{Weak-Lensing} software also maintained to use the FIAT 1.0 format, taking advantage of some of these FIAT tools. On the other hand, DLS pipeline shear-measurement programs are not public. The availability of the DLS pipeline programs is recorded in Table A.1, and the pseudocode or main algorithm of the DLS program may be explained in the bibliographical sources provided during the Final Degree Project report as they were introduced in the text.

\textsc{Weak-Lensing} software is registered under the GNU GENERAL PUBLIC LICENSE, and all rights and responsibilities can be inferred from it.

\textbf{A.2 Author Information}

Guadalupe Cañas Herrera is the only author of \textsc{Weak-Lensing software}, and she is the responsible of new updates of version in the future. For any query regarding the source code, the user may request her via email at gch25@alumnos.unican.es or canasg@ifca.unican.es.

\textsuperscript{1}Professor D. Wittman works currently at UC Davis. His personal web-page is \url{http://wittman.physics.ucdavis.edu/contact.html}. Most of the software he developed is not public as it belongs to the DLS Collaboration. Therefore, any modification of those codes performed by the author of this documentation will be explicitly remarked.
A.3 General Requisites

In order to run WEAK-LENSING software, it is necessary to have installed in the computer and/or virtual machine the following requisites. Furthermore, the use of a grid, such as any computational cluster or supercomputer, is highly recommended due to the great amount of data to process.

- **UNIX Operative System**: This is not an essential requisite, but the use of the software has not been tested in Windows. If you have Windows as a OP, you may try to work with ENTHOUGHT CANOPY available at https://www.enthought.com/products/canopy/.

- **PYTHON 2.7 and IPYTHON**

- **Python Library Packages**: ASTROPY, NUMPY, MATPLOTLIB, PYMODELFIT, SEABORN, SCIPY. Most of them are included in ANACONDA package, a completely free Python Distribution containing most famous Python Packages available at http://docs.continuum.io/anaconda/install, or they are easily downloadable thanks to the CONDA command given by ANACONDA itself such as PYMODELFIT and SEABORN. I strongly recommend using ANACONDA to avoid further problems with hidden dependencies (for instance, SEABORN depends on the package PANDAS, but this one is directly included in ANACONDA). If you choose working with ANACONDA, be sure that you are calling the PYTHON compiler included in this package, and you are not using by default the PYTHON compiler included in most UNIX distributions. When ANACONDA is installed, it usually builds up automatically the PATH to ANACONDA location in order to simply start to use this PYTHON distribution and their compilers. However, it is useful to check that the PATH is corrected by typing in a terminal:

```
which python
```

You should see something like:

```
/thepath/anaconda/bin/python
```

In case you observe something different, it is probably because you are using the PYTHON compiler not placed in ANACONDA. Consequently, you should change yourself the PATH to ANACONDA into the .bashrc file,

```
ls -a
vim .bashrc
```

and write down the path to python compiler in Anaconda,

```
#Call Anaconda
export PATH="/thepath/anaconda/bin/python:$PATH"
```
You can also work with the default PYTHON compiler and install the required packages yourself thanks to a Package Installer such as PIP (also available in ANACONDA or at https://pip.pypa.io/en/stable/) or HOMEBREW.

- **SOURCE EXTRACTOR** v.2.3.2 (older versions such as 2.19.5 may work although measurement results are not as precise as v. 2.3.2 and further modifications in the scripts may need to be carried out). Installation is complicated in OS X (it is suggested trying with HOMEBREW instead of re-compiling from the source code) and in Linux (if you lack of administrator permission to use apt command to re-compiling from scratch). Once it is installed, it is recommended to open a terminal and attach the the PATH to SOURCE EXTRACTOR executable into the .bashrc file,

```
ls -a
vim .bashrc
```

and write down the path to SOURCE EXTRACTOR executable,

```
#Call Source Extractor
export PATH="yourpath/sextractor/bin:$PATH"
```

This way, it is not needed to modify any program from WEAK-LENSING software in order to make SOURCE EXTRACTOR work, as WEAK-LENSING software scripts use sub-processes via terminal to call SOURCE EXTRACTOR. It also requires default files as an input to start to work. You can find those default files in SOURCE EXTRACTOR directory.

- **SAOIMAGE DS9** in order to visualize pictures automatically from some scripts. Installation is really simple, and it can be found at http://ds9.si.edu/site/Download.html. One it is installed, it is recommended to open a terminal and attach the the PATH to SAOIMAGE DS9 into the .bashrc file,

```
ls -a
vim .bashrc
```

and write down the path to SAOIMAGE DS9 executable,

```
#Call ds9
export PATH="yourpath/ds9:$PATH"
```

- **DLS PIPELINE SOFTWARE**: SEX2FIAT, FIATFILTER, ELLIPTO, DLSCOMBINE, FIATMAP, FIATREVIEW. Again, it is highly recommended to attach the the PATH to these executables into the .bashrc file. For instance,
ls -a
vim .bashrc
#Call Fiatmap
export PATH="yourpath/fiatmap:$PATH"

<table>
<thead>
<tr>
<th>DLS software</th>
<th>Privacy</th>
<th>Author</th>
<th>Language</th>
<th>Dependencies</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEX2FIAT</td>
<td>public</td>
<td>D. Wittman</td>
<td>PERL</td>
<td>fiat.c, fiat.h</td>
<td>Traduce SOURCE EXTRACTOR Catalogue into a FIAT catalogue</td>
</tr>
<tr>
<td>FIATFILTER</td>
<td>public</td>
<td>D. Wittman</td>
<td>PERL</td>
<td>fiat.c, fiat.h</td>
<td>obtain only required data that fulfils a logical condition from the catalogue</td>
</tr>
<tr>
<td>DLSCOMBINE</td>
<td>private</td>
<td>D. Wittman/G. Cañas</td>
<td>C</td>
<td>fiat.c, fiat.h, dlscombine_utils.c, dlscombine.c,</td>
<td>correct PSF anisotropies of the images</td>
</tr>
<tr>
<td>FIATMAP</td>
<td>private</td>
<td>D. Wittman</td>
<td>C</td>
<td>fitsio.h, fiat.h</td>
<td>Generates fits image based on mass-maps thanks to invlens algorithm</td>
</tr>
<tr>
<td>ELLIPTO</td>
<td>private</td>
<td>D. Wittman</td>
<td>C</td>
<td>fiat.h</td>
<td>Corrects objects’ shapes</td>
</tr>
<tr>
<td>FIATREVIEW</td>
<td>public</td>
<td>D. Wittman</td>
<td>C</td>
<td>PGPLOT, cfitsio, wcstools</td>
<td>Plots a FITS image and a FIAT catalogue at the same time</td>
</tr>
</tbody>
</table>

Table A.1 Description of the different DLS software used during the project, their characteristics and authors. The public fiat tools may be found at http://dls.physics.ucdavis.edu/fiat/fiat.html

I can provide with the binaries/executables of DLS PIPELINE used in this Final Degree Project for both OS X and Linux. To obtain the source code, I suggest contacting directly to Professor Wittman.
A.4  Source Installation

For installation, simply copy the Github repository and execute the main script via terminal. Open a new terminal tab and type:

```
git clone https://github.com/gcanasherrera/Weak-Lensing.git
cd Weak-Lensing/
```

To execute any script, I recommend using IPYTHON in interactive mode (so variables kept in memory at the end of the script and you can work with them). To execute the main script, you will need to pass to the script a FITS picture and a catalogue to start the analysis. To try, type:

```
ipython -i WL_Script.py
```

A.5  Content

The WEAK-LENSING software contains a set of several programs divided into four different types:

1. Auxiliary programs destined to assist main scripts.
2. Validation programs to check catalogues.
3. Programmes destined to the study of systematics and classification of compact objects.
4. Dark matter halo estimation programmes.

A.5.1  Auxiliary Programs

Read Catalogues

To read catalogues obtained from SOURCE EXTRACTOR it is essential to transform them into a FIAT format first, and then translate the information into NUMPY arrays. For that, I have defined the PYTHON class CLASS_CATALOGREADER.PY that contains general methods to read FIAT catalogues and transform them into structured NUMPY arrays. To work, it uses the following logical pseudocode to understand if the catalogue comes from SOURCE EXTRACTOR or it is already in FIAT format:

```
if catalogue_name ends with .cat (#Catalogue is in Source Extractor format)
call sex2fiat to transform
    then read
if catalogue_name ends with .fcat (#Catalogue is in FIAT format)
    then read
```
Therefore, the endings ".cat" is reserved for the catalogue given by SOURCE EXTRACTOR and ".fcat" to those in FIAT format.

To access the information of the catalogue we just create an object of this class and read the catalogue. The information is saved as an attribute of the object, for instance:

```python
#Object from the CatalogReader class
from Class_CatalogReader import CatalogReader
c = CatalogReader('string_nameofcatalogue')
c.read()
information = c.fcat
x_coordinate = information['x']
```

**General tools**

Here we can find PYTHON files that complete main scripts.

First, the PYTHON file WL_UTILS.PY contains three different types of designed methods to:

1. call quickly executables such as SOURCE EXTRACTOR, ELLIPTO, DLSCOMBINE and SAOIMAGE DS9 through the Python Library SUBPROCESS.

2. short-cut the creation of ellipticity components plots and general FWHM vs. $m_{iso}$ plots.

3. create text files necessary for DLSCOMBINE to work (see details below).

Second, the PYTHON file WL_ELLIP_FITTER.PY works a a fitter of the ellipticity components as a function of Legendre polynomials.

**A.5.2 Validation Programs: Cross-Matching and Add-Object Simulation**

**Cross-Matching Process**

This software checks which objects are found in two different catalogues 1 and 2 regarding their coordinates (the numbers may refer to any band filter used in the work). It performs the following algorithm:

1. Transform objects’ coordinates from catalogue 1 and 2 into spherical coordinates,

\[ x = \sin \phi \cos \theta; y = \sin \phi \sin \theta; z = \cos \theta, \quad (A.1) \]

where $\phi$ is the known as the declination DEC and $\theta$ is the right ascension RA. It is necessary to use spherical coordinates defined from the declination and the right ascension due to the fact the fields for which the cross-matching are performed may not share the same Cartesian coordinates $(x', y')$, but they do regarding spherical coordinates (DEC and RA are the same for both). In order to avoid any extra transformation to one Cartesian coordinates system to another, we just simply use coordinates defined at Eq. (A.1).
2. Look for the objects from catalogue 2 in catalogue 1 (cross-matching code 1 to 2). The search is performed using the KD-Tree algorithm [1] with a confidence radius \( r = kr_p \), being \( k \) an integer and \( r_p \) the size of the pixel in spherical coordinates. The objects of the catalogue 2 are referred as keys. Count the number of object that are not found in catalogue 1 as \( n_{lost} \). Write a new catalogue (called catalogue 1 to 2) containing the objects found in catalogue 1 according to catalogue 2 keys.

3. Repeat step 3, in this case looking for the objects from catalogue 1 in the catalogue 2. Write a new catalogue (called catalogue 2 to 1) that contains the objects found in catalogue 2 according to catalogue 1 keys.

To carry out this task, the cross-matching process uses three different programs:

- **CLASS_CROSSMATCHING.PY**: PYTHON class. It contains methods to translate Cartesian coordinates to spherical ones, to construct the KD-Tree structure from the library SCIPY, and to write the final cross-matched catalogue thanks to a for-loop.

- **WL_CROSSMATCHING.PY**: PYTHON file. It is the main script to perform the cross-matching process between two catalogues. To produce the cross-matched catalogue, you should first read both catalogues using CLASS_CATALOGREADER.PY and then create an object from the class CLASS_CROSSMATCHING.PY. You have to choose the confidence radius \( r \) for the cross matching when calling the method KDTREE() and if you want to compare 1 to 2 or 2 to 1.

```python
#Read catalogues catalogue_1 and catalogue_2 to obtain numpy st. array
#Create object from class CrossMatching
crossmatching = CrossMatching(catalogue_1, catalogue_2)
k=3
crossmatching.kdtree(r=k*1e-06)
#argument 'compare' may be set to '1to2' or '2to1'
crossmatching.catalog_writter('name of the array', compare = '1to2')
```

### Add-Object Simulation

This simulation is designed to throw mock galaxies, whose features are known, into a FITS file to test a detection tool such as SOURCE EXTRACTOR. The algorithm is as follows:

1. Read the catalogue obtained by SOURCE EXTRACTOR.

2. Calculate the mean values of the detected objects’ magnitudes \( \langle m \rangle \), ellipticities \( \langle \varepsilon \rangle \) and minor axis \( \langle B \rangle \) directly from the data contained in the catalogue.

3. Calculate the mean value of the detected objects eccentricity as,

\[
\langle ec \rangle = \sqrt{1 - \left(1 - \langle \varepsilon \rangle \right)^2}
\] (A.2)
4. Calculate the mean value of the detected objects’ major axis \( \langle A \rangle \) as,
\[
\langle A \rangle = \frac{\langle B \rangle}{\sqrt{1 - \langle ec \rangle^2}}
\] (A.3)

5. Establish a relation between the \textit{isophotal flux} \( F_{iso} \) and the \textit{isophotal magnitude} \( m_{iso} \) data from the catalogue according to the general expression in [11],
\[
F = a 10^{-m/2.5+b}
\] (A.4)
where \( a \) and \( b \) are parameters to be calculate thanks to a non-linear Least Square Fitting that minimizes the sum of the squares of the residuals between the data and the model.

6. Read the FITS image in order to obtain a matrix whose size agrees with its from the picture. Each element of the matrix corresponds to a image pixel. We give matrix elements a value equals 1 if \textsc{source extractor} detected celestial objects in their matching pixels and give value equals 0 otherwise.

7. Calculate \( n \) pairs of random numbers for values of \( x \) and \( y \) belonging to the image size. We check if the matrix element noted by the position of those random pairs \((x, y)\) is equals to 0. Save those pairs into a text file.

8. Create elliptical own-made galaxies for which their \textit{maximum magnitude} \( m_{max} \) are fixed. Thanks to relation (A.4), calculate the corresponding \textit{maximum flux} \( F_{max} \) according to the \( m_{max} \). The distribution of the intensity along the ellipse is defined by elliptical Gaussian function expressed as,
\[
f(x, y) = Fe^{-\frac{1}{2} \left[ \frac{(x-x_0)^2}{(A)^2} + \frac{(y-y_0)^2}{(B)^2} \right]}
\] (A.5)

9. Launch \( n \) number of mock galaxies into the \( n \) pairs of random numbers for values of \( x \) and \( y \) in the null elements of the matrix.

10. Fulfil the matrix elements still equals zero after launching the own-made galaxies with the corresponding values of the original image pixels. Copy that matrix into a FITS image.

11. Analyse with \textsc{source extractor} the new FITS image containing the new \( n \) galaxies. Extract a new catalogue.

12. Look for the information corresponding to the new \( n \) mock galaxies in the new catalogue. Compare the mean value of the mock galaxies \textit{isophotal magnitude} attached by \textsc{source extractor} regarding the value of the input magnitude that was given to the galaxies when were launched.

13. Count the number of own-made galaxies not found by \textsc{source extractor} \( n_{lost} \).
14. Repeat steps from 7 to 13 for own-made galaxies whose given magnitude goes from $m_{\text{max}} = 1$ until $m_{\text{max}} = 30$.

Note the difference between the isophotal magnitude $m_{\text{iso}}$ and the maximum value of the magnitude, or maximum magnitude $m_{\text{max}}$. The maximum value of the magnitude corresponds to the magnitude value corresponding to the faintest pixel in the elliptical gaussian function, whereas the isophotal magnitude refers to the total value attached to the integrated magnitude for the object. Therefore, the maximum magnitude $m_{\text{max}}$ will be always smaller than the isophotal magnitude $m_{\text{iso}}$, as the magnitude is the inverse of the intensity flux $F$ so that $F_{\text{iso}} > F_{\text{max}}$ and consequently $m_{\text{iso}} < m_{\text{max}}$. Concretely, the relationship between $F_{\text{iso}}$ and $F_{\text{max}}$ is, $F_{\text{iso}} \approx 2\pi \langle A \rangle \langle B \rangle$, where $2\pi \langle A \rangle \langle B \rangle = 7.2$ and experimentally we obtain around 7. This relation is necessary because the introduced input data is the value of the maximum magnitude $m_{\text{max}}^i$, we want to attach to mock galaxies.

To carry out the task explained in the algorithm, the simulation uses three different programs:

- **CLASS_CATALOGREADER.py**.

- **CLASS_OBJECTCREATOR.py**: PYTHON class. Contains general methods to fit magnitudes and fluxes (through the package PYMODELFIT), open and write FITS files (using the package ASTROPY), plots histograms of the properties $A$, $B$, $\varepsilon$ and $\varepsilon c$, creates mock galaxies and search tools based on three different tools:
  - **SEARCHER(c1.fcat, c2.fcat)**: loops-for the first catalogue and looks for each element of the first catalogue looping-for the second catalogue. It works but it is highly inefficient and slow.
  - **SEARCHER_DIC(c1.fcat, c2.fcat)**: looks for elements of the first catalogue in the second catalogue by means of python dictionaries. It offers an advantage with respect the first method meaning that it is quick and efficient, however it reduces $(x, y)$ coordinates to integers values, missing significant digits.
  - **SEARCHER_KDTREE(c1.fcat, c2.fcat)**: looks for elements included in the first catalogue in relation to those in the second catalogue by the implementation of the KD-Tree algorithm available at the library SCIPY.

- **WL_SIMULATION_EXECUTE_MAG.py**: PYTHON file. It is the main script containing instruction to develop steps described in the algorithm above. It saves the mean values of the quantities $n_{\text{lost}}$, $m_{\text{iso}}^i$, $m_{\text{iso}}^o$, $F_{\text{iso}}^i$, $F_{\text{iso}}^o$, $F_{\text{max}}^i$, $F_{\text{max}}^o$ and their uncertainties associated in text files (upper index $i$ and $o$ mean input and output, respectively). The mean value is obtained from $n$ mock galaxies. The simulation is also useful to test the reliability of different smoothing filters such as Mexican Hat or Gaussian. To start the simulation pass as parameters the name of the picture in which you want to throw the mock galaxies and the name of the filter SOURCE EXTRACTOR will use, typing:
ipython -i WL_Simulation_execute_mag.py PICTURE_NAME FILTER_NAME

- **PLOTTER_OBJECTCREATOR.PY**: PYTHON file. Script that reads the text files saved by the script explained above and plots the relations between $m_{iso}^i$ and $m_{iso}^o$, $F_{iso}^i$ and $F_{iso}^o$, $F_{\max}^i$ and $F_{\max}^o$, and $m_{iso}^i$ and $n_{lost}$.

This simulation needs to open and close at least 60 times pretty heavy FITS files, and run *Source Extractor* at least 30 times. For this reason, the use of a supercomputer is highly recommend. In the development of this work, and in order to send works to the tail of ALTAMIRA in an user-friendly way, I split the simulation and the plotting part (PLOTTER_OBJECTCREATOR.PY). Therefore, once that the features of the galaxies are saved in files by running just one time the simulation, it is possible to make any change on the plots just executing PLOTTER_OBJECTCREATOR.PY easily.

### A.5.3 Study of Systematics and Classification of Compact Objects

The classification of compact objects into galaxies and stars, as well as the re-estimation of objects shapes and the correction of PSF anisotropies is performed by a general script named WL_SCRIPT.PY. This script performs the following pseudocode:

1. Read the number of pictures that perform the first cross-matching process and their corresponding observations bands (for instance $R$ and $Z$). Then, reads the name of the two cross-matched catalogues and the name of the corresponding FITS files.

2. Read the first cross-matched catalogue and save the information it contains.

3. Plot Cartesian coordinates $(x, y)$ corresponding to the compact objects in order to filter saturate edges. Perform the cut in $F_{iso}$ to avoid saturated stars.

4. Plot the $FWHM$ vs. the magnitude $m_{iso}$ of the compact objects without saturate detections.

5. Filter stars according to the study of the $FWHM$ vs. the magnitude $m_{iso}$ and *SOURCE EXTRACTOR* neural network. Plots the ellipticity vector components.

6. Filter galaxies according to the study of the $FWHM$ vs. the magnitude $m_{iso}$ and *SOURCE EXTRACTOR* neural network. Plots the ellipticity vector components.

7. Plot the histogram corresponding to the stars and galaxies population.

8. Recalculate galaxies and stars shapes using ELLIPTO. Filter those objects that have an error code higher than 2 (meaning that the shapes given by ELLIPTO differ considerably from those given by *SOURCE EXTRACTOR*).

9. Plot the ellipticity vector components of the new re-estimate compact object shapes.

10. Fit the ellipticity vector components $(e_1, e_2)$ of the stars as a function of Legendre polynomials $L(x)L(y)$. Write the fitting coefficients into a text file.
11. Write a text file suitable to be understood by DLSCOMBINE containing the path to the fitting coefficients text file and the name of the analysed image (denominated `specfile`).

12. Call DLSCOMBINE to correct PSF anisotropies from the analysed image.

13. Repeat the steps from 1 to 12 with a different image observed in another band and/or cross-matched catalogue of the same region of study.

14. Carry out cross-matching process between the two corrected images.

15. Repeats steps from 2 to 9 obtaining a final galaxies catalogue.

WL_SCRIPT.PY takes advantage of the software FIATFILTER to perform the several filtering of the catalogues. Furthermore, it is helped by two other programs:

- WL_UTILS.PY.
- WL_ELLIP_FITTER.PY: PYTHON file. Contains general methods to fit the ellipticity components \((e_1, e_2)\) as a function of Legendre polynomials by a procedure of non-linear least square fitting minimizing the following quantity,

\[
\chi_i^2 = \sum_x \sum_y \left( e_i(x,y) - \sum_n \sum_m c_{nm}(x,y) \frac{L_n(x)L_m(y)}{\sigma_e(x,y)} \right)^2,
\]  

where \(L_n(x)\) and \(L_m(y)\) are Legendre polynomials, of order \(n\) and \(m\), respectively, \(c_{nm}(x,y)\) are the fitting coefficients and \(\sigma_e(x,y)\) are the relative error associated to the ellipticity \(e\) by ELLIPTO used to weight the fitting. Due to the fact that DLSCOMBINE is designed to read polynomial fittings up to a maximum grade equals to 4, the value of \(n\) and \(m\) must be \(n + m \leq 4\), so that \(n, m = \{0, 1, 2, 3, 4\}\). The fitting is performed thanks to the package ASTROPY using a 2-Dimensional Legendre polynomial base as fitting model and the non-linear least square statistics provided by the package SCIPY following the Levenberg-Marquardt algorithm. Moreover, the fitting is retroactive, meaning that it rejects those values of the ellipticity components \(e_1\) and \(e_2\) that gives a residual higher than 3\(\sigma\) after the fitting, and then it carries out a second fitting without those objects. The rejection is performed by masking the \(e_1\) and \(e_2\) arrays with the package NUMPY to gain in speed. Finally, it writes a text files on the form,

```plaintext
# The fitting is in the form of \(P_n_m=C_n_m L_n(x)L_m(y)\) where
# the maximum degree is 4, meaning \(x(\text{degree})\times y(\text{degree})\leq4\).
# Number of coefficients \(npar=15\) for each \(e_i\)
# Coefficient Matrix for \(e_1\)
# c0_0  c0_1  c0_2  c0_3  c0_4
# c1_0  c1_1  c1_2  c1_3  0
# c2_0  c2_1  c2_2  0  0
# c3_0  c3_1  0  0  0
```
# c4_0 0 0 0 0
#
# Coefficient Matrix for e_2
# c0_0 c0_1 c0_2 c0_3 c0_4
# c1_0 c1_1 c1_2 c1_3 0
# c2_0 c2_1 c2_2 0 0
# c3_0 c3_1 0 0 0
# c4_0 0 0 0 0
#

DLSCOMBINE was designed to accept a fit based on simple polynomial bases such as \( c_{nm}^k(x,y)x_ny_m \). Therefore, a change in its code was performed to make available the possibility of read Legendre polynomials. The changes were introduced in the file fiat.c of the DLS PIPELINE (see below for the new code attached).

```c
/*
 * Calculates \( f(\text{ellip}_\,)=P_{n\,m}=C_{n\,m}\,L_n(x)L_m(y) \)
 * Introduced by Guadalupe Canas in order to use
 * Weak-Lensing Software
 * Version 1.0
 */

/*
 * Function Legendre: generates Legendre polynomials
 * corresponding to the index \( n \) through the iterative formulae
 */

float Legendre(int n, float variable){
    float P[n];
    if (n==0){
        P[0]=1;
        return P[0];
    }
    else if (n==1){
        P[1]=variable;
        return P[1];
    }
    else{
        P[n]=((2.0*n-1.0)*variable*Legendre(n-1,variable)
             -(n-1)*Legendre(n-2,variable))/n;
        return P[n];
    }
```
/*
 * Function fiat_xylookup_legendre: modified from fiat_xylookup.
 * It takes the coefficients of the fitting obtained from
 * fiat_read and calculates the expression
 * \( f(\text{ellip}_-) = P_{n-m} C_{n-m} \text{Ln}(x)\text{Lm}(y) \) for \( e_1 = f(x, y) \) and \( e_2 = f(x, y) \)
 * in terms of a legendre orthonormal basis.
 */

int fiat_xylookup_legendre(float x, float y, float *e1, float *e2, float *p, int order)
{
    int npar, e_1indx, e_2indx, l_xorder, l_yorder;
    // define npar (position of coefficients in text files),
    double legendre_xtmp, legendre_ytmp;

    /* setup */
    npar = (order + 1)*(order + 2);
    *e1 = *e2 = 0.0; // set the ellipticities at 0.0
    e_1indx = 0; // e_1 is at the beginning of the file
    e_2indx = npar/2; // e_2 is at position npar/2 = 15

    for(l_yorder=0, legendre_ytmp=1.0;
        l_yorder <= order;
        l_yorder++, legendre_ytmp*=Legendre(l_yorder, y))
    {
        for(l_xorder=0,legendre_xtmp=1.0;
            l_xorder <= order-l_yorder;
            l_xorder++, legendre_xtmp*=Legendre(l_xorder, y))
        {
            *e1 += p[e_1indx]*legendre_xtmp*legendre_ytmp; // expression for e_1
            *e2 += p[e_2indx]*legendre_xtmp*legendre_ytmp; // expression for e_2
            e_2indx++; // go ahead
            e_1indx++; // go ahead
        }
    }
}
Finally, WL_UTILS.PY writes another text file that makes DLSCOMBINE to know the path to the fitting coefficients text file and the path to the analysed image we need to correct (denominated specfile).

A.5.4 Estimation of Dark Matter Halos

The estimation of DM halos is performed thanks to the software FIATMAP. Therefore, WEAKLENSING software only contain two scripts designed to call FIATMAP properly and read the resulting convergence maps to transform then into signal-to-noise ones. It is extremely important to use a grid or supercomputer in order to carry out this performance.

- **WL_FILTER_MAG_GAL.PY**: PYTHON file. It splits the final galaxies catalogue into three different galaxies sub-catalogues attending to several cuts in magnitude (by default the cuts are \( m_{\text{cut}} = \{20, 22\} \)). It calls the Monte Carlo algorithm implemented in FIATMAP 1000 times obtaining 1000 randomized convergence maps and the real convergence \( \kappa \) map for each sub-catalogue (if the default cuts are maintained, 3000 randomized maps are obtained).

- **WL_MONTECARLOANALIZER.PY**: It reads the 1000 randomized maps corresponding to each galaxies sub-catalogues and calculates the standard deviation \( \sigma \) of the convergence \( \kappa \) value for each pixel by fitting a Gaussian function thanks to the packages SCIPY and NUMPY. Then, it obtains signal-to-noise maps dividing each pixel of the original convergence \( \kappa \) maps by the standard deviation \( \sigma \) associated to each pixel. Finally, it filters signal-to-noise maps for a cut in the signal-to-noise value taking into account a value passed as argument.

A.6 Acknowledges

I acknowledge Santander Supercomputacion support group at the University of Cantabria who provided access to the supercomputer ALTAMIRA Supercomputer at the Institute of Physics of Cantabria (IFCA-CSIC), member of the Spanish Supercomputing Network, for performing simulations/analyses.