EPFL

Development and evaluation of galaxy shape measurement algorithms for radio interferometric data

Vasileios ANGELOPOULOS

■ École
polytechnique
fédérale
de Lausanne

2021

The Cosmos is within us. We are made of star-stuff.
We are a way for the Universe to know itself.

— Carl Sagan

To my family …

Acknowledgements

The path to carry out this Thesis was not the easiest to follow, but it was probably the most exciting that I could imagine taking. Here, I would like to express my gratitude to all individuals who were involved in doing this research.

Profound appreciations are to my supervisor, Professor Jean-Paul Kneib, for challenging my knowledge and directing me in the difficulties of this long journey. He was the guide in my academic exploration, providing stimulation, encouragement, constructive criticism, and advice during my dissertation.

My sincere gratitude goes to the Thesis jury members for the time dedicated to reviewing my work, providing valuable feedback, and contributing to its improvement with their comments.

In addition, I would want to recognize the important role of our MOOC team, Fred, Kim, Oleg, Sphe, and Griffin, in developing the courses for radio astronomy. Without a doubt, it was a fascinating experience and a captivating opportunity to prepare and share knowledge on the area with people all over the world. I will never forget the uncountable hours spent to make the material as better as possible. It was a great pleasure working with you.

I would like to take the opportunity to thank EPFL, LASTRO, and LTS5 colleagues for their contributions, directly and indirectly. I know very well how important was your help in times of straggling and that this work would not be materialized without you.

Finally, my deepest appreciation belongs to my family and my wife for their understanding, consolation, support, and giving strength. I desire to dedicate the success of finishing this study to them. They have had a profound impact on my ability to complete my Thesis.

Lausanne, March 25th, 2021   V. A.

Abstract

The purpose of this Thesis is to develop, test, and characterize different models attempting to tackle the problem of measurement of galaxy shapes applied in interferometric observations. Shape measurement is a tool for estimating the underlying shear due to weak gravitational lensing by large-scale foreground matter distributions. The recent enormous progress in radio interferometric imaging during the last years, with projects such as MeerKAT, ASKAP, and in future SKA, motivates the development of advanced algorithms that utilize radio observations for this task. In most of these algorithms, the main disadvantage is that the proposed models' parameters have to be estimated in advance, assuming certain prior knowledge on the form of the objects. Our motivation is to develop and evaluate frameworks that do not require significant prior information on this form to make accurate measurements.

To achieve this goal, we move in two implementation directions. In the first one, we are based on an existing approach that applies an object decomposition procedure using a dictionary of shapelet functions. Our proposal goes some steps forward trying to estimate this decomposition's key size parameter from a model built on advanced regularization. More specifically, we create an over-redundant dictionary formed as the concatenation of several groups of orthogonal shapelet basis functions with different parameters, and we use structured sparsity penalties to estimate the size parameter of the best group. As an alternative development strategy, we form an algorithm that employs multi-resolution least-squares analysis, which attempts to identify the size value that minimizes the relative residual in the fitting.

The second path, instead of making the shape estimation directly on radio interferometric data, performs image reconstruction from the visibilities and measures the ellipticity of the objects in the resulting images. For this purpose, we adopt a sparsity averaging analysis algorithm that has been developed in the past and restores with high precision the intensity image from the visibilities. We also implement a similar framework that uses the CLEAN algorithm for image reconstruction, which helps compare the results between the two options and evaluate the quality of the measurements achieved.

All our models are tested using a collection of simulated data that include objects

of different profile types, ellipticity, size, orientation, and position in the field of view. The visibilities generation is done using either a simulated Gaussian profile coverage or a realistic SKA-like one. Additionally, we present an initial study of the same algorithms on real objects from the COSMOS survey.

Résumé

Le but de cette Thèse est de développer, tester et caractériser différentes méthodes permettant la mesure des formes de galaxies utilisant des observations en radio interférométrie. La mesure de la forme des galaxies est un outil essentiel pour estimer les déformations produites par l'effet de lentille gravitationnelle faible résultant de la distribution non homogènes de la matière dans l'Univers. Les performances en imagerie radio interférométrique obtenu ces dernières années, avec les observatoires tels que MeerKAT, ASKAP et bientôt avec SKA, motivent le développement de nouveaux algorithmes performants en radio-interférométrie pour la mesure des formes des galaxies. Pour la plupart des algorithmes, la principale difficulté est la connaissance a priori des paramètres des modèles caractérisant la forme des objets. La motivation de cette Thèse est de développer et d'évaluer différentes méthodes qui ne nécessitent pas d'informations préalables sur les paramètres des galaxies pour effectuer des mesures de forme non biasées.

Pour atteindre cet objectif, nous présentons deux méthodologies. Pour la premiere méthode, nous utilisons une approche décomposant les objets à l'aide d'un dictionnaire de fonctions de type 'shapelet'. Tout d'abord nous estimons le paramètre de la taille de la galaxie à partir d'une méthode de régularisation. Plus précisément, nous créons un dictionnaire sur-redondant formé par la concaténation de plusieurs groupes de fonctions de base orthogonales avec différents paramètres et, nous utilisons des pénalités de 'sparsité structurée' pour estimer le paramètre de taille du meilleur groupe. De plus, comme stratégie alternative, nous proposons un algorithme qui utilise une analyse multi-resolution des moindres carrés, mesurant la valeur de taille qui minimise le résidu relatif de l'ajustement.

Pour la deuxième méthode, au lieu de faire de l'estimation de la forme directement sur les données radio interférométriques, nous effectuons tout d'abord une reconstruction d'image à partir des 'visibilités'. Ensuite, nous faisons la mesure de l'ellipticité des objets dans les images reconstruites. Pour cela, nous adoptons un algorithme d'analyse de la moyenne de 'sparsité' qui a été précédemment développée et reconstruit avec une grande précision l'intensité des images à partir des 'visibilités'. Nous avons également développé un algorithme similaire utilisant CLEAN pour la reconstruction d'images permettant de comparer les résultats entre

les deux méthodes, permettant aussi d'évaluer la qualité des mesures obtenues. Tous nos modèles ont été testés à l'aide d'un ensemble de données simulées qui comprennent des objets de différents profil, ellipticité, taille, orientation et position dans le champ de vue. Ensuite, les 'visibilités' sont générées à l'aide d'une couverture d'antennes de profil Gaussien ou de type SKA. Pour finir, nous présentons une étude préliminaire de ces mêmes algorithmes sur les objets réels provenant du sondage COSMOS.

Mots clefs :
lentille gravitationnelle faible, estimation du cisaillement, forme des galaxies, observations interférométriques, décomposition en shapelet, moindres carrés, minimisation avancée, acquisition comprimée, optimisation convexe

Contents

List of Figures

List of Tables

General Note

This Ph.D. Thesis is a continuation of my Master's Thesis pursued in 2016 at EPFL, as a fulfillment of the requirements for the M.Sc. in Communication Systems.

Parts of this work are also included in the following journal article publications (titles and state are given as at the moment this Thesis was written):

- Galaxy shape estimation from radio interferometric data through group sparsity priors - V. Angelopoulos et al. - 2021 - *Monthly Notices of the Royal Astronomical Society journal* (under revision after the 1st review)

- Sparsity averaging analysis for galaxy shape measurement based on radio interferometric observations - V. Angelopoulos et al. - 2021 - *Monthly Notices of the Royal Astronomical Society journal* (in preparation)

Introduction

$\mathcal{T}$ HIS chapter introduces the reader to the scientific field examined in this Thesis and presents the problem background and how we attempt to tackle this in our research. In Section 1.1, we describe the historical background of radio astronomy and why this area is distinctive in studying many astronomical phenomena. Section 1.2 analyses the gravitational lensing effect and how this can reveal information on the Universe that scientists could not reveal in other ways. Section 1.3 makes the connection between the cosmological problems and the galaxy shape measurement that we work on. A short analysis of the related studies performed in this domain is the topic of Section 1.4. Finally, in Section ,1.5 we explain how this Thesis was developed and the contributions through our research in the area, and in Section 1.6, we give an outline of the material included in our study chapter by chapter.

1.1 Radio astronomy and recent progress in the area

1.1.1 The field of radio astronomy

As explained in Swarup, 2015, Karl Jansky was experimenting in Bell Telephone Laboratory (New Jersey, USA), when, in 1933, he accidentally identified that from the direction of the Milky Way there is an emission of radio waves at 19 MHz (Jansky, 1933). Although, due to the bad financial situation at that time, his finding was-not followed at the laboratory. In 1937, this finding was validated by Grote Reber. He used simple wooden and steel pieces to construct a parabolic dish with 9 meter diameter, and the observation was made at 160 MHz (Reber, 1940). In World War II period, strong radio wave emission from the Sun was identified by Hey (Hey, 1946) and Southworth (Southworth, 1945), who worked in independent studies and
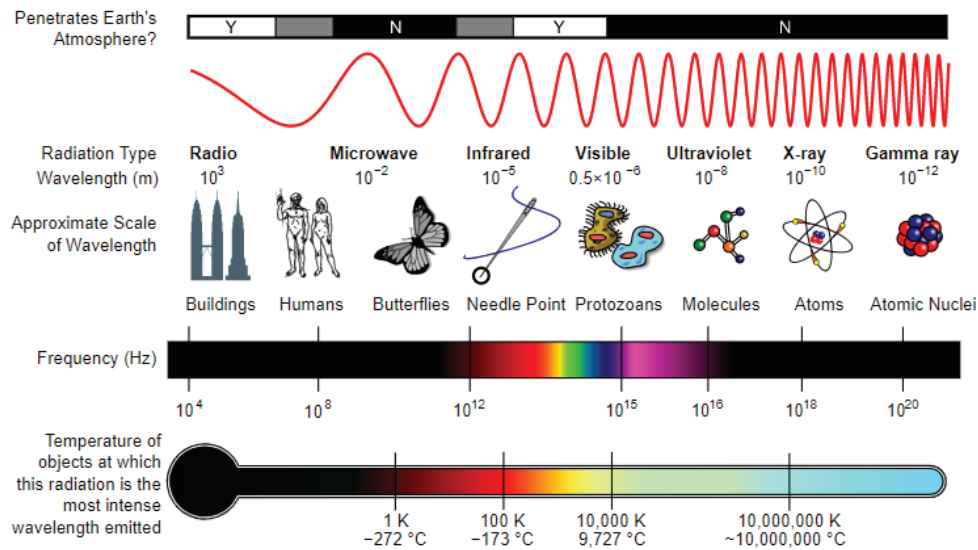
Figure 1.1 – The electromagnetic spectrum and the characteristics of radio radiation analysed in this section. *Credits: NASA and Wikipedia user Inductiveload, distributed under a Creative Commons Attribution-Share Alike 3.0 Unported license and presented without modifications*

waited until the end of the war to make the publication of their discoveries. Nevertheless, due to the huge technological developments during the same period, many researchers from the United Kingdom and Australia got the motivation to construct last-generation radio antennas that would serve this new area in astronomy, radio astronomy. The next decades will be marked by many developments in astronomy that started from Jansky's discovery, such as "the 21 cm hydrogen line, the evolution of distant radio sources, quasars, pulsars, and the cosmic microwave background" (Burke et al., 2019).

A short definition of radio astronomy would be "the study of natural radio emission from celestial sources" (Condon & Ransom, 2016). Radio wavelengths have not strict limits on the electromagnetic spectrum, but their corresponding frequencies can range from 1 MHz to 1 THz. Figure 1.1 presents this spectrum and the approximate scale of the relevant wavelengths compared to other objects that help the reader appreciate the relative sizes. As we can see, the comparative scale for the radio wavelengths is a very tall building.

The observation and study of signals in radio wavelengths have significant advantages in astrophysics. We can observe galaxies in molecular clouds due to the waves' ability in these frequencies to avoid absorption and scattering by dust and gas molecules that are dominant in other types of wavelengths (Burke et al., 2019). A close and well-known radio wave emission source is the Sun at the center of our solar system. Nevertheless, the radio signals coming from other sources may be

hugely powerful, even if they are very distant, with distances from Earth ranging up to several billion light-years. Another example where radio observation can be advantageous is the detection of synchrotron radiation. This type of radiation is also identifiable in other wavebands (such as the X-ray band), but it is a phenomenon better observed in long wavelengths. This gives us the opportunity of studying many of the objects with the highest level of energy in the Universe (Burke et al., 2019).

1.1.2 Aperture synthesis

The study of radio signals is essential in advancing our knowledge of the Universe; however, their detection is a significantly difficult task, especially when dealing with objects with very faint footprints in the radio waveband. The Signal-to-Noise (SNR) ratio needs to be adequately large to achieve good detection due to the relation between angular resolution and dish diameter, and this is not the case with single-dish telescopes. The construction of ever-larger single dish telescopes has the limitation in the strength of the materials.

This problem was solved by the introduction of aperture synthesis in radio astronomy. This technique allows us to observe radio waves from an array of antennas that in total behave as a single virtual telescope, also called an interferometer. In this case, a correlator performs a mixture of the signals received by the dishes to conclude in only one output containing the information of the whole array. In Section 2.1, we describe in more detail how the signals coming from many dishes produce the final observation data collected by the interferometer. Moreover, the angular resolution of the virtual telescope is not more dependent on the diameter of a single dish but on the maximum distance between any two dishes in the array we have created, which is called baseline length. Therefore, we can improve our interferometer's detection capability by increasing the distance between the participating dishes.

In aperture synthesis, the dish pair's output is a complex number, which is a function of the baseline length. This includes the amplitude and phase information of the signal. Assuming that we observe a point source, its flux density will have a proportional relation to the observed signal's amplitude. On the contrary, its phase will be dependent on the difference in length of the paths to each of the dishes (Marr et al., 2019).

In optical observations, there is a relation between the ratio of the observation wavelength and the dish diameter from one side, and the instrument's resolution from the other, which are two proportional measures. Similarly, in radio interferometry, there is proportionality to the ratio of the observation wavelength and the baseline

length. Hence, in an interferometer consisting of many antennas, we get as many responses as the number of antenna pairs in the array. These correspond to data of different orientations or baseline lengths, and the final image of the observed object is constructed by source structure information from many angular scales. Then, we can transform the interferometric data to the source image using a 2D Fourier transform (Marr et al., 2019).

1.1.3   Modern radio telescopes and the SKA

Although a relatively new field of astronomy, radio astronomy has made substantial steps of progress during the past decades. This is combined with the technological progress in building modern radio facilities, as well as new imaging techniques that offer enormous improvement in the quality of the observations.

LOFAR (Low-Frequency Radio Array) is a radio interferometer with many stations in the Netherlands and some more in Germany, UK, Sweden, and France. The stations include antennas that observe in different radio bands, from 30 to 240 MHz. Its frequency resolution is 0.76 kHz, and its maximum baseline length is 1500 km (Hellbourg, 2014). Its role is to study through several surveys a band of the spectrum analyzed the least compared to other bands. The JVLA (Jansky Very Large Array) is based in New Mexico and consists of 27 radio dishes of 25 meters in diameter. They are installed in a Y-like structure, and they provide a sensitivity equivalent to a 130-meter diameter antenna. The ASKAP (Australian SKA Pathfinder) has 36 radio antennas with a diameter of 12 meters, and its mission is to perform fast HI surveys with HDR imaging. The antennas study the field of view of 30 deg^2 at 700-1800 MHz. Furthermore, MeerKAT is a radio interferometer operating since 2018 in Karoo desert, South Africa. It performs at a centimeter wavelength and has 64 antennas with a diameter of 13.5 meters. It is planned to be integrated into the Square Kilometer Array (SKA) telescope and, in particular, the mid-frequency part of it (J. P. Kneib et al., 2020).

Nonetheless, the most challenging project ever designed in this area is the SKA telescope that will bring a revolution in our capabilities of studying and understanding cosmological phenomena and fundamental Physics. Its name originates from the size of its collective area when it will be fully developed. The resolution that scientists will have the opportunity to observe will be below arcsec, utilizing several thousands of telescope components located thousands of kilometers far apart. Its frequency coverage will range between approximately 30 MHz and approximately 10 GHz, with three different configurations. Figure 1.2 gives the reader an artistic view, showing a part of the telescope collective area when this will be finalized. The organization created for this purpose has a core of 16 countries, with 100 or-

Figure 1.2 – A creative view on how SKA will look like when the project will be finalized. This is a bending of photos of already existing real hardware with an impression of the SKA antennas to be built. *Credits: SKA Observatory, distributed under a Creative Commons Attribution 3.0 Unported license and presented without modifications*

ganizations and 20 countries taking part in designing and developing the project. There are two main sites in Australia and South Africa, chosen because of the deficient noise level from human activity, which is very limited in the specified areas (Swarup, 2015). Of course, this telescope's launch will increase the scientific and technological challenges that researchers must face. The load of data produced by the antennas will require new technologies in data storage and management, as well as the algorithms processing this amount of data will demand high-performance computers and techniques.

1.2   Weak gravitational lensing effect

1.2.1   Light deflection by a gravitational field

When a light ray traverses a gravitational field, it is deflected. For many years, scientists hypothesized that this effect exists without being able to confirm it, but only after Einstein formulated the general relativity theory (Bartelmann & Schneider, 2001). One can compute light deflection using Newtonian Physics, where the light ray can be presented as a collection of photons acting as particles. Nonetheless, Einstein's theory measures this effect having double intensity. As mentioned in Eddington, 1920, the validation of the fact that the gravitational lensing effect was double the estimated one according to the traditional theory was the crucial point

Figure 1.3 – A gravitational lensed galaxy, where the deformation causes an almost ideal ring, the so called Einstein's ring. The observation (LRG 3-757) was made by Hubble Space Telescope. *Credits: NASA, ESA, HST*

in accepting Einstein's general relativity theory as the correct one for gravity. An interesting historical analysis for this aspect is provided in Schneider et al., 1992.

Our Sun, like any other astronomical object that can make light deflection, but if the object that causes this effect has a much larger mass and compactness, being also much further than the Sun to us, rays that come from a single light source can be distorted in such a degree that the observer will see multiple images of the same initial object (Bartelmann & Schneider, 2001). In this case, each light ray from the deflection forms an image in their direction, and the observer ends up with multiple copies of the same distant source. Even if general relativity theory had validated the gravitational lensing theory, the first observation of a real case with multiple images was made in 1979 (Walsh et al., 1979). This event was the beginning of the spring in this area, with several teams reporting since then their observations with gravitationally lensed sources and many researchers studying the characteristics of the deflectors.

When the light is deflected due to a gravitational field, there is no emission nor absorption of photons, which results in unaffected surface brightness for the sources experiencing this effect (Bartelmann & Schneider, 2001). Although, the gravitational fields causing the distortion induce changes in the size and shape of these sources. Therefore, each of the multiple images created from an initial distant source will have a different total flux. When the studied source is not a point source but an extended one (e.g., galaxies), this change is not easily recognized since their intrinsic

Figure 1.4 – The central region of this galaxy cluster present system of images from galaxies tangentially aligned regarding the center. This cluster's (Abell 2218) observation was made by Hubble Space Telescope and was first presented in J.-P. Kneib et al., 1996. *Credits: NASA, ESA, HST*

shapes are, in general, non-circular. In Figures 1.3, and 1.4, we can see two examples where we can observe the product of gravitational lensing. When the outcome of the gravitational lensing is a luminous arc, as in Figure 1.4, the caused deformation is so intense that around the strongly lensed objects, it is very likely to have other sources that are experiencing much smaller changes, usually in their shape and size. This weak version of the lensing is not easily identified in a small set of observations. On the other hand, research with a larger data collection can reveal its presence and intensity.

1.2.2 Lens system performance

We consider a mass accumulation (L) in angular diameter distance D_d (which corresponds to redshift z_d) which performs as a lens for the light rays arriving from a source (S) in angular diameter distance D_s (i.e., redshift z_s). We also assume that no other structures in the line of sight may intervene with their gravitational field. The deflecting masses' extent along the line of sight is significantly smaller than D_d
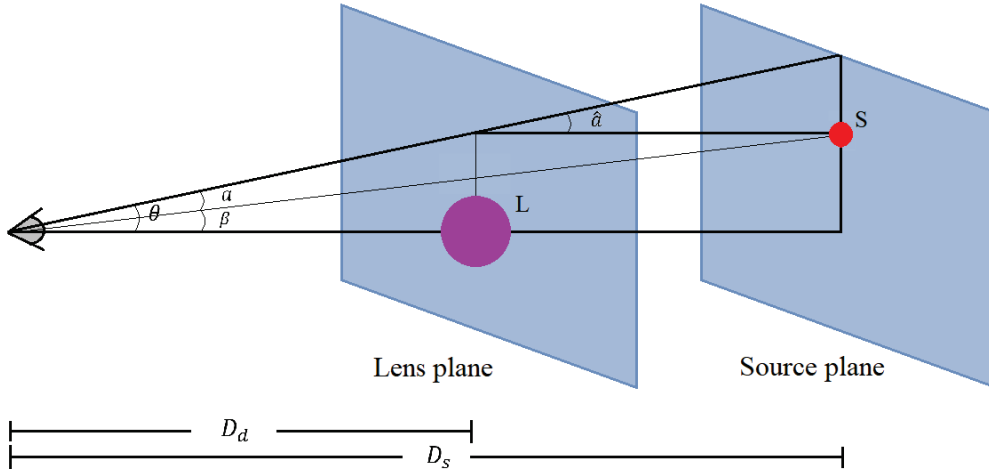
Figure 1.5 – A scheme presenting how a lens-system performs.

and D_s, which is true for most cases in astrophysics. The light rays passing nearby L are then replaced by two straight lines slightly bent close to L. The deflection angle â gives the direction and the intensity of this distortion. This angle is correlated to the impact vector of the light ray, and the mass distribution of L (Bartelmann & Schneider, 2001). Figure 1.5 describes the situation in this case.

Following this description, the source position (β) can be expressed as a function of the position of the observed object image (θ) as:

$$\beta = \theta - \alpha(\theta), \tag{1.1}$$

where $\alpha(\theta)$ is the scaled deflection angle. Additionally, we assume that a point mass M causes the deflection, and there is no light propagation within the strong gravitational field along the horizon (Bartelmann & Schneider, 2001). This implies that the lens' Schwarzschild radius is much smaller than the value of the impact parameter ξ:

$$R_S \equiv 2GMc^{-2} \ll \xi. \tag{1.2}$$

Thus, according to Einstein's general relativity, the deflection angle will be (Schneider et al., 1992):

$$\hat{a} = \frac{4GM}{c^2 \xi} \ll 1. \tag{1.3}$$
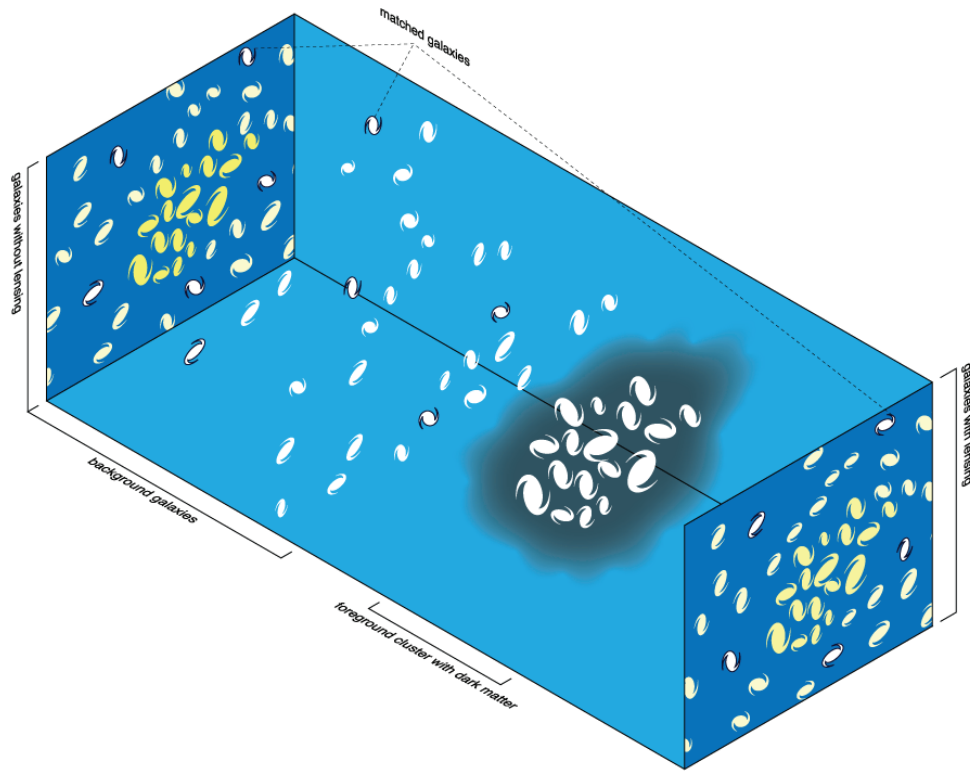
Figure 1.6 – A 3-dimensional depiction of the gravitational lensing effect. *Credits: Created by Michael Sachs, distributed under a Creative Commons Attribution-Share Alike 3.0 Unported license and presented without modifications*

1.2.3 Weak lensing regime

Although the research domain is very interesting and active lately, the weak gravitational lensing, the so-called 'cosmic shear', still has no clear definition, and the analysis of the consequences of this phenomenon is always made statistically (Bartelmann & Schneider, 2001). In order to study the deflector's mass distribution in strong lensing it suffices to have one multiple-imaged source, whereas in weak lensing we need to investigate a collection of sources to explore the results. As presented in Figure 1.4, the shapes of a collection of galaxies have been distorted due to a galaxy cluster with a very large mass in the foreground. The same happens to the size of objects that are affected by an intervening mass accumulation. In Figure 1.6 we show a 3-dimensional representation of the weak lensing effect.

A tool to statistically probe the matter distribution in the Universe is to use the deformations caused by weak lensing. For instance, when we want to study an object with intrinsic luminosity directly deducible by the flux we observe, which in astronomy is called a standard candle (Bartelmann & Schneider, 2001). In this case,

if we achieve to measure its flux we can conclude on the distortions caused along its line of sight. Unfortunately, these effects' intensity is very low, and scientists must tackle big observational issues to study them. Such problems include distortions in the observed objects misleading to gravitational lensing, while they are either systematic (e.g., telescope) or due to propagation effects in Earth's atmosphere. Therefore, researchers have to keep the influence of these parameters to a degree small enough to allow the recognition of the gravitational lensing effect.

Observational cosmology has adopted gravitational lensing to explore solutions to many scientific problems. The motivation behind it is that the matter distribution's gravitational field along a light ray's path helps discover its deflection angle. Additionally, after the determination of this angle, the task becomes a simple geometric problem (Bartelmann & Schneider, 2001).

A powerful application of weak gravitational lensing by large-scale structures is also the measurement of the dark mass distribution in the Universe (Chang et al., 2004; Refregier, 2003b; Van Waerbeke & Mellier, 2003). The gravitational tidal field of the intervening large structure produces this effect by distorting the images of galaxies with a dark matter concentration in the foreground. The distortions induced by this field are correlated since the light rays from nearby background galaxies travel through similar mass inhomogeneities along their way. In order to investigate the mass distribution projected along the line of sight, we have to measure the coherent lensing distortions of the background galaxies. The level of the caused distortion is determined by the mass fluctuations and the mean mass density of the Universe. Furthermore, as a projected effect, its amplitude depends on the geometry of the Universe and the background source distribution in the redshift space.

A square 2×2 tensor field based on two parameters may quantify the lensing distortion. These parameters are the convergence and the directly observable shear. The shear signal must be measured statistically, and the amplitude of this signal under the weak gravitational lensing regime is a small fraction on arcminute scales, making measuring a difficult task. To measure the lensing effect, here, we sample random regions of the sky. Nevertheless, large areas of the sky have to be analyzed, aiming to decrease any possible statistical error.

1.3 Why galaxy shape measurement?

The deflector's gravitational field bends the light rays coming from a set of galaxies in the field of view. The weak signal coming from the galaxies and their considerable distance from us complicates the observational task. The objects' appearance is distorted, with their shape being changed by the gravitational tidal field and their

brightness being altered by the magnification of the gravitational lens (Bartelmann & Schneider, 2001). Assuming that each galaxy has zero intrinsic ellipticity (i.e., a perfect circle), the observation of each object and the measurement of its shape could give us the local distribution of the gravitational field. Although, in reality, and generally speaking, this case is not true. Almost all galaxies have a non-zero intrinsic ellipticity, which does the work of estimating the gravitational field distribution in the majority of the cases much more challenging.

In reality, the intrinsic orientation of any galaxy is assumed to be random (Brainerd et al., 1995; Villumsen et al., 1997). This indicates that one may conclude on the intensity of the tidal gravitational field and the local mass (luminous, baryonic matter, dark matter) distribution by studying a small set of images from neighboring objects. In this case, we must take into account the Poisson noise produced due to the limited number of studied objects. Therefore, their mean intrinsic ellipticity must be larger than the added noise (Bartelmann & Schneider, 2001). When lensing occurs by a deflector, the surface brightness remains constant for the object before and after the lensing. Hence, at a fixed value for surface brightness, the size of the observed galaxy is enhanced by magnification. This increase produces an observed flux that allows the study of objects with intrinsically fainter characteristics. Therefore, the lensing effect changes the local number density of observed galaxies with a flux larger than a specified value.

In the following lines, we explain in a few words how the intrinsic ellipticity of a galaxy is affected because of the weak gravitational lensing effect, following the description in Patel et al., 2010 and Rivi et al., 2016.

When a light ray travels from a distant source to the observer, and it is deflected by the tidal gravitational field that the intervening matter creates, the position of the object θ on the image plane is a mapping of the deflection angle of its position β on the source plane. Extended source objects, like galaxies, present a change in their shape due to the fact that the deflection angle is not constant in all places. One can approximate this angle by the first-order Jacobian matrix that maps the source to the image plane:

$$\mathbf{A} \equiv \frac{\partial \beta}{\partial \theta} = \begin{bmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{bmatrix}^{-1}, \tag{1.4}$$

where κ denotes the convergence and the projected surface mass density, and γ_1 and γ_2 are the real and imaginary shear components giving the change of the object size along the axes $y = 0$ and $y = x$, respectively. Hence, the complex shear reads:

$$\gamma = \gamma_1 + i\gamma_2. \tag{1.5}$$
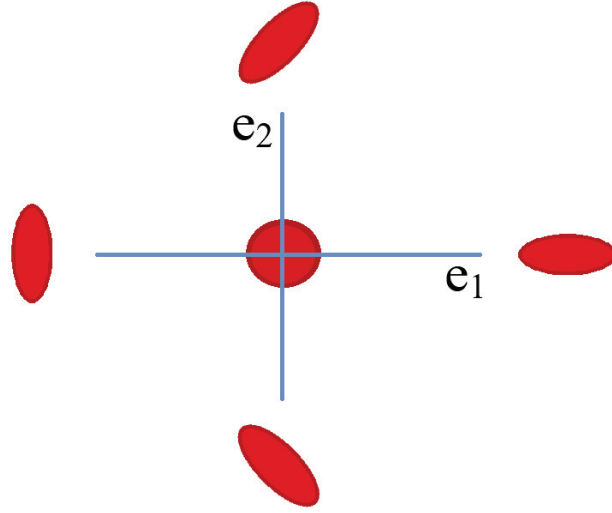
Figure 1.7 – This scheme illustrates the geometric representation of the ellipticity e_i. When the component e_1 is positive or negative, there is a stretching or squeezing, respectively, along axis $y = 0$. Similarly, when the component e_2 is positive or negative, there is a stretching or squeezing, respectively, along axis $x = y$. In case of a circular object, both the ellipticity components are zero.

Figure 1.7 illustrates this geometric representation, which is common for shear and object ellipticity.

Under the weak gravitational lensing regime, the lensing effect is very weak on the galaxies, and the values of convergence and shear parameters are only a few percent. Moreover, since we are interested in measuring the galaxy shape, we may define the 'reduced shear':

$$g = \frac{\gamma}{1 - \kappa},$$ (1.6)

which describes the induced ellipticity of any sheared image. For the weak gravitational case, where we know that $\kappa \ll 1$, it can be simplified in $g \approx \gamma$. The reduced shear develops the same polar transformation properties as shear.

For each of the galaxies, we can define an intrinsic, complex source ellipticity e^s, which comes from the fact that the shape of the source object is not, in general, circular. Then, the ellipticity of the observed object, under the weak gravitational lensing effect, will be:

$$e^{obs} = \frac{e^s + g}{1 + g^* e^s},$$ (1.7)

while the ellipticity and shear are defined as:

$$e^{obs} = \left| e^{obs} \right| e^{i2\theta} \tag{1.8}$$

and

$$\gamma = |\gamma| e^{i2\varphi}, \tag{1.9}$$

where the complex magnitude measures the shape and the phase measures the orientation of the object. Moreover, the complex magnitude of the observed ellipticity reads in an alternative expression:

$$\left| e^{obs} \right| = \frac{a - b}{a + b}, \tag{1.10}$$

with a and b being the major and minor axes of the ellipse, respectively.

In the absence of gravitational lensing, we assume that the galaxies develop random orientations, meaning that $\langle e^s \rangle = 0$. In this case, measuring the observed ellipticity, we get an estimator for the shear since $\langle e^{obs} \rangle = g \approx \gamma$.

1.4   Related studies in the domain

The theoretical background for the weak gravitational lensing by large-scale foreground matter distributions was first set in Gunn, 1967 and was further investigated in modern cosmological researches (Blandford et al., 1991; Jain & Seljak, 1997; Kaiser, 1992, 1998; Miralda-Escude, 1991; Schneider, 1998). However, the phenomenon was clearly identified in the 21st century because of the difficult observational task (D. J. Bacon et al., 2000; Kaiser et al., 2000; Van Waerbeke et al., 2000; Wittman et al., 2000), making many studies to be involved with the investigation of its implications (D. J. Bacon et al., 2003; Jarvis et al., 2003; Refregier et al., 2002).

Since the initial detection of the phenomenon, most of the measurements have been done in the optical band or, alternatively, in the near-infrared one. The large majority of the surveys that have been done are optical, too. The difficulty in interpreting the measurements coming from radio-detected galaxies, since some faint sources have unknown redshift distributions, has made the radio surveys rarer (Rivi et al., 2016).

This issue is attempted to be tackled by the new generation of radio telescopes, which are designed to provide a large number density and resolve the radio emission of ordinary galaxies through sufficient sensitivity. One of these telescopes is the SKA, which at its final implementation phase is expected to reach a number density of up

to ~10 galaxies arcmin^{-2} (Brown et al., 2015). Moreover, the last generation radio surveys have the ability to access the largest scales in the Universe, which, in terms of explored redshifts, is going beyond optical surveys, such as Euclid. Despite that, scientists have to face the issue regarding the systematics of radio observations caused by the large field of view and the higher sensitive instruments. This makes the development of highly accurate and precise shape estimation techniques and algorithms designed for interferometric data necessary.

The techniques currently used are mainly based on the measurement of galaxy shapes in images coming from optical surveys. In Mandelbaum et al., 2015 there is a summary of the methods developed for these surveys. Most of them are based on fitting model surface brightness distributions to the galaxies and make an estimation of the cosmic shear through a combination of the measurements from the distributions.

A major problem that radio observations face is that radio telescopes measure the visibility data from the observed sky instead of directly producing images. These data are formed by a sampled version of the Fourier transform of the sky image, given a set of sampling points *uv* in the Fourier domain. The sampling points come from the projected baseline formed between two antennas on the plane orthogonal to the antennas' pointing direction. In Patel et al., 2014, and Patel et al., 2015, the authors have simulated images of several radio surveys through a radio pipeline, showing that the iterative deconvolution methods currently used give images with analysis-induced bias far from the required one. According to this finding, it is speculated that the source shape measurement would be more precise and accurate if the method avoids reconstructing the intensity image.

In the most recent works that have been done using interferometric data, Patel et al., 2015 provides a study using decomposition with the shapelet basis in the visibility domain, given data from SKA1 simulations, showing the advantages of shapelet models in the frequency domain to describe galaxy image details. Authors in Rivi et al., 2016 present an adaptation to the visibility domain of a model-fitting approach developed for optical surveys, called *lensfit*. Their method models galaxy shapes using an exponential profile and applies a Bayesian marginalization of the likelihood over uninteresting parameters. The estimation of the likelihood is done directly from the visibility data. They noted that additive biases might be better controlled in radio interferometer data than in optical surveys, as any anisotropy of the visibility data may be mitigated in the shape measurement process by a suitable weighting of the measurements.

Researchers in Rivi et al., 2019 propose an extension of the Bayesian Inference fitting model that performs galaxy ellipticity estimation in radio data. The regular Monte

Carlo Markov Chain is not ideal for problems where the parameters' space is very large. For this reason, a Hamiltonian Monte Carlo approach is used to sample the posterior, and the relevant likelihood gradient is computed analytically after the adoption of an exponential galaxy fitting profile. This technique allows sampling that is more accurate and faster. The group has found that this model may also be extended for fitting further galaxy and scientific parameters with marginalization over the effects caused by the system or the instrument.

In one equally recent study, Nammour et al., 2019 work on a solution that tries to tackle the new challenges that arise after the development of modern instruments, with the weak lensing study being possible in new frequencies and the available amount of data being very demanding. For this purpose, they present a radio astronomical image restoration method that adds a shape constraint to the deconvolution algorithms. According to the published findings, the result is a plug-and-play model where the measurements are made with at least 20% improvement compared to techniques without the shape constraint.

1.5   Contributions of this Thesis

In this work, we focus on the problem of unbiased estimation of galaxy shapes applied in radio observations. Working with simulated data, the purpose is to measure the shape of each individual galaxy as well as the accuracy of the measurements. To achieve this goal, we implement several solutions that move in two different main directions.

The first one focuses on the type of fitting we make and how we compute the missing parameters before the measurement. Our approach is the creation of a shape estimation framework based on shapelets (Refregier, 2003a), where galaxy shapes are decomposed through an orthonormal basis of functions corresponding to perturbations around a circular Gaussian. A similar study has been done in the past (Chang & Refregier, 2002; Chang et al., 2004), working on data from the FIRST survey conducted with the VLA. Deploying the shapelets property to be invariant up to rescaling under Fourier transform, the shape measurements were made directly from the visibilities. In our framework, the interferometric data are generated using a realistic set of artificial images and a pre-constructed pattern for the visibilities. In most of the studies mentioned in this chapter that use radio data, the main disadvantage is that the model parameters have to be estimated in advance before the measurement, which assumes certain prior knowledge on the form of the input data. Especially for the approaches using shapelets, one needs to possess very good estimators for the scale and center of the basis functions used for the analysis. The novelty in our proposal is that the framework employs two

alternative techniques to tackle this problem. In one of them, we use group sparse regularization to estimate automatically the scale parameter requiring no prior knowledge on our data, while the other one deals with the same issue applying multi-resolution least-squares fitting. For comparison reasons, we are also testing a baseline approach, where the parameters are computed in advance with good estimators that use information from the intensity images.

The second main direction considers the problem as a two-step task. Again, the visibilities are generated using a pre-defined pattern and an appropriate simulated image database. Initially, this framework attempts to reconstruct the intensity images from the visibilities, while the ellipticity measurements are done on the resulting intensity images. Unlike the fitting-based methods, this approach avoids the time-consuming and often inaccurate parameter estimation procedure for the basis functions, which is necessary for a good ellipticity estimation. On the contrary, in this proposal, the implemented models require almost no prior configuration before making the measurements, and they are suitable to all object types and characteristics. Similarly to the fitting-based solutions, we compare the performance of our algorithms with the measurement of the ellipticity directly in the original optical images using the quadruple moments' technique.

Finally, a significant part of this work was dedicated to developing two Massive Open Online Courses (MOOCs) focused on modern radio astronomy. As explained in more detail in Section 1.1, this scientific area has enormously evolved in the past decades and is expected to progress further during the following period. Many astronomers concentrate their studies in the domain, and researchers from other disciplines apply their expertise in radio astronomy. These facts have got the attention of other people who do not have expertise in this area and want to learn and potentially contribute scientifically. Motivated by this, we have created a series of two courses presenting the theoretical basis of the field, as well as the practical application of it, getting the interest of many participants with different scientific backgrounds, ages, and education levels.

1.6 Thesis outline

The remainder of this Thesis is organized as follows:

- In Chapter 2, we give a brief presentation of the background for the methods used in our work, also describing the theoretical tools used for their development.

- Chapter 3 demonstrates the proposed frameworks of the respective measure-

ment models in detail, showing the algorithmic schemes and the steps followed for the implementation.

- In Chapter 4, we present the experimental results after the application of our models (and their several variants) to our image database, as well as a preliminary study of the same approaches using real data.

- Chapter 5 summarizes the whole research and some ideas on how this could be extended in the future in the same or different directions.

- Chapter 6 is dedicated to introducing the MOOCs developed during my doctoral studies aiming at the popularization of modern radio astronomy.

- Finally, Chapter 7 highlights the main points of the entire work presented in this Thesis.

Scientific Background

*I*N developing the methods described later in this work, we had to apply knowledge from diverse sub-fields of radio astronomy and signal processing. The chapter herein gives a presentation of the scientific theory in these domains and shortly introduces us to the different tools employed in our implementations. More specifically, Section 2.1 shows the procedure to acquire the visibilities through the measurements. In Section 2.2, we include a small description of the compressed sensing theory and why this is important in the problem we attempt to solve. Section 2.3 defines the shapelet basis functions and the parameters required to perform a decomposition using a dictionary of this type. The background of the group sparse regularization procedure is demonstrated in Section 2.5. In Section 2.6, we give the details of an imaging algorithm based on compressed sensing and convex optimization, called SARA, whereas the CLEAN deconvolution procedure is shown in Section 2.7. Finally, Section 2.8 is dedicated to a very popular ellipticity estimation tool, the quadruple moments' technique.

2.1   Radio-interferometric observations

This section describes how the radio interferometric observations are made and the mathematical background of the corresponding measurements. The experiments presented in this Thesis are done on intensity images. Therefore, we follow a procedure for transforming them into radio interferometric data, which will be the input to the implemented models described in Chapter 3. In the following paragraphs, we present briefly how the observations in radio interferometry are formed, following the description in R. Carrillo et al., 2012.

Aiming at the observation of a sky region, we use an interferometric array that

consists of appropriately placed radio telescopes. The field of view is the subject of our study, represented in the Cartesian coordinate system, where the Earth is in the center, and the pointing directions of the dishes are aligned with it. This direction is symbolized with $\hat{p}_0 \in \mathbb{R}^3$ on the unit celestial sphere. A visibility is measured by a telescope pair for an instant observation and represents the correlation of the incoming electric fields at the dishes' positions. These values are dependent only on the baseline that describes the relative positions of the telescopes in the pair. Furthermore, this scheme considers radiation without polarization.

We assume a monochromatic input signal $\mathbf{x}$, consisting of spatially incoherent sources. We also consider a field of view relatively small, aiming to accurately approximate the signal using a plane orthogonal to the pointing direction and the projection of the signal on it. In this context, each visibility point for a planar signal is formed by its Fourier transform at the spatial frequency $\mathbf{k}$. The pair of values $\mathbf{k} = (u, v)$ corresponds to the baseline components in the image plane. Thus, according to the van Cittert-Zernike theorem (A. R. Thompson et al., 2008), we can define the visibility value as:

$$\mathbf{y}(\mathbf{k}) = \int \mathbf{D}(\mathbf{l}) \, \mathbf{x}(\mathbf{l}) \, e^{-2\pi i \mathbf{k} \mathbf{l}} d^2\mathbf{l}, \tag{2.1}$$

where the primary beam is denoted by $\mathbf{D}(\mathbf{l})$, and it is limited to the field of view of our observation. Additionally, $\mathbf{l}$ represents the 2-D coordinates in the image plane. The main characteristic of an interferometer is the total number of visibility points that produces that depends on its coverage $\mathbf{k}$, which gives an incomplete signal representation in the Fourier plane.

The reconstruction of the original signal $\mathbf{x}$ from these incomplete visibilities is a discretized inverse problem variation. Many recent methods focus on the image restoration problem starting from interferometric observations (R. E. Carrillo et al., 2014; R. Carrillo et al., 2012; Onose et al., 2016). The most widely-used algorithm for image reconstruction from radio data is the CLEAN algorithm, which is very effective and is based on non-linear deconvolution with local iterative beam removal (Högbom, 1974; Schwarz, 1978; A. Thompson et al., 2004). Section 2.7 gives a short description of this algorithm, and Section 2.6 presents another imaging algorithm that has been proposed by R. Carrillo et al., 2012 and is used in our model implementation.

This work introduces two types of methods. A group of them does not consider the inverse problem since the implementations make the shape measurement directly on the radio data, ignoring the prior knowledge on the sky brightness information. On the contrary, the other type of methods tries to recover the image intensity data by solving an inverse problem. A more detailed explanation of how these models

work is given in Chapter 3.

2.2   Compressed sensing

In the regular Nyquist-Shannon theory of sampling, when we increase the sampling rate, we can achieve a better representation of the high frequencies of the signal. According to it, the threshold for the sampling frequency is double the maximum frequency of the signal, and that is the minimum sampling frequency that ensures accurate signal acquisition. Although, practically, this task is very demanding because of the limitations of the resources. In the compressed sensing scheme, the signal recovery procedure moves a step forward by introducing an accurate sparse signal representation from incomplete information (Baraniuk, 2007; Candès et al., 2006a; Donoho, 2006; Fornasier M., 2011).

Compressed sensing establishes a non-linear iterative model and regularises the inverse problem that is ill-posed, offering a low-dimensional signal representation. This framework considers that the studied signal is sparse and can be compressed using a suitable decomposition basis. We observe a signal with complex values, expressed by $\mathbf{x} \in \mathbb{C}^N$, and some orthonormal decomposition basis, expressed by $\boldsymbol{\Psi} \in \mathbb{C}^{N \times N}$, where $\mathbf{x}$ is considered sparse. Consequently, among the N decomposition coefficients, only $K : K \ll N$ are non-zero. Additionally, $\mathbf{x}$ can be considered adequately compressible if among the N coefficients the choice of $K : K \ll N$ will not affect the good approximation of it. Then, the signal can be written as:

$$\mathbf{x} = \boldsymbol{\Psi}\boldsymbol{\alpha}, \tag{2.2}$$

where $\boldsymbol{\alpha}$ is the vector of the decomposition coefficients (Onose et al., 2016).

Furthermore, we consider a sensing basis $\boldsymbol{\Phi} \in \mathbb{C}^{M \times N}$ that contributes to the total measurement model, which now will read:

$$\mathbf{y} = \boldsymbol{\Phi}\mathbf{x} + \mathbf{n} \equiv \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{\alpha} + \mathbf{n}, \tag{2.3}$$

where $\boldsymbol{\Phi}\boldsymbol{\Psi} \in \mathbb{C}^{M \times N}$, $\mathbf{y}$ represents the measurement vector, and $\mathbf{n}$, the additive noise. As a general case, we result in an ill-posed inverse problem, because $M < N$ (R. Carrillo et al., 2012).

Therefore, we must solve the following problem:

$$\min_{\hat{\boldsymbol{\alpha}} \in \mathbb{C}^N} \|\hat{\boldsymbol{\alpha}}\|_0, \qquad \text{subject to} \quad \|\mathbf{y} - \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{\alpha}\|_2^2 \leqslant \epsilon, \tag{2.4}$$

where $\hat{\boldsymbol{\alpha}}$ denotes the sparsest coefficients' vector that approximates the measure-

ments. Moreover, ϵ represents the upper bound of $\|\mathbf{y} - \mathbf{\Phi\Psi\alpha}\|_2$, which corresponds to the residual noise. Nevertheless, the aforementioned problem is non-convex and NP-hard. Thus, we must transform it into the closest convex relaxation that will allow the application of convex optimization techniques (Candès et al., 2006a; Donoho, 2006). The new minimization problem will have the form:

$$\min_{\hat{\mathbf{\alpha}} \in \mathbb{C}^N} \|\hat{\mathbf{\alpha}}\|_1, \qquad \text{subject to} \quad \|\mathbf{y} - \mathbf{\Phi\Psi\alpha}\|_2^2 \leqslant \epsilon, \qquad (2.5)$$

which have the same structure with the initial one, but we have substituted the ℓ_0 norm, $\|\hat{\mathbf{\alpha}}\|_0$, with the respective ℓ_1 (R. Carrillo et al., 2012). It has been proved that these two problems can be considered equivalent when the matrix $\mathbf{\Phi\Psi}$ have specific characteristics (Candes, 2008; Candès et al., 2006a; Candès et al., 2006b; R. Carrillo et al., 2012). An alternative expression of the same problem following the Lagrangian formalism will be:

$$\hat{\mathbf{\alpha}}_{L_1} = \operatorname*{argmin}_{\mathbf{\alpha} \in \mathbb{C}^N} \frac{1}{2} \|\mathbf{y} - \mathbf{\Phi\Psi\alpha}\|_2^2 + \lambda \|\mathbf{\alpha}\|_1, \qquad (2.6)$$

where λ is a regularization parameter that has the balancing role between the data fidelity and the regularization terms.

In compressed sensing regime, the sensing basis $\mathbf{\Phi}$ must follow the restricted isometry property, so the signal $\mathbf{x}$ will be efficiently reconstructed from the measurements' vector $\mathbf{y}$ (Fornasier M., 2011), which practically happens when $M > 2K \ll N$. One more condition to be respected is the orthonormality of the decomposition basis $\mathbf{\Psi}$, but studies have shown that over-complete dictionaries frequently result in adequately sparse signals, making the research on these dictionaries widespread recently (Bobin et al., 2007; Gribonval & Nielsen, 2003; Starck et al., 2010). This approach requires the decomposition of the signal $\mathbf{x}$, using an appropriate dictionary $\mathbf{\Psi} \in \mathbb{C}^{N \times D}$, and the corresponding coefficients $\mathbf{\alpha} \in \mathbb{C}^D$, with $D : D > N$. The problem's dimensionality will be increased by this modification from N to D, making it strictly undetermined (R. E. Carrillo et al., 2014).

2.3   Shapelets

The present section briefly describes in mathematical terms how we can linearly decompose objects by constructing a series of localized basis functions of diverse shapes (Refregier, 2003a; Refregier & Bacon, 2003). This basis functions shall be considered as perturbations around the 2-dimensional Gaussian forming the corresponding function for $\mathbf{n} = (0, 0)$, and natural bases for describing the shapes of most astronomical objects. These functions are based on weighted Hermite polynomials,
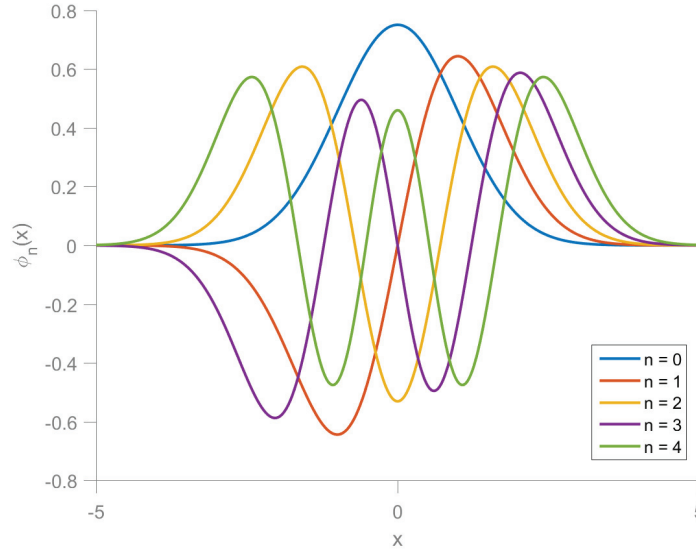
Figure 2.1 – 1-dimensional shapelet basis functions $\phi_n(x)$, for $n = 0, 1, 2, 3, 4$.

and they are eigenstates of the 2-dimensional Quantum Harmonic Oscillator (QHO), which permits us to use the relevant developed theory.

2.3.1 Decomposition functions

In the beginning, we define the 1-dimensional Cartesian function form, and then we will proceed with the extension of the formulation in two dimensions, which is the form used in our implementation. They are defined as orthonormal basis functions with the following formation:

$$\phi_n(x) \equiv \left[2^n \pi^{1/2} n!\right]^{-1/2} H_n(x) e^{-x^2/2}, \tag{2.7}$$

where n represents the shapelet 'order' that is a non-negative integer, and $H_n(x)$ correspond to a Hermite polynomial of order n, the same to the subject basis function. In Figure 2.1 we can see some of these functions with order 0 to 4. To decompose an object practically, we use their dimensional counterpart functions:

$$B_n(x; \beta) \equiv \beta^{-1/2} \phi_n\left(\beta^{-1} x\right), \tag{2.8}$$

where β is the 'characteristic scale of the shapelet function' or simply 'scale', which is a parameter typically translated to the approximation of the size of the studied object.

In this way, any object $o\left(x\right)$ with sufficient and well behaving characteristics, may be decomposed using this infinite set of shapelet functions, which form a complete basis for any integrable smooth function. As a result, we can have the following expansion:

$$o\left(x\right) = \sum_{n=0}^{\infty} f_n B_n\left(x; \beta\right),\tag{2.9}$$

where the decomposition coefficients are denoted with f_n. A well-selected scale parameter approximating the size of the object and a centre of the object close enough to the origin $x = 0$, speeds up the convergence of the resulting expansion series representation.

Shapelets follow a collection of noteworthy properties, with one of them being valuable for our implementation. This is their invariance under the Fourier transform up to a scaling phase factor that reads:

$$\widetilde{B}_n\left(k; \beta\right) = i^n B_n\left(k; \beta^{-1}\right).\tag{2.10}$$

We can use the analogy with the QHO to understand this property better in physical terms. For more information on this analogy, the reader may refer to Refregier, 2003a, containing a full explanation on it.

If we take the tensor product between two 1-dimensional functions, we can move to two dimensions by defining the 2-D shapelet basis functions:

$$\mathbf{B_n}\left(\mathbf{x}; \beta\right) \equiv B_{n_1}\left(x_1; \beta\right) B_{n_2}\left(x_2; \beta\right) = \beta^{-1}\phi_\mathbf{n}\left(\beta^{-1}\mathbf{x}\right),\tag{2.11}$$

with $\mathbf{x} = \left(x_1, x_2\right)$ and $\mathbf{n} = \left(n_1, n_2\right)$.

In this case, the condition of orthonormality is still valid, because:

$$\int d^2x \mathbf{B_n}\left(\mathbf{x}; \beta\right) \mathbf{B_m}\left(\mathbf{x}; \beta\right) = \delta_{n_1 m_1}\delta_{n_2 m_2}.\tag{2.12}$$

Figure 2.2 demonstrates the 2D shapelet functions $\mathbf{B_n}\left(\mathbf{x}; \beta\right)$ for $\mathbf{n} = \left(n_1, n_2\right) \, \forall n_1, n_2 \in [0, 4]$ and scale value equal to 2 pixels. This is the set of the shapelet functions used in a part of our study. As in the 1-dimensional case, the 2-dimensional functions shall be seen as perturbations around the 2-dimensional Gaussian corresponding to the basis function for $\mathbf{n} = \left(0, 0\right)$.

Using this infinite set of basis functions, we can perform the decomposition of any
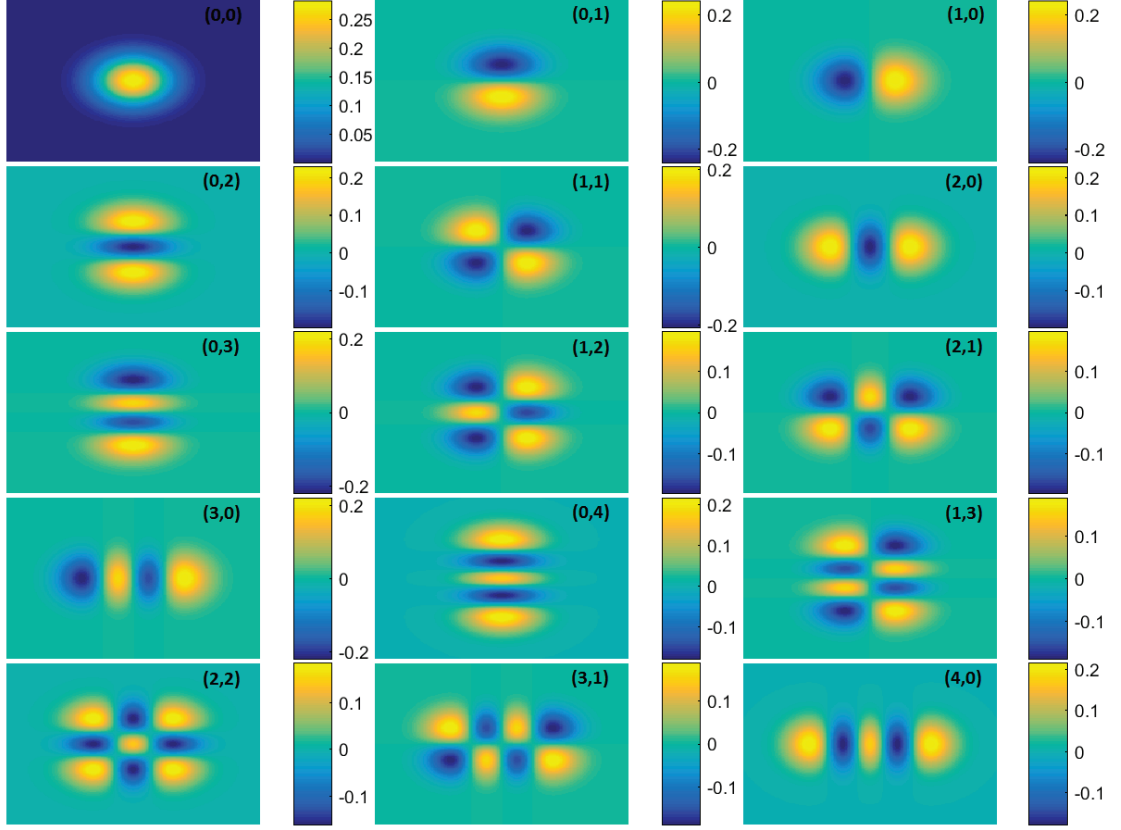
Figure 2.2 – The first few 2-dimensional Cartesian shapelet basis functions $B_{\mathbf{n}}(\mathbf{x}; \beta)$ that are used in our model in their space domain representation, $\forall n_1, n_2 \in [0, 4]$. On the top right corner of each image there is the pair of values (n_1, n_2).

2D object $o(\mathbf{x})$ with sufficient and well-behaving characteristics as:

$$o(\mathbf{x}) = \sum_{n_1, n_2=0}^{\infty} f_{\mathbf{n}} B_{\mathbf{n}}(\mathbf{x}; \beta). \tag{2.13}$$

This basis functions are a natural choice for processing radio interferometric data, due to their invariance under the Fourier transform that is valid also in two dimensions. Therefore, we may decompose the Fourier transform of an object's intensity using:

$$o(\mathbf{k}) = \sum_{n_1, n_2=0}^{\infty} f_{\mathbf{n}} \widetilde{B}_{\mathbf{n}}(\mathbf{k}; \beta), \tag{2.14}$$

with:

$$\widetilde{B}_{\mathbf{n}}(\mathbf{k}; \beta) = i^{n_1 + n_2} B_{\mathbf{n}}(\mathbf{k}; \beta^{-1}). \tag{2.15}$$

In our definitions and equations, we assume that the source centroid is located at $(0,0)$; thus, we do not consider the centroid parameter by not including it in them. Our implemented solutions based on shapelet decomposition extract each object in the preprocessing part from the original image and center it on the origin.

In our models we work with Cartesian basis functions, although the shape measurement requires another type of information that does not accept the direct use of the coefficients of the decomposition with the Cartesian shapelets. Instead, after the decomposition, the resulting coefficients are transformed to the corresponding polar coefficients, as the decomposition was done using the polar basis functions. This transformation is necessary to avoid significantly more complicated computations in the processing of the latter form of shapelet functions. In that case, the following pattern is used:

$$
\begin{aligned}
f_{n,m} = 2^{-n/2} i^m & \left[\frac{n_1! n_2!}{\left(\frac{n+m}{2}\right)! \left(\frac{n-m}{2}\right)!} \right]^{1/2} \delta_{n_1+n_2, n} \times \\
& \sum_{n_r'=0}^{n_r} \sum_{n_l'=0}^{n_l} i^{m'} \binom{\left(\frac{n+m}{2}\right)}{n_r'} \binom{\left(\frac{n-m}{2}\right)}{n_l'} \delta_{n_r'+n_l', n_1} f_{n_1, n_2},
\end{aligned}
\tag{2.16}
$$

with the Cartesian decomposition coefficients being denoted as f_{n_1, n_2}, where $n_1 + n_2 \leqslant n_{max}$. In addition, n is any non-negative integer limited at $n \leqslant n_{max}$, where n_{max} is an integer selected to help truncate the series, and m is an integer with $-n \leqslant m \leqslant n$ in step of two. It is worth mentioning that the states allowed are those with both n and m either even or odd. Finally, (n_r, n_l) is an integer pair such that:

$$
n = n_r + n_l \tag{2.17}
$$

$$
m = n_r - n_l, \tag{2.18}
$$

with $n \leqslant n_{max}$. A further analysis on the topic of polar shapelets and the transformation used is offered in Massey and Refregier, 2005.

2.3.2   Parameters of the decomposition

The decomposition of an object using the shapelets approach requires three parameters to be specified: the characteristic scale, β, the maximum shapelet order that will be used, n_{max}, and the centroid of the object, x_c (or k_c if the analysis is done in the Fourier domain).

For the latter one, our approach is to extract the object from the source image, generating a new image where the object is centered. Therefore, we assume that the

centroid is always located at the origin $(0,0)$. For the other two parameters, there are several approaches in the bibliography on how they can be estimated (Chang et al., 2004; Massey & Refregier, 2005; Refregier, 2003a) where some of them, for radio data, cause many failures due to the incorrect estimation of β (Patel et al., 2010).

Instead of this, for one variant of our model, we choose as β and n_{max} values coming from the estimations proposed in Chang and Refregier, 2002. Accordingly, we define two parameters describing the root-mean square (RMS) radii of the maximum and minimum sized objects that can be analysed by the chosen set of shapelet basis functions, the θ_{max} and θ_{min}, respectively. They can be computed as:

$$\theta_{max} \approx \beta \left(n_{max} + 1 \right), \tag{2.19}$$

$$\theta_{min} \approx \frac{\beta}{n_{max} + 1}. \tag{2.20}$$

Having this in mind, and considering the minimum RMS radius of an object as $\theta_{min} = 1$, the values of the parameters of our model may be approximated by:

$$\beta \approx \sqrt{\theta_{max}}, \tag{2.21}$$

$$n_{max} \approx [\theta_{max} - 1], \tag{2.22}$$

where θ_{max} is the radius of the object that we want to measure and $[\cdot]$ denotes the rounding to the closest integer operator.

As we will see in Section 3.1, the value of n_{max} is not crucially important for our work since a model with constant $n_{max} = 4$ may produce beneficial results, avoiding extending our shapelet function set to higher-order functions. On the other hand, the computed value for n_{max}, following the above-mentioned equations, is essential when we face an image reconstruction problem but less important for the shape measurement task.

2.4   Convex optimization

In radio interferometry, an important aspect is the solution of convex optimization problems. A comprehensive presentation of this type of techniques is given in Bauschke, Combettes, et al., 2011, and in Onose et al., 2016, where we base the analysis in this section.

A way to get scalable algorithms is by adopting proximal splitting methods. All of

these methods concentrate on optimization using the following setup:

$$\min_{\mathbf{z}} g_1(\mathbf{z}) + g_2(\mathbf{z}) + \dots + g_n(\mathbf{z}), \tag{2.23}$$

where g_i, $\forall i \in \{1, 2, \dots, n\}$, are convex functions with lower semi-continuous properties. Every non-differentiable function included in this type of optimization problems has a proximity operator (Moreau, 1965) defined to be:

$$\mathrm{prox}_g(\mathbf{z}) \equiv \underset{\mathbf{z}}{\mathrm{argmin}}\, g(\bar{\mathbf{z}}) + \frac{1}{2}\|\mathbf{z} - \bar{\mathbf{z}}\|_2^2. \tag{2.24}$$

Concerning the constraint minimization problems, they can be expressed using an indicator function:

$$\iota_{\mathcal{C}}(\mathbf{z}) \equiv \begin{cases} 0 & \mathbf{z} \in \mathcal{C}, \\ +\infty & \mathbf{z} \notin \mathcal{C}, \end{cases} \tag{2.25}$$

where the constraints define the convex set $\mathcal{C}$. Generally speaking, proximal splitting algorithms face problems in the form of equation 2.23 by using the proximity operator of a function g_i in an iterative manner. Combettes and Pesquet, 2011 includes a description of the applications of such methods in signal and image processing.

A full splitting can be attained by the primal-dual algorithms that use the proximal splitting philosophy (Komodakis & Pesquet, 2015). A computational improvement in comparison to other models applying splitting techniques can be achieved if the inversion of operators is avoided (Combettes & Pesquet, 2012). In primal-dual approach, all operators used, including the linear ones, can be manipulated separately. In this case, the minimization problem is transformed as:

$$\min_{\mathbf{z}} g_1(\mathbf{z}) + g_2(\mathbf{L}\mathbf{z}), \tag{2.26}$$

which expresses the primal task, with $\mathbf{L}$ being a linear operator. The dual counterpart of this problem would be:

$$\min_{\mathbf{v}} g_1^*\left(-\mathbf{L}^\dagger \mathbf{v}\right) + g_2^*(\mathbf{v}), \tag{2.27}$$

with $\mathbf{L}^\dagger$ being the adjoint of $\mathbf{L}$, and g_i^* being related to g_i with:

$$g^*(\mathbf{v}) \equiv \sup_{\mathbf{z}} \mathbf{z}^\dagger \mathbf{v} - g(\mathbf{z}), \tag{2.28}$$

which defines the Legendre–Fenchel conjugate. With this in mind, we can imply that when the primal problem has a solution, then powerful methods that treat both

the primal and dual tasks can be developed (Combettes & Pesquet, 2012; Condat, 2013; Vũ, 2013).

Another type of algorithms are the Augmented Lagrangian ones, which transform the constrained optimization tasks to equivalent unconstrained. Thus, problems of the form described in equation 2.26 can be transformed with the help of a slack variable to:

$$\min_{\mathbf{z},\mathbf{r}} g_1(\mathbf{z}) + g_2(\mathbf{r}), \qquad \text{subject to} \quad \mathbf{r} = \mathbf{Lz}. \tag{2.29}$$

This is translated into the following optimization problem using the Lagrangian scheme:

$$\max_{\mathbf{s}} \min_{\mathbf{z},\mathbf{r}} g_1(\mathbf{z}) + g_2(\mathbf{r}) + \frac{\mathbf{s}^\dagger}{\mu}(\mathbf{Lz}-\mathbf{r}) + \frac{1}{2\mu}\|\mathbf{Lz}-\mathbf{r}\|_2^2, \tag{2.30}$$

where $\mathbf{s}$, and μ represent the Lagrange multipliers of the problem (Bertsekas, 1982). Methods that implement this approach are the ADMM (Boyd et al., 2011; Yang & Zhang, 2011), and the SDMM (Setzer et al., 2010). In these methods, each step in the minimization procedure alternates between the minimization with respect to $\mathbf{z}$ and $\mathbf{r}$, and the maximization over the Lagrange multipliers $\mathbf{s}$.

2.5   Group sparse regularization

Regularization is a common approach when dealing with ill-posed problems or possibly over-fitting by introducing additional information to the problem. In sparse regularization, we aim to choose the input variables that describe the output in the best way. Hence, we increase the interpretability of the model, which may have input data with dimensionality much higher than the number of available observations. Other reasons for introducing sparsity in a model may be the requirement for compression or reducing the overall computational complexity.

As in Huang et al., 2011 and in Huang, Zhang, et al., 2010 we will form a sparse learning problem considering least-squares regression. A set of p basis vectors $\mathbf{x}_i \in \mathbb{R}^n, \forall i \in \{1, ..., p\}$, with n being the sample size is defined, in order to create a $n \times p$ data matrix, $\mathbf{X}$, such that each of its columns j corresponds to vector $\mathbf{x}_j$. We are interested in estimation of a vector of coefficients, $\tilde{\beta} \in \mathbb{R}^p$, that corresponds to the sparse linear decomposition of a given observations vector $\mathbf{y} = [y_1, ..., y_n] \in \mathbb{R}^n$,

using the matrix $\mathbf{X}$, such that:

$$\mathbf{y} = \mathbf{X}\tilde{\boldsymbol{\beta}} + \boldsymbol{\epsilon} = \sum_{i=1}^{p} \tilde{\beta}_i \mathbf{x}_i + \boldsymbol{\epsilon}, \tag{2.31}$$

where $\boldsymbol{\epsilon} \in \mathbb{R}^n$ is a stochastic noise vector and $\tilde{\boldsymbol{\beta}}$ is assumed to be sparse.

The support of the sparse vector of coefficients is then defined as:

$$\mathrm{supp}\,(\boldsymbol{\beta}) = \{i : \beta_i \neq 0\}. \tag{2.32}$$

A method that could be used as a natural choice for problems involving sparse learning is L_0 regularization, which is formed as:

$$\hat{\boldsymbol{\beta}}_{L_0} = \underset{\boldsymbol{\beta} \in \mathbb{R}^p}{\mathrm{argmin}}\ \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2, \qquad \text{subject to} \quad \|\boldsymbol{\beta}\|_0 \leqslant k, \tag{2.33}$$

with k being the sparsity. Here, we may use the support of $\boldsymbol{\beta}$ and say $\|\boldsymbol{\beta}\|_0 = \left|\mathrm{supp}\left(\hat{\boldsymbol{\beta}}\right)\right|$. Although, practically, we would be called to face a NP-hard optimization problem. Therefore, usually, in similar problems we proceed with the standard convex relaxation of L_0 regularization:

$$\hat{\boldsymbol{\beta}}_{L_1} = \underset{\boldsymbol{\beta} \in \mathbb{R}^p}{\mathrm{argmin}}\ \frac{1}{2} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda \|\boldsymbol{\beta}\|_1, \tag{2.34}$$

which is the respective L_1 problem, also called as LASSO method. Here, the parameter λ is appropriately chosen to achieve the required regularization.

The group structure of the coefficients vector denotes that coefficients that belong to the same group tend to be zero or non-zero as a whole at the same time. In this case, we assume that the set $\{1, ..., p\}$ is partitioned in m disjoint groups such that:

- $\{1, ..., p\} = \bigcup_{i=1}^{m} G_i$,

- $G_1, ... G_m : G_i \bigcap G_j = \varnothing$, when $i \neq j$,

- if $F \subset \{1, ..., p\}$, then for a coefficient vector $\boldsymbol{\beta} \in \mathbb{R}^p$, we define its identical in $F : \boldsymbol{\beta}_F \in \mathbb{R}^{|F|}$.

To take advantage of the group structure, we use an updated version of the method described above, defined as:

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta} \in \mathbb{R}^p}{\mathrm{argmin}}\ \frac{1}{2} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 + \sum_{i=1}^{m} \lambda_i \|\boldsymbol{\beta}_{G_i}\|_2. \tag{2.35}$$

Here, instead of the standard assumption of sparsity, where the number of non-zero coefficients k is used to measure the complexity, we use the group sparsity as an addition to the coefficient one. Thus, k is used for the measurement of sparsity of $\tilde{\beta}$ instead of $\|\tilde{\beta}\|_0$. Then, if the subject signal has been covered efficiently by the groups, the ratio $k \,/\, \|\tilde{\beta}\|_0$ will be small (strong group sparsity). On the other hand, for large ratio (weak group sparsity), this method does not give good results, and $\|\tilde{\beta}\|_0$ should be used for the measurement of the sparsity of the signal. The regularization term must be strong enough to overpower the noise since there is no variation of the method that includes this information in its formulation and can overcome this issue.

The group version of the LASSO method is superior to the standard one when the signal under processing is strongly group-sparse. The group structure and the stability associated with it make the method more robust to noise, whereas the sparse eigenvalue condition can be satisfied with smaller sample size. On the contrary, for weakly group-sparse data and those covered with very small or overlapping groups, the group LASSO method produces inferior quality results. For a more comprehensive analysis of the benefits and learning with group sparsity, as well as for a study on the consistency of the relevant methods, one may refer to Bach, 2008, Huang, Zhang, et al., 2010 and Huang et al., 2011.

2.6   Sparsity Averaging Reweighted Analysis (SARA)

This section is dedicated to presenting an innovative algorithm for radio inter-ferometric imaging, called Sparsity Averaging Reweighted Analysis (SARA). The algorithm was proposed in 2012, and it employs convex optimization techniques under the compressed sensing framework to achieve an accurate image reconstruction. Two main ideas form the novelty of this approach: the average sparsity over a set of wavelet bases and the reweighted ℓ_1 analysis that advances the aforementioned averaging. We will describe this scheme based on the structure and content of the article R. Carrillo et al., 2012, where the method was published.

An important motivation for proposing the sparsity averaging is the observation that different types of sources have sparse representations in different decomposition bases. As an example, for compact or point sources, the best sparse approximation is achieved with a Dirac basis. In contrast, wavelets are suitable for extended continuous structures or piecewise smooth ones with gradient sparsity. However, when we want to analyze a field of view with multiple objects with several structure types expressed in the same image, using a single basis is not the ideal choice. On the contrary, we can consider a set of multiple decomposition bases as an effective prior and the average sparsity over the whole dictionary. As a remark, we need to

mention that the decomposition bases are incoherent between each other, implying that an individual signal can be arbitrary sparse only in one of them (Wiaux et al., 2009).

In this work, the authors made the assumption that a generalized uncertainty relation may be determined by a lower bound on a signal's average sparsity over each pair of bases from the defined dictionary. The consideration practically is that the previously stated uncertainty relation is saturated by the signals of interest. In this context, a long dictionary is formed as a concatenation of many dictionaries containing individual orthonormal bases. This will have the following structure:

$$\boldsymbol{\Psi} = \frac{1}{\sqrt{q}} \left[\boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2, ..., \boldsymbol{\Psi}_q \right], \tag{2.36}$$

where $\boldsymbol{\Psi}_i \in \mathbb{R}^{N \times N}$, with $i = 1, 2, ..., q$, and q is the number of the selected bases in our dictionary. Moreover, $\boldsymbol{\Psi} \in \mathbb{R}^{N \times D}$, and $D = qN$. SARA employs a dictionary of 9 decomposition bases, with the first one being the Dirac basis, ideal for compact or point sources. The remaining bases are the wavelet bases for the first 8 Daubechies wavelets (Daubechies, 1992), which are suitable for sparse representation of piecewise smooth signals. Finally, the wavelet bases are used with a decomposition depth of fourth order.

Considering the reweighted analysis applied in this algorithm, we will begin with explaining the weighted ℓ_1 problem, which can be described as follows:

$$\min_{\hat{\boldsymbol{x}} \in \mathbb{R}^N} \left\| \mathbf{W} \boldsymbol{\Psi}^T \hat{\boldsymbol{x}} \right\|_1, \quad \text{subject to} \quad \|\mathbf{y} - \boldsymbol{\Phi}\hat{\boldsymbol{x}}\|_2 \leqslant \epsilon, \quad \text{and} \quad \hat{\boldsymbol{x}} \geqslant 0. \tag{2.37}$$

In the definition above, we note the extra constraint imposed, specifying the positivity prior on $\mathbf{x}$, and the inclusion of the matrix $\mathbf{W} \in \mathbb{R}^{D \times D}$, which is diagonal and contains the positive weights of the problem. Another point to consider is the noise vector $\mathbf{n}$, which is implicitly related to the bound ϵ in our setting. In the common case where the noise is an independent and identically distributed complex Gaussian one, it can be proved (R. Carrillo et al., 2012) that this noise level estimator's bound must be set at:

$$\epsilon = \sqrt{\left(2M + 4\sqrt{M}\right) \frac{\sigma_n^2}{2}}, \tag{2.38}$$

with $\sigma_n^2 / 2$ denoting the variance of the noise for both its real and imaginary parts. The problem's solution is represented by a function Δ of the input visibilities, the matrices $\boldsymbol{\Phi}$ and $\mathbf{W}$, and the above-computed bound ϵ.

Under the reweighting scheme, a problem similar to the one described above, is

solved in each iteration of the method. The difference is that every next step uses as prior information the solution of the previous step. Therefore, the matrix $\mathbf{W}$ is updated with the inverse of the values of the output of the previous step. The function that is used for this update is:

$$f(\gamma, \mathbf{x}) = \frac{\gamma}{\gamma + |\mathbf{x}|}, \tag{2.39}$$

with γ being a stabilization parameter. When this parameter approaches zero, the weighted ℓ_1 problem approximates the ℓ_0 one. In this context, an iterative method of ℓ_1 problems would be geometrically comparable to an ℓ_0 if the γ value converged to zero through the iterations. For this purpose, SARA solves t weighted ℓ_1 problems, employing a homotopy strategy and producing a decreasing sequence $\left\{ \gamma^{(0)}, \gamma^{(1)}, ..., \gamma^{(t)} \right\}$, until a stationary solution is achieved (Nocedal & Wright, 2006).

The detailed framework is described in Algorithm 1, with the following characteristics:

- $\beta \in (0, 1)$ is a rate parameter and controller of the convergence of γ to zero for $t \to \infty$.

- σ_c is the standard deviation of the noise in the representation space, with value $\sqrt{M \ / \ qN}\sigma_n$ and approximately defines the threshold that below a significant signal component detection is not possible. This is also the lower bound for the values that practically γ can take, considering that in real cases the presence of noise does not allow the sequence $\left\{ \gamma^{(0)}, \gamma^{(1)}, ..., \gamma^{(t)} \right\}$ to converge to zero.

- $\rho, \eta \in (0, 1)$ are the computed relative variation between two successive solutions and the relevant threshold below which convergence is confirmed, respectively.

- The algorithm is initialized with the identity matrix being the weight matrix, the approximate solution before the first iteration being the solution of the first problem, and γ having the value $\sigma_s \left(\mathbf{\Psi}^T \hat{\mathbf{x}}^{(0)} \right)$, with $\sigma_s \left(\cdot \right)$ being the empirical standard deviation of the signal.

- The iterative part stops when either the computed ρ is smaller than the bound we set, or when a certain number (N_{max}) of iterations has been made.

SARA has been a very useful tool that we use in one of the shape measurement

models developed and explained in Section 3.4.

Algorithm 1: SARA Imaging Algorithm

Input: y, Φ, ϵ, σ_c, β, η, N_{max};

Initialize: $t = 1$, $\mathbf{W}^{(0)} = \mathbf{I}$, $\rho = 1$;

Compute: $\hat{x}^{(0)} = \Delta\left(y, \Phi, \mathbf{W}^{(0)}, \epsilon\right)$, $\gamma^{(0)} = \sigma_s\left(\mathbf{\Psi}^T\hat{x}^{(0)}\right)$;

while $\rho > \eta$ and $t < N_{max}$ **do**

$\quad$ Update the weight matrix: $W_{ij}^{(t)} = f\left(\gamma^{(t-1)}, \hat{\alpha}_i^{(t-1)}\right) \delta_{ij}$, for i,j = 1, 2, ..., D,

$\qquad$ with $\hat{\boldsymbol{\alpha}}^{(t-1)} = \mathbf{\Psi}^T\hat{x}^{(t-1)}$;

$\quad$ Compute a solution: $\hat{x}^{(t)} = \Delta\left(y, \Phi, \mathbf{W}^{(t)}, \epsilon\right)$;

$\quad$ Update: $\gamma^{(t)} = \max\left(\beta\gamma^{(t-1)}, \sigma_c\right)$;

$\quad$ Update: $\rho = \left\|\hat{x}^{(t)} - \hat{x}^{(t-1)}\right\|_2 / \left\|\hat{x}^{(t-1)}\right\|_2$;

$\quad$ $t = t + 1$;

end

Return $\hat{x}$

2.7 CLEAN: the basic iterative beam removing technique

A revolution in radio synthesis imaging for sources in astronomy was the invention of the CLEAN deconvolution algorithm. The limited number of antennas in an interferometric array and the partial rotation synthesis of the Earth translate into partial coverage in the Fourier space. CLEAN (Högbom, 1974) is the most popular algorithm since it was initially proposed in 1974 that scientists in the domain use to solve this problem.

Although CLEAN is very efficient in restoring point-like sources from the radio data, it is not ideal for extended sources with more complex profiles. This was the motivation for the development of variants of the algorithm that perform better in these cases. Here, we will focus on the MS-CLEAN version of the method through a short presentation of it adopting the description given in the original publication (Cornwell, 2008). This is also the variant that is applied in our work (see Section 3.5).

Consider a pair of antennas and a baseline $[u, v, w]$. Then, the Fourier component of the sky total brightness, $\mathbf{I}$, is given as:

$$V(u, v) = \int I(x, y) e^{2\pi i(ux+vy)} dx dy, \qquad (2.40)$$

and, in the case of complete coverage, the inverse Fourier transform gives back the sky brightness as:

$$I(x, y) = \int V(u, v) e^{-2\pi i (ux + vy)} du dv. \tag{2.41}$$

In reality, the coverage is always incomplete, and the Fourier space is sampled using a sampling pattern:

$$S(u, v) = \sum_k w_k \delta(u - u_k) \delta(v - v_k), \tag{2.42}$$

and we can obtain the dirty image $\mathbf{I}^D$, by Fourier inversion of the sampled visibility function $\mathbf{V}$, following:

$$\mathbf{I}^D = \mathcal{F}^{-1}[\mathbf{SV}]. \tag{2.43}$$

Moreover, we can define the dirty beam $\mathbf{B}$ that is the inverse Fourier transform of the sampling mask $\mathbf{S}$:

$$B(x, y) = \mathcal{F}^{-1}[S(x, y)] = \sum_k w_k \cos 2\pi(u_k x + v_k y). \tag{2.44}$$

This gives us the dirty image $\mathbf{I}^D = \mathbf{B} * \mathbf{I}$.

The CLEAN method assumes that the sky brightness can be modeled iteratively as a collection of point-like objects:

$$\mathbf{I}_n^C = \sum_{q=1}^{n} \mathbf{I}_q \delta(x - x_q) \delta(y - y_q), \tag{2.45}$$

In each iteration, the values $\mathbf{I}_i$, x_i and y_i are computed using a least-squares fit and the peak is found in the residual image:

$$\mathbf{I}_i^R = \mathbf{I}^D - \mathbf{B} * \mathbf{I}_{i-1}^C. \tag{2.46}$$

In the residual update, we subtract a centered and correctly scaled version of the point-spread function (PSF).

The multi-scale CLEAN (MS-CLEAN) method models the sky brightness as a sum of suitably scaled and centered extended elements, so that:

$$\mathbf{I}_n^M = \sum_{q=1}^{n} \mathbf{I}_q \mathcal{H}(x - x_q, y - y_q, \alpha_q), \tag{2.47}$$

where α_q represents the element width. Consequently, the i-th iteration will try to find a peak in the residual image:

$$\mathbf{I}_i^R = \mathbf{I}^D - \mathbf{B} * \mathbf{I}_{i-1}^M. \tag{2.48}$$

Here, each iteration must determine the values for $\mathbf{I}_i$, x_i, y_i and α_i.

The component shape $\mathcal{H}(x, y, \alpha)$ definition considers that it must have the properties of independence of the pixel orientation and the possibility to use a support constraint. Thus, it is set to be:

$$\mathcal{H}(r, \alpha) = \Psi(r)\left(1 - \frac{r^2}{\alpha^2}\right), \tag{2.49}$$

with Ψ being a prolate spheroidal wave function, which bypasses issues related to the use of a constraint for image plane support or evaluation at a high dynamic range.

The algorithm tries to avoid the exhaustive search over scale values and finds the maximum residual for a set of them. Then, the selected scale size is the one with the maximum adjusted residual. Practically, it applies:

$$S(\alpha) = 1 - 0.6\alpha\,/\,\alpha_{max}, \tag{2.50}$$

with the chosen peak per iteration being decided by the scale size.

Algorithm 2: MS-CLEAN Deconvolution Algorithm

Initialize: $\mathbf{I}^M = 0$, $\mathbf{I}^R = \mathbf{I}^D$, $N = 0$;
foreach α_q **do**
    Compute: $\mathbf{I}_\alpha^R = \mathcal{H}(\alpha) * \mathbf{I}^R$;
    Compute: $S(\alpha) = 1 - 0.6\alpha\,/\,\alpha_{max}$;
end
while $\max\left(\mathbf{I}_\alpha^R\right) \geqslant T$ and $N \leqslant N_{max}$ **do**
    Find strength and location of peak residual for each scale;
    Multiply scales with scale-dependent bias terms;
    Select scale with maximum residual;
    Apply loop gain to the component;
    Add the component to the current model;
    Update the residuals using the precomputed terms;
end
Return $\hat{\mathbf{I}} = \mathbf{B}_G * \mathbf{I}^M + \mathbf{I}^R$

Algorithm 2 presents this scheme in a short step-by-step manner. Here, T is a threshold that we set for the convergence of the method, N_{max} is the maximum number of components that can be identified, B_G is the clean beam, and $\hat{I}$ is the restored sky brightness.

The authors have implemented a version of the method that precomputes several quantities of the procedure to minimize the computational complexity. Furthermore, its convergence is guaranteed when the component shape is selected correctly, and the point-spread functions are computed precisely.

2.8 Quadruple moments

As described in Section 1.3, the galaxies' ellipticity estimation in a set of neighboring objects can reveal information on the tidal gravitational field and the local mass distribution. The study in the domain is commonly done using the KSB method (Kaiser et al., 1994) or variations of it (D. J. Bacon et al., 2000; Van Waerbeke et al., 2000). In these works, the quadruple light distributions in the source plane are mapped to those of the lens plane, and the sheared field is inverted using estimations of the observed ellipticity.

Here, following the notations and the explanation in Goldberg and Natarajan, 2002, we define the galaxy n-th order moments:

$$\langle x^n y^m \rangle = \frac{\int I(x,y)(x-\bar{x})^n (y-\bar{y})^m \, dxdy}{\int I(x,y) \, dxdy}. \tag{2.51}$$

In the equation above, the $\bar{x}$ and $\bar{y}$ represent the corresponding means, and the denominator can be translated to the flux of the object. For quadruple moments, the expression must satisfy $n + m = 2$. Then, the observed ellipticity of a galaxy can be calculated as:

$$e^{obs} = \frac{\langle x^2 \rangle - \langle y^2 \rangle + 2i \langle xy \rangle}{\langle x^2 \rangle + \langle y^2 \rangle + 2\sqrt{\langle x^2 \rangle \langle y^2 \rangle - \langle xy \rangle^2}}. \tag{2.52}$$

This estimator is applied in a subset of our developed shape measurement methods for measuring objects in intensity images. Chapter 3 provides a detailed description of how this is done.

Shape Estimation Algorithms

$\mathcal{U}$SING the background knowledge given in the previous chapters, here, we analyze the shape measurement methods developed and evaluated in this Thesis. Section 3.1 introduces the reader into the decomposition framework using the shapelet basis functions, with a simple model using prior knowledge to make accurate estimations. Sections 3.2 and 3.3, describe two more complex methods for shapelet decomposition that try different ways to maximize the fitting of the data in the decomposition dictionary, aiming at the most precise shape measurement. Section 3.4 gives a demonstration of the method SARA-QMM that uses the SARA algorithm to reconstruct the true sky brightness and measures the ellipticity of the objects in the resulting images. A similar approach is explained in Section 3.5, where the SARA algorithm has been substituted by the MS-CLEAN one. Finally, Section 3.6 describes how we measured our objects directly on the optical data, which is used as a performance measure for the rest of the models.

3.1   Baseline least-squares fitting model (LS-simple)

Here, we introduce a relatively simple model based on shapelet decomposition, which can also be used as a baseline for comparison with the more complex proposed methods described in 3.2 and 3.3. In this model, the preprocessing and decomposition sections are treated individually for every image I_s in the input. On the other hand, the measurement part combines the results of the object's decomposition and the information after analyzing all items in the dataset.

In the beginning, the object is extracted from the input image, creating a new image considered for the next step as the source image. The visibility points $y(\mathbf{k})$ are constructed using the image with the extracted object, following the procedure

presented in Section 2.1. A predefined set that is constant during all our experiments includes several vectors **k** that implement different versions for the coverage points of the radio telescope. More details on this aspect can be found in Section 4.2.

The main processing starts by selecting the model parameters' values. In this method, the parameter values for β and n_{max} are chosen with the help of equations 2.21 and 2.22. We have worked on two sub-models for the assignment of the value of n_{max}, where one of them chooses an individual value for every image, whereas the other fixes it to 4 for all images in the dataset. By the definition of the approximation equations, we need to have some knowledge of the form of our data to proceed with the processing. In this case, it is the RMS radius of the data, and it is coming from the intensity images that we possess. This makes our model dependent on the information that in some way is provided to the algorithm before performing the measurements.

After the estimation of the model parameters, we focus on the decomposition by constructing the dictionary, $\mathbf{B} = \left\{ \overline{B_n} \right\}$, including the evaluation of the shapelet basis functions. The columns of this matrix are shapelet functions of order n, different for each column, which we evaluate at the coverage points contained in the **k** vector. Then, a χ^2 fit is used to represent the source in the field (Chang & Refregier, 2002):

$$\chi^2 = (\mathbf{y} - \mathbf{Bf})^\top (\mathbf{y} - \mathbf{Bf}), \tag{3.1}$$

with vector $\mathbf{f} = \{f_n\}$ containing the respective coefficients for each shapelet order n.

This problem can be transformed into the relevant least-squares one aiming to identify the optimal coefficient values for the decomposition, described as:

$$\mathbf{f} = (\mathbf{B}^\top \mathbf{B})^{-1} \mathbf{B}^\top \mathbf{y}. \tag{3.2}$$

The problem solution represents the optimal coefficient values of the fit using the Cartesian shapelets. However, we need to convert them to those representing the decomposition with polar shapelets for the measurement part. A transformation matrix dependent only on n_{max} is precomputed in our model for this purpose.

For this solution, we define the relative residual that shows the accuracy of fitting our data to the dictionary as:

$$r = \frac{\|\mathbf{y} - \mathbf{Bf}\|_2}{\|\mathbf{y}\|_2}. \tag{3.3}$$

In the last step of the procedure, we define an estimator that gives an unbiased measurement of the ellipticity utilizing the result of the objects' fitting to the shapelet basis functions.
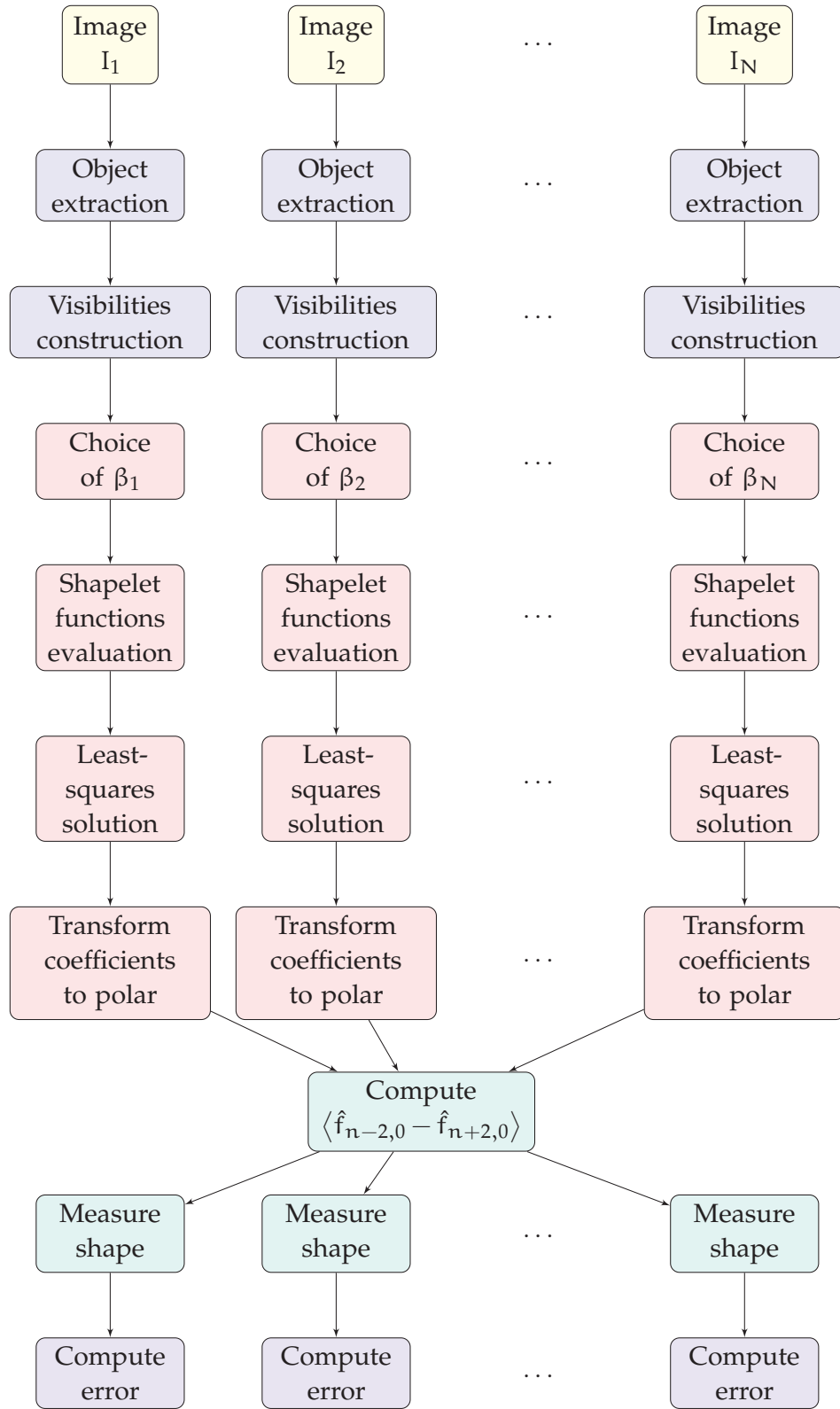
Figure 3.1 – Diagram presenting the shape measurement framework using the LS-simple approach.

Following (Massey et al., 2007; Patel et al., 2015; Refregier & Bacon, 2003), this estimator reads:

$$e^m = \frac{\sqrt{2}\hat{f}_{2,2}}{\langle \hat{f}_{0,0} - \hat{f}_{4,0} \rangle}. \tag{3.4}$$

In this expression, the denominator averages the information from a set of objects and gives a complex number in the output. The quantities $\hat{f}_{n,m}$ symbolize the complex coefficients as a result of the decomposition with the polar shapelets functions, which in our model are acquired from the transformation of the respective Cartesian ones. This estimator is the Gaussian weighted quadruple moment cast into shapelet space, and it has been found to capture most of the shape information of the faint radio sources (Chang et al., 2004). In agreement with this ellipticity estimator, the shapelet functions that are required for the decomposition are those with order up to 4. Therefore, a model with $n_{max} = 4$ has sufficient complexity to make the measurements that we need.

Algorithm 3: LS-simple Algorithm

Input: x, k;
Extract object x_o from x;
Construct y by applying non-uniform FFT on x_o using k;
Compute β by applying equation 2.21;
Compute n_{max} by applying equation 2.22;
Construct Cartesian-polar transformation matrix T;
Construct B columns by evaluating equation 2.15 in k points $\forall$ pair (n_1, n_2);
Solve equation 3.2 using B and y to get f_C;
Apply: $f_p = T(f_C)$;
Compute e^m from equation 3.4 using information from an ensemble of objects
  for the denominator;
Return e^m

The whole procedure is described in Algorithm 3 and demonstrated in Figure 3.1, with blue units denoting the image preprocessing, red units the decomposition steps, green units showing the object's ellipticity measurement, and the purple unit depicting the error estimation. For the rest of the work, the method where the chosen parameters come from the approximation equations 2.21 and 2.22 will be called 'LS-simple'. Similarly, the one with constant n_{max} will be referred as 'LS $n_{max} = 4$'.

3.2 Group sparse regularization measurement model (GLASSO)

Through sparse regularization, we enjoy an effective and flexible scheme to manipulate sparse linear inverse tasks, very much like the ones we face in radio interferometry. Many solutions treat radio data using sparse regularization. In their majority (R. E. Carrillo et al., 2014; R. Carrillo et al., 2012; Onose et al., 2016), they are concentrated in image reconstruction using sparse data. For our purposes, $\mathbf{x} \in \mathbb{C}^N$ will be assumed to be a sparse signal, when, for a given orthonormal basis $\mathbf{M} \in \mathbb{C}^{N \times N}$, its decomposition using this basis gives $L \ll N$ coefficients with non-negligible values, or when we can approximate the signal accurately in the reconstruction using only the L largest of them.

The purpose of employing sparse regularization in this method is, instead of minimizing the number of coefficients used in the signal restoration, to automatically approximate the scale parameter of the shapelet functions without requiring prior knowledge of the data form. The motivation behind the development of this model was the idea that the scale value that provides the optimal fitting also gives the most precise shape estimation. The validity of this idea is explained in detail in Section 4.5.

Our analysis considers several variants, with the common main concept of constructing a long dictionary containing shapelet functions of multiple scale values, which is the outcome of horizontally concatenated dictionaries that represent a single β, similar to the ones of the baseline least-squares approach. Here, the difference is that the constructed dictionary has many more columns compared to this of the baseline approach, where the used scale value is only one and fixed. The problem that we need to solve in this case is:

$$\underset{\mathbf{f}}{\text{argmin}} \ \frac{1}{2} \|\mathbf{y} - \mathbf{Bf}\|_2^2 + \lambda \|\mathbf{f}\|_{2,1} \,, \tag{3.5}$$

which is a group LASSO (hereafter GLASSO) problem. In this context, vector $\mathbf{y}$ includes the visibilities, dictionary $\mathbf{B}$ is the long one constructed as described above that we evaluate at the points in vector $\mathbf{k}$, and vector $\mathbf{f}$ contains the coefficients for estimation. Parameter λ is used for regularization, balancing the contribution of the fidelity and regularization terms. The value for the regularization parameter has no analytical form for its computation, and it depends mainly on the form of the input data. To select it, we use a good estimator, which seems to work well on our data (Garsden et al., 2015), and states:

$$\lambda = \|\mathrm{At}\,(\mathbf{y})\|_\infty \,, \tag{3.6}$$

where $At(\cdot)$ is the operator that implements the multiplication by the transpose of the dictionary matrix.

Unlike other studies that have been done on the problem in the past for other purposes (Meier et al., 2008; Zhang et al., 2010), this approach aims to provide a group sparse result that, in the ideal case it will have only a single coefficient group with members of non-null values. The best shape measurement is expected to be achieved using the β value that corresponds to this coefficient group. If more than a single group has elements with non-zero values, we select the one with the maximum ℓ_2 group norm. After choosing the optimal scale value, similarly to the baseline model, we construct a short dictionary of basis functions using only that value, making χ^2 fit. Then, by solving the resulting least-squares problem, we get the coefficients' estimation required for the shape measurement, according to equation 3.4.

Algorithm 4: GLASSO increasing precision Algorithm

Input: x, k, b, Q, g_{step}, $n_{max} = 4$;

Extract object xo from x;

Construct y by applying non-uniform FFT on xo using k;

Construct Cartesian-polar transformation matrix T;

for $q \leftarrow 1$ **to** Q **do**

    Construct $\mathbf{B}_L$ columns by evaluating equation 2.15 in k points $\forall$ pair
      (n_1, n_2), and $\forall \beta_i \in \mathbf{b}$;

    Compute λ from equation 3.6;

    Solve minimization problem in equation 3.5 using $\mathbf{B}_L$, y, and λ to get $\mathbf{f}_L$;

    Choose from $\mathbf{f}_L$ the group $\mathbf{f}_q$ with maximum ℓ_2 norm;

    Locate β_q that correspond to $\mathbf{f}_q$;

    Update $\mathbf{b}$ with elements around β_q with step size 10^{1-q};

end

$\beta_{opt} = \beta_Q$;

Construct $\mathbf{B}_S$ columns by evaluating equation 2.15 in k points $\forall$ pair (n_1, n_2)
  using β_{opt};

Solve equation 3.2 using $\mathbf{B}_S$ and y to get f_C;

Apply: $\mathbf{f}_p = \mathbf{T}(\mathbf{f}_C)$;

Compute e^m from equation 3.4 using information from an ensemble of objects
  for the denominator;

Return e^m
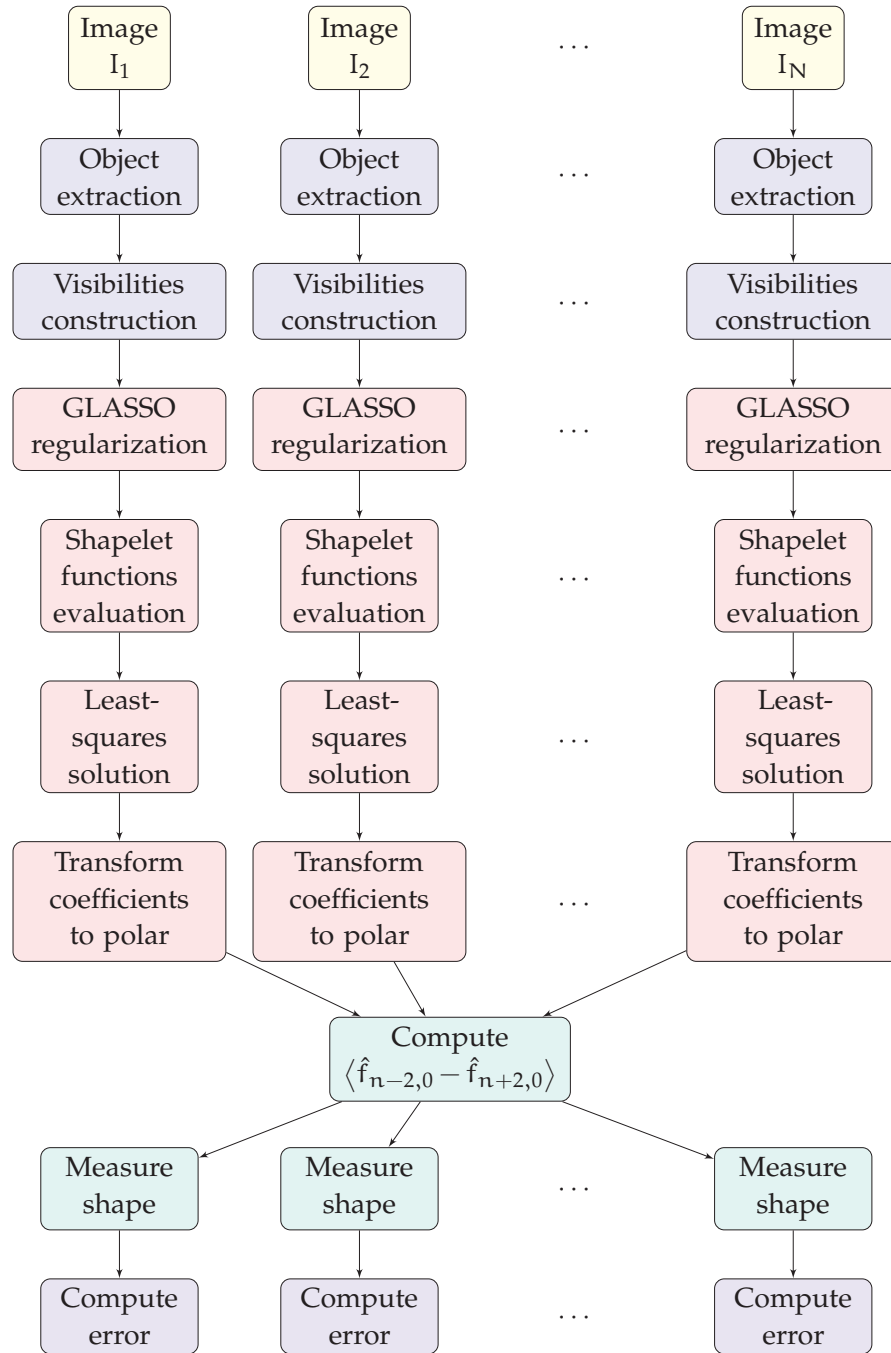
Figure 3.2 – Diagram presenting the shape measurement framework using the GLASSO regularization approach.

We have developed and studied a series of variants of this framework applying group sparsity. In the two sub-models, β is automatically approximated, whereas the rest use a semi-automatic approach, assisted by the information from the approximation equation 2.21. The optimal value is searched in an interval created around

the scale value coming from this equation. Algorithm 4 gives the procedure for the **GLASSO increasing precision** variant presented in subsection 3.2.4. There, **b** is the vector of the chosen β values for our dictionary, Q is the number of iterations of increasing precision, and g_{step} is the gradient step used in the regularization part (further details on this aspect are given in Section 4.5). We can also observe the diagram of the general model in Figure 3.2, where the main modification in the framework is the way that we determine the β value, employing the sparse regularization technique. In the subsections that follow, we outline the algorithm variations considered in our analysis.

3.2.1   Automatic selection using a constant predefined dictionary (GLASSO automatic)

This variant constructs a dictionary containing multiple β values chosen uniformly from a large interval. We consider the range of values given by the baseline model as a good guess, and we observe that the objects in our dataset can be covered with values belonging to the interval $[0.5, 20.5]$. Thus, we select values in this range with a step of 0.1, and we process all images using the common dictionary, getting a β value approximation. Then, a short dictionary is constructed with the selected β value, and a least-squares problem is solved to estimate the fitting coefficients. In the last step, there is the ellipticity measurement and error estimation similarly to **LS-simple**. In this model version, we test if the group sparsity approach can select the optimal scale, assuming that there is no significant prior knowledge on the properties of the studied object.

3.2.2   Semi-automatic selection using the initial guess equations (GLASSO initial guess)

The variant of this subsection offers a similar solution to the one explained above, where the main modification concerns constructing an individual dictionary of multiple scale values per image. It considers as an initial guess the approximation coming from the equation 2.21. Although, it does not use the RMS radius-related value, but instead, it sets up an interval of test values around it. Considering the approximated value from the equation β_0, the interval is defined to be $[\beta_0 - 2, \beta_0 + 2]$, and the step size for the tested values is 0.02. Subsequently, we follow the same steps as the case above. The target of this study is, if possible, to optimize the parameter selection further than the initial guess, leading eventually to an even better measurement.

3.2.3 Baseline LS combined with group sparsity approach (GLASSO mixed with LS)

In our analysis of the baseline model, we saw that for objects with a complex ellipticity magnitude greater than 0.5, the errors in the measurement are usually larger. Motivated by this observation, in this method variant, we initially analyze all objects using the baseline least-squares solution. Then, the cases measured with complex ellipticity magnitude in the first phase greater than 0.5 are re-analyzed with the **GLASSO automatic** that applies group sparsity. Here, the study assumes that the ellipticity estimation for the objects that are not reprocessed is adequately satisfactory because of the smaller errors produced. Thus, the goal is to decrease the total average measurement error by better fitting the initially measured objects with large ellipticity. The threshold set is an empirical one that came after testing several magnitudes in our experiments and is chosen to balance the performance in measurement accuracy and computation time.

3.2.4 Automatic selection with increasing precision iterations (GLASSO increasing precision)

The major change that this version introduces is the increasing precision of the fitting dictionary. An automatic selection of the parameter β, with the required accuracy to produce a minimal measurement error, demands a dictionary of scale values with precision at the magnitude of 10^{-2}. Considering that the values must range, as we mentioned above, from 0.5 to 20.5, the generated dictionary would have a very large number of columns, and the resulting computational complexity for the minimization would be substantial. The novelty of our proposal is the iterative solving of several GLASSO minimization problems, where each time, the dictionary includes scale values with increasing precision. Hence, we begin the processing with a dictionary with precision at the magnitude of 10^0, and at each iteration, we use the solution that we got from the previous one, and we proceed. The new dictionary contains values with one magnitude less than the preceding in the interval around the solution of the previous step. Thus, a β value is assigned, combining the purely automatic choice with the precision that the prior knowledge offers.

3.3 Multi-resolution least-squares fitting model (MR-LS)

In Section 3.2, we presented a method that allows the shape estimation with the minimum knowledge on the data form. However, the construction of a long dic-

tionary containing shapelet functions with different scale values and evaluated at the coverage points $\mathbf{k}$ is an approach that results in a complex regularization problem. This type of problem has increased computational complexity, i.e., increased processing time, and is usually demanding in terms of parameter choice to achieve a successful convergence to the true solution. As we saw in the GLASSO model, we expect convergence to the group of coefficients that correspond to the correct scale value of the problem, while the rest of the coefficients in the vector vanish. Practically, there are many cases where the groups in the coefficients vector are all non-zero. This usually happens when the evaluated shapelet groups have very close values to each other, either because of the object characteristics or the chosen β values in the dictionary. Then, the regularization parameter and the gradient step value must be selected very precisely to perform a successful convergence.

The motivation in developing the approach of this section was to base on the same main principle of minimizing the reconstruction error but build a less complex solution, which guarantees the faster convergence to the scale value that minimizes this error.

In the absence of visibilities, the input to this pipeline is an image I_s, which contains a single galaxy for measurement. After appropriate object extraction, we find the source image for the rest of the algorithm. The processing of a collection of images is done separately at each step until the shape measurement, where the information from the whole dataset is combined to obtain a good estimation. The visibilities are generated from the intensity data, using one of the coverage point collections we have created for this purpose.

In the main part of the method, we apply a grid search using a predefined grid of scale values to find the best β for our problem. Grid search is a traditional way to achieve hyperparameter optimization through a manually stated hyperparameter space's subset in an algorithm. In our case, we have fixed a set of β values that range from 0.5 to 20.5 with a step size 1, which is used in the majority of our experiments. This subset is based on the logic described in 3.2.1. Then, we construct multiple simple shapelet dictionaries $\mathbf{B_i} = \left\{ \overline{B_{i,n}} \right\}$, where $i = 1, 2, ..., D$, with D being the number of the members of the chosen subset (here, $D = 21$), and n the order of the shapelet function. These dictionaries contain evaluations of the basis function of a single scale, and we are called to solve multiple least-squares fitting problems similar to the one in Equation 3.2. Additionally, a point to consider in building a grid search optimization method is the performance metric used to guide it. Here, we want to minimize the relative residual in Equation 3.3, which controls the fitting of

our data to each dictionary. Essentially, the scale value that produces the minimum relative residual in the fitting is the value to be used in the shape measurement.

Algorithm 5: MR-LS Algorithm

Input: x, k, b, Q, $n_{max} = 4$;
Initialize: $r_{opt} = \infty$;
Extract object x_o from x;
Construct y by applying non-uniform FFT on x_o using k;
Construct Cartesian-polar transformation matrix T;
for $q \leftarrow 1$ **to** Q **do**
    foreach $\beta_i \in b$ **do**
        Construct B_i columns by evaluating equation 2.15 in k points $\forall$ pair
          (n_1, n_2), and for β_i;
        Solve equation 3.2 using B_i and y to get f_i;
        Compute r_i from equation 3.3;
        if $r_i < r_{opt}$ **then**
            $r_{opt} = r_i$;
            $\beta_q = \beta_i$;
        end
    end
    Update b with elements around β_q with step size 10^{1-q};
end
$\beta_{opt} = \beta_Q$;
Construct B_S columns by evaluating equation 2.15 in k points $\forall$ pair (n_1, n_2)
  using β_{opt};
Solve equation 3.2 using B_S and y to get f_C;
Apply: $f_p = T(f_C)$;
Compute e^m from equation 3.4 using information from an ensemble of objects
  for the denominator;
Return e^m

Inspired by the solution shown in 3.2.4, we implement a multi-resolution variant of the concept presented above. Hence, when a solution has been found in the first iteration (i.e., the first parameter subset), we proceed with constructing another subset around the solution of the first step, which differs by one order of magnitude. We continue the same procedure until we locate the β that minimizes the relative residual. This iterative scheme can continue as much as we want, but typically in our experiments, we have seen that 3 iterations are adequate to achieve a very good measurement (see Section 4.6).
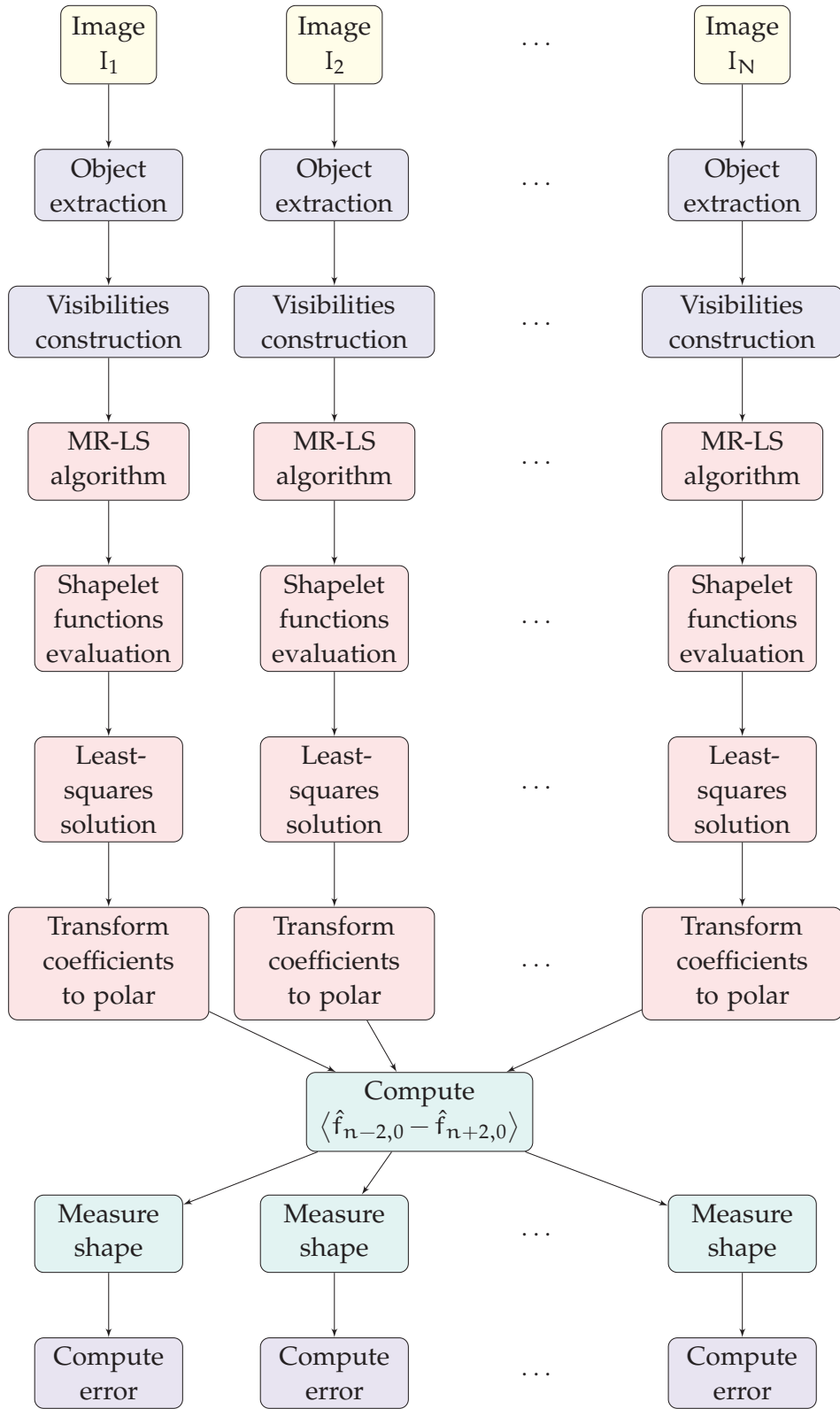
Figure 3.3 – Diagram presenting the shape measurement framework using the MR-LS fitting approach.

Once we have found the scale that best fits our data in the corresponding dictionary and the decomposition coefficients for the Cartesian shapelet analysis, we transform them to the respective polar ones. When the decomposition of the whole database is finished, we use the estimator in Equation 3.4.

One more advantage of this implementation compared to the more complex GLASSO is that, although grid search optimization suffers from the curse of dimensionality, the problem here is formed in such a way that a complete parallelization can be accomplished with each iteration due to the independence in the solution of the individual least-squares problems. Therefore, we can obtain a total computational complexity significantly smaller than the one in the GLASSO model. This approach, hereafter referred to as MR-LS, is described in the scheme of Figure 3.3. Readers can also use Algorithm 5 for a better understanding of the MR-LS procedure. Again, **b** describes the vector of the chosen β values for our grid search optimization, and Q is the number of iterations of increasing precision.

3.4 SARA-based measurement model (SARA-QMM)

Before going into the details of the framework we have developed and considered for our measurement scheme, we will revisit shortly the inverse problem that we are called to solve efficiently to make our measurements, from the point of view that explains the motivation of implementing a solution based on compressed sensing. For this purpose, we follow the presentation given in R. Carrillo et al., 2012, where the relevant setting is explained in more detail.

We start from a map **x** and we have the goal to produce a discretized image from it. For this reason, we construct a discrete, square, and uniform grid with a total of N points in real space, consisting of $\sqrt{N}$ points in each dimension. Additionally, we assign the respective discrete spatial frequencies, u_i, $i = 1, ..., N$, to each grid point. Then, this grid can be used to completely identify by Nyquist-Shannon sampling the original signal multiplied by the primary beam. We also consider that the observation spacial frequencies **u** are part of the aforementioned grid.

In this context, we can define the following ill-posed inverse problem that attempts to reconstruct the signal **x**:

$$\mathbf{y} = \mathbf{LFAx} + \mathbf{n}, \tag{3.7}$$

with the following characteristics:

- $\mathbf{y} \in \mathbb{C}^M$ is a complex vector $\mathbf{y} = [y_1, y_2, ..., y_M]^T$ containing the radio interferometric data coming from the observation measurements at M spatial

frequencies, such that $y_j \equiv y\left(u_j\right), j = 1, 2, ..., M,$

- $\mathbf{L} \in \mathbb{R}^{M \times N}$ is a binary mask that characterizes the interferometer and is true only for the spacial frequencies probed during the observation,

- $\mathbf{F} \in \mathbb{C}^{N \times N}$ contains the respective Fourier coefficients that perform the Discrete Fourier Transform (DFT),

- $\mathbf{A} \in \mathbb{R}^{N \times N}$ is a diagonal matrix that corresponds to the primary beam pattern, and

- $\mathbf{n} \in \mathbb{C}^{M}$ is a vector containing the additive complex noise induced in our visibilities due to systematic or other factors.

Here, the matrix $\mathbf{LFA}$ describes the full linear relation of the signal and the radio interferometric data. Considering that the signal for recovery contains only real values, the measurements could be acquired initially only from the half of the plane and compute the rest using conjugate relations. Regarding the Fourier transform, a complete coverage of the plane would correspond to $N/2$ visibility points. However, we have an incomplete coverage for the discretized scheme, with $2M < N$ real constraints. Finally, if we want to compute the dirty beam using the inverse transform of the mask $\mathbf{L}$, we must use $\mathbf{F}^{\mathsf{T}}\mathbf{L}^{\mathsf{T}}\mathbf{1}_M$, with $\mathbf{1}_M$ being a real vector of ones of size M. Similarly, to obtain the dirty image, we must apply the operator $\mathbf{F}^{\mathsf{T}}\mathbf{L}^{\mathsf{T}}\mathbf{y}$ that corresponds to the inverse transform of the measurements, where we have set to zero all visibilities that have not been observed.

The above-mentioned problem has an infinite number of solutions, and we must apply a regularization method to provide adequate prior information that will allow the estimation of a unique solution.

This section proposes a novel method for galaxy shape measurement, which enhances the already state-of-the-art techniques with an image reconstruction algorithm based on convex optimization and compressed sensing. We decided to use the SARA algorithm to tackle the complex problem described above, which is part of the shape measurement for the estimation models working on the intensity data after image reconstruction.

There were two main motivations in developing this measurement method. The first one, similarly to the previously presented solutions, was implementing a framework that can make accurate and precise estimations without significant prior knowledge of the data. The second was to promote the measurement on the reconstructed intensity images as an alternative, comparing its performance with the methods making the procedure directly on the interferometric data. In such a way, we also propose another option to handle problems with radio interferometric data, similar

to the galaxy shape measurement. For these reasons, we adopted an algorithm that has shown excellent efficiency in radio interferometric imaging. Readers may refer to the original article R. Carrillo et al., 2012 for more details on this aspect.

Algorithmically, this method (hereafter referred to as SARA-QMM) differs from the ones given in the previous sections in the way we achieve the result. Firstly, the pipeline's input is not mandatory to be an image with a single object for measurement. A case with only one galaxy in the input image can be processed, but the advantage of the algorithm is that the whole field-of-view can be processed simultaneously. In principle, the framework works directly with the radio data. However, when we do not have the visibilities available and want to test the method's performance, we need to generate them. Therefore, similarly to the other models, we perform a visibilities construction step, where we use one of the predefined coverage patterns.

In the next step, we use the visibilities as the input to the imaging part of our algorithm based on SARA. Practically, we try to solve an inverse problem similar to the one described above. In our implementation, we have experimented with the specific parameters of SARA, and we found that the efficiency of the model is maximized when we work with a dictionary containing the Dirac basis and the first 8 Daubechies wavelets, with a total of 9 decomposition bases used, and a fourth-order decomposition depth. Moreover, for the rate parameter, we found that a good value is 10^{-1}, while the threshold for the relative variation between two successive solutions, below which we consider successful convergence, is fixed at 10^{-6}. For further information on how the reconstruction with SARA works, one may refer to Section 2.6.

When we have achieved a convergence on the image reconstruction step, we proceed with processing the resulting intensity image. Here, we want to isolate the several objects presented in the field-of-view in separate sub-images. After finding the maximum value in the image, we compute a binary image by thresholding the initial one using an appropriate threshold value, T. This value is determined according to the data we want to process, but essentially suitable values are those between 10% and 20% of the maximum value, depending on how strictly we consider the object's borders. In shape measurement, a small difference at the chosen threshold does not significantly affect the method's efficiency. At this stage, thresholding the image helps in removing the background noise.

The next level in this procedure is the 'hole fill', which assists in eliminating any background pixels inside the objects. This is particularly useful when we work with astronomical images where the included objects are not only galaxies but stars and other types of structures. In this case, another preprocessing step is required to

exclude pixels of very high intensity level that 'hide' any other object. Therefore, some background pixels may end up in the bodies of galaxies. Another possible reason to perform this task is to cover any noise that has been induced in our objects.

In the next phase, we employ the MATLAB built-in function `regionprops` to measure the properties of the objects we are interested in, and more specifically, their centroids and their bounding box, which will help us to identify them. Then, we apply a real object recognition function that we have developed to exclude possible artifacts from the final list.

Algorithm 6: SARA-QMM Algorithm

Input: x, k;
Initialize: $\beta = 10^{-1}$, $\eta = 10^{-6}$, $N_{max} = 1000$;
Construct y by applying non-uniform FFT on x using k;
Construct Φ;
Compute: ϵ, σ_c;
Apply: Algorithm 1 with input y, Φ, ϵ, σ_c, β, η, N_{max}, to get $\hat{x}$;
Compute: $T = 0.15 \max (\hat{x})$;
Compute binary image from $\hat{x}$ using T, to get $\hat{x}_b$;
Perform hole fill to $\hat{x}_b$;
Isolate objects using `regionprops` function;
Apply real object recognition procedure and produce objects $\hat{x}_i$ $\forall i \in \{1, \dots, K\}$;
foreach $\hat{x}_i$ **do**
   | Compute $\langle x^2 \rangle$, $\langle y^2 \rangle$, and $\langle xy \rangle$ from equation 2.51;
   | Compute e_i^m from equation 2.52;
end
Return e_i^m $\forall i \in \{1, \dots, K\}$

In the final point of the procedure, we measure the shapes of the objects we are interested in sequentially by separating each sub-image from the previous step. For this purpose, we work with the quadruple moment technique, analyzed in Section 2.8, which is very popular for similar estimation problems and produces very accurate results. We also use the object properties that we collected in the previous stages to identify the object in the image and compare the shape estimation with the ground truth when this is available. Figure 3.4 depicts in a compact way the whole procedure schematically. Algorithm 6 presents the main steps of the method, where K is the number of objects identified in the input image.
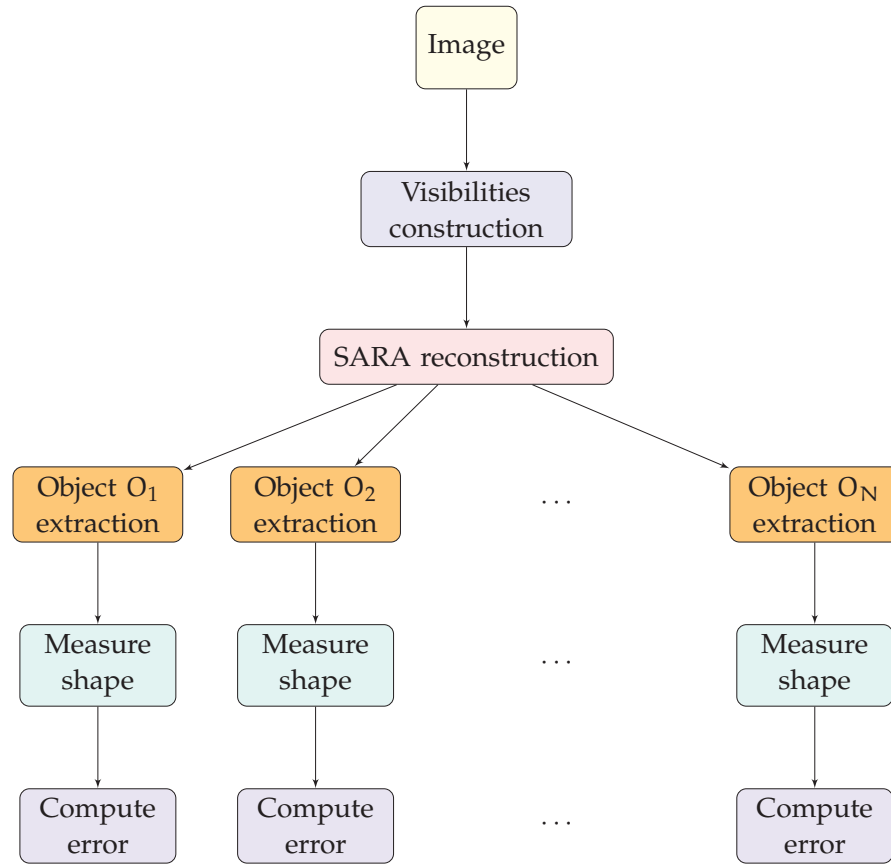
Figure 3.4 – Diagram presenting the shape measurement framework using the SARA-QMM approach.

3.5 CLEAN-based measurement model (CLEAN-QMM)

This approach could be considered a variation of the one described in Section 3.4 since they have similarities in many of their steps. The idea was to create a model that applies the MS-CLEAN algorithm for image reconstruction instead of the more modern SARA. MS-CLEAN is a state-of-the-art deconvolution algorithm that is well appreciated among researchers to find a true sky brightness estimation based on incomplete visibility data. Thus, we can evaluate if and how much the performance of the whole framework improves or declines because of the modification in the algorithm for image reconstruction. It is a clear way to learn if a more complex algorithm, like SARA, is necessary to achieve a better estimate in the shape measurement problem.

As in the other methods that we have developed, the input of this pipeline is the collection of visibilities from our observation. If the input data is not acquired by

radio interferometry, the optical data must be transformed into visibilities using a telescope coverage. For our experiments, we use the same set of coverage collections that help the direct comparison among the models. The CLEAN-based measurement (hereafter CLEAN-QMM), again, is constructed in such a way that it can accept the whole field-of-view as an input. Thus, the measurement can be operated on multiple objects in the same execution.

Then, the radio interferometric data are available, and we load them to the MS-CLEAN deconvolution part. This is the main difference with the SARA-QMM approach. Here, we expect to receive a deconvolved intensity image where the objects can be easily identified, and their shape information will be intact. This will allow an accurate shape estimation in the next steps. Section 2.7 gives a more detailed description of CLEAN and how the deconvolution problem is solved.

Algorithm 7: CLEAN-QMM Algorithm

Input: x, k;
Initialize: N_{max} = 30000;
Construct y by applying non-uniform FFT on x using k;
Compute I^D by applying inverse non-uniform FFT on y using k;
Compute: P;
Apply: Algorithm 2 with input I^D, P, N_{max}, to get $\hat{x}$;
Compute: $T = 0.15 \max(\hat{x})$;
Compute binary image from $\hat{x}$ using T, to get $\hat{x}_b$;
Perform hole fill to $\hat{x}_b$;
Isolate objects using `regionprops` function;
Apply real object recognition procedure and produce objects $\hat{x}_i \; \forall i \in \{1, \ldots, K\}$;
foreach $\hat{x}_i$ **do**
> Compute $\langle x^2 \rangle$, $\langle y^2 \rangle$, and $\langle xy \rangle$ from equation 2.51;
> Compute e_i^m from equation 2.52;

end
Return $e_i^m \; \forall i \in \{1, \ldots, K\}$

The following two phases, including the object extraction into sub-images and the shape measurement working with the quadruple moment technique, are the same as the respective phases of the model described in Section 3.4. Algorithm 7 shows the procedure including the most important steps, with P being the PSF computed using the inverse non-uniform FFT on the non-uniform FFT of a 2-dimensional Dirac function with the help of the coverage k. Figure 3.5 also gives a diagrammatic view of the framework presented in this section.
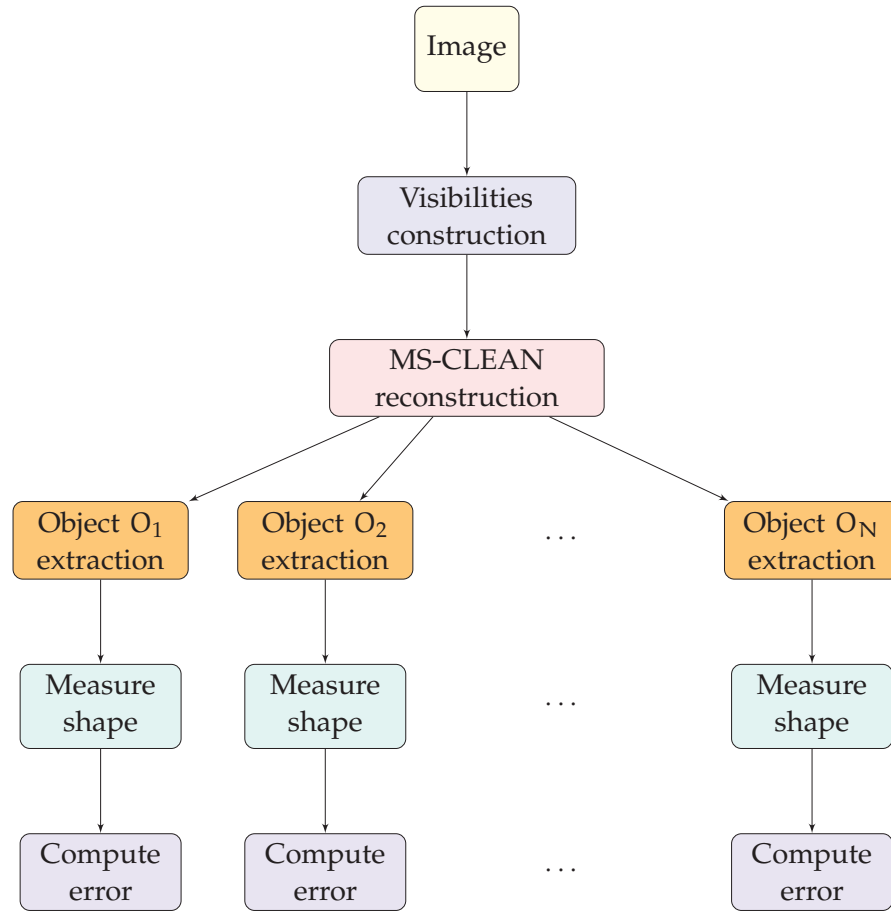
Figure 3.5 – Diagram presenting the shape measurement framework using the CLEAN-QMM approach.

3.6 Measurement directly on optical data

This model represents a reference measurement approach, with the main difference that it performs the required estimation directly on the optical data without considering the visibilities. When we developed the solutions presented in the previous sections, particularly those using the quadruple moments' technique for the measurement, we considered that their evaluation must include all their framework parts. In this way, we could demonstrate the increase in performance compared to all state-of-the-art tools for this problem.

In Section 3.5, we investigate the difference in the model's performance in Section 3.4 when we substitute the image reconstruction part with the widely popular MS-CLEAN deconvolution algorithm. This section is dedicated to providing an evaluation of the efficiency of the quadruple moment measurement technique in our dataset. Thus, we can conclude on the share of the total measurement error that
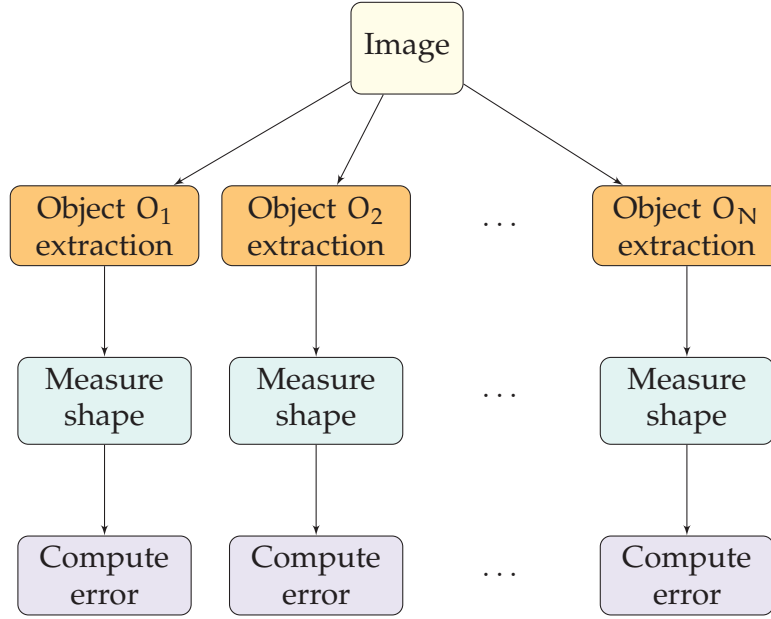
Figure 3.6 – Diagram presenting the shape measurement framework when the measurement is made directly on the optical data.

corresponds to SARA reconstruction, and the part that is a result of the quadruple moment technique.

Algorithm 8: Direct Optical Measurement Algorithm

Input: x;
Compute: $T = 0.15 \max(x)$;
Compute binary image from x using T, to get x_b;
Perform hole fill to x_b;
Isolate objects using `regionprops` function;
Apply real object recognition procedure and produce objects $x_i \ \forall i \in \{1, \ldots, K\}$;
foreach x_i **do**
  Compute $\langle x^2 \rangle$, $\langle y^2 \rangle$, and $\langle xy \rangle$ from equation 2.51;
  Compute e_i^m from equation 2.52;
end
Return $e_i^m \ \forall i \in \{1, \ldots, K\}$

An additional reason that the current model is useful is as potential ground truth for data without known true shape information. More specifically, when we do not have available such information, we can use direct measurements on the optical

data to guide the evaluation of the main measurement solutions of this work. For more details on using this measurement type as a performance measure, one may see Section 4.10.

Figure 3.6 gives a clear depiction of how this model works, while the reader can use Algorithm 8 for a clear presentation of the steps followed. In more detail, each image includes many objects and is processed directly by extracting them in many sub-images that contain a single galaxy. The algorithm for object extraction is the same as the one used for the methods in Sections 3.4 and 3.5. Then, the measurement is done employing the quadruple moments computed for each input sub-image. Finally, the evaluation is done using the ground truth information when this is applicable.

4

Experimental Evaluation

$\mathcal{E}$ ARLIER we gave the algorithmic description of the implemented shape measurement solutions. After introducing the background of the experiments, this chapter demonstrates the results of their evaluation. Section 4.1 gives a presentation of the database of simulated objects we generate to help us test the performance of our models. In Section 4.2, we give the background details for the experiments and the parameters' selection for the intensity image models. We use Section 4.3 to describe the quantities we have set up to evaluate the experimental results. Sections 4.4 to 4.9 show and comment the results from each model when applied to the simulated data, in terms of errors, biases, and resistance to noise. Additionally, we test our implementations using a real dataset, and Section 4.10 describes the outputs from this type of experiments. Finally, Section 4.11 discusses the total behavior of the studied approaches and the advantages and disadvantages of each one.

4.1   Simulation database description

The main part of the experimental evaluation for our implementations was done on synthetic images. Real data in the domain are limited and the experiments demand a large collection of these. Therefore, a database of synthetic images with very similar properties to real objects, was considered a very good alternative.

We have generated a large dataset of about 15000 images grouped in smaller sets of similar properties. More specifically, a set of 10000 images corresponds to items of size 201×201 with a single object in them with Gaussian profile. The galaxies in

these images are 2-dimensional Gaussian formations such that:

$$f(x,y) = A e^{-\left(a(x-x_0)^2 - 2b(x-x_0)(y-y_0) + c(y-y_0)^2\right)}, \tag{4.1}$$

with:

$$a = \frac{\cos^2\theta}{2\sigma_x^2} + \frac{\sin^2\theta}{2\sigma_y^2} \tag{4.2}$$

$$b = \frac{\sin 2\theta}{4\sigma_x^2} - \frac{\sin 2\theta}{4\sigma_y^2} \tag{4.3}$$

$$c = \frac{\sin^2\theta}{2\sigma_x^2} + \frac{\cos^2\theta}{2\sigma_y^2}, \tag{4.4}$$

where θ is the clockwise angle between the major axis of the ellipse and axis xx', used for the counter-clockwise rotation of the object. Additionally, A represents the height of the peak, whereas, σ_x and σ_y, are the standard deviations of the 1-dimensional Gaussian functions along the major and minor axes, respectively. An example of these images is presented in Fig. 4.1a.

The second set of 4000 images used in our experiments include multi-profiled objects with a single entity in each of them. Their size is again 201×201. They contain galaxies formed as 2-dimensional Sérsic-profiled objects of different indices. Sérsic model is a mathematical function expressing how the intensity I of a galaxy varies with the distance R from its center. It was published by José Luis Sérsic in Sérsic, 1963, as a generalization of de Vaucouleurs' law. This reads:

$$I(R) = I_e e^{-b_n \left[\left(\frac{R}{R_e}\right)^{1/n} - 1\right]}, \tag{4.5}$$

where I_e is the half-light intensity, R_e is the corresponding radius, and n represents the index parameter that defines the profile's degree of curvature. The parameter b_n is calculated as $\frac{1}{2}\Gamma(2n)$.

Finally, we have created a set of 1000 images of size 256×256 containing 20 multi-profiled objects as described above in each of them. They are located in random positions, with the restriction of avoiding overlapping. These are useful to test these methods that do not require a single object at each run, but they are able to measure multiple galaxies in the field of view, depending on the input. Figure 4.1b gives an example of an image with multiple simulated entities for extraction and measurement.

(a) Image with a single object   **(b)** Image with multiple objects
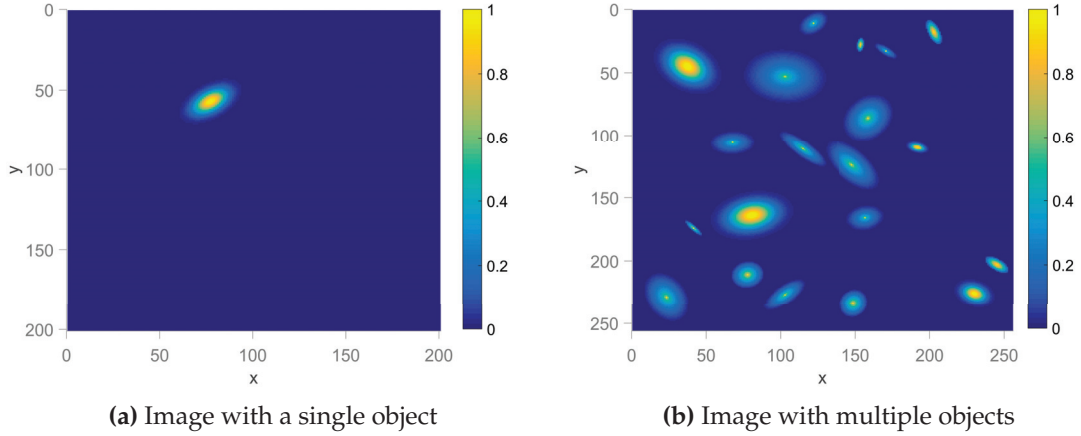
Figure 4.1 – Examples of the intensity images that are contained in the simulated databases we constructed, with: **(a)** a single object in it, and **(b)** multiple objects included.

4.2   Experimental set-up

In this part, we describe the details of the several data subsets explained in Section 4.1, giving more information on selecting the parameters in the intensity image models and the generation of the interferometric data.

For the set of Gaussian-type objects, the value of θ, as defined in 4.1, is chosen in the interval $[0, \pi/2)$ in a uniform way, and A is fixed to be 1, producing galaxies with a unit maximum pixel value. To achieve a final image with values in the interval $[0, 1]$, the minimum value is determined after a normalization.

For σ_x and σ_y, the first is uniformly selected in the interval $[2, 10]$ and the latter corresponds to the value that gives the desired ellipticity, with the greater of them to be assigned to the major axis. The ellipticity for each object is the realization of a Rayleigh random variable with mean 0.25, as shown in Figure. 4.2a), assumed to be a realistic mean ellipticity for our purpose (Geisler & Hodge, 1980; Rivi et al., 2016; van Uitert et al., 2012). Figure 4.2b outlines the distribution of the major axis (diameter) size for the galaxies included in this set, in terms of the number of pixels. The pixels considered for this estimation have values greater than or equal to half of the maximum value in the object (i.e., for our simulation 0.5). The red line in the same figure shows the size of the PSF for the main coverage of the Gaussian type, which is 4 pixels.

The object's center is randomly located using a uniform distribution, considering that the whole object's body (values $\geqslant 0.5$) must be included inside the image. During the preprocessing of each image in our framework, the object is extracted, forming a new smaller image for processing. This aims not only to decrease the

(a) Distribution of the galaxies ellipticity.

(b) Distribution of the galaxies ellipse major axis (diameter) size in number of pixels.
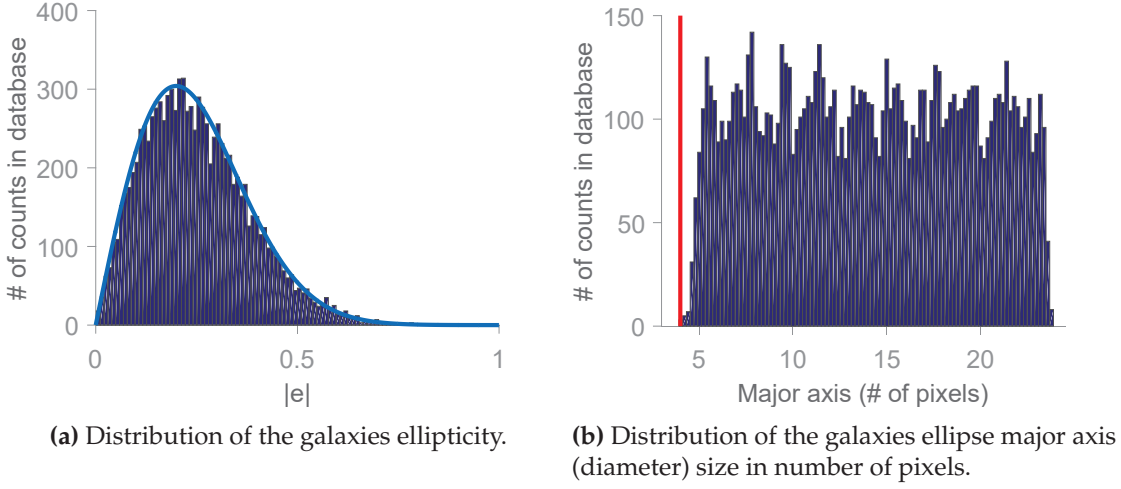
Figure 4.2 – (a) The distribution of the galaxies ellipticity in the whole database. **(b)** The distribution of the galaxies' ellipse major axis (diameter) size in the whole database, in terms of the number of pixels. The red line represents the corresponding PSF size in number of pixels.

computational complexity but also to center the objects, which is a requirement for an accurate decomposition with the shapelet basis.

The dataset with multi-profile objects follows the same main principles, with the difference of the parameters in the intensity model. The effective radius for the generated entities ranges in $[5, 50]$, with an index randomly chosen between 0.5 and 4. The images are again normalized in the interval $[0, 1]$, and the presented result is given in a logarithmic scale. The same rules apply for the multi-object images, where we create 20 items of the same logic in the image.

The images contained in the developed database include intensity data; therefore, a system to create the visibilities is required. For this purpose, we have generated a set of coverage patterns as a collection of visibility points in the uv plane. We have simulated two types of patterns for our needs. In the first one, we consider a coverage that follows the Gaussian profile, and the points **k** are random values of a 2-dimensional distribution with zero mean and standard deviation $\pi/2$, which belong to the square frame $[-\pi, \pi] \times [-\pi, \pi]$. We have simulated four collections following this pattern, with sizes M, 14365, 22155, 32131, and 128778 points. Figure 4.3a shows the visibility coverage with M = 32131 used for our experiments, which can be generated by 254 radio antennas.

The second type of pattern we simulate is a more realistic one, which follows this of the future SKA telescope. In Figure ,4.3b we see a coverage simulating a SKA mid-frequency array with 254 dishes, which uses the earth rotation aperture synthesis

(a) Gaussian-type coverage   **(b)** Realistic SKA-like coverage
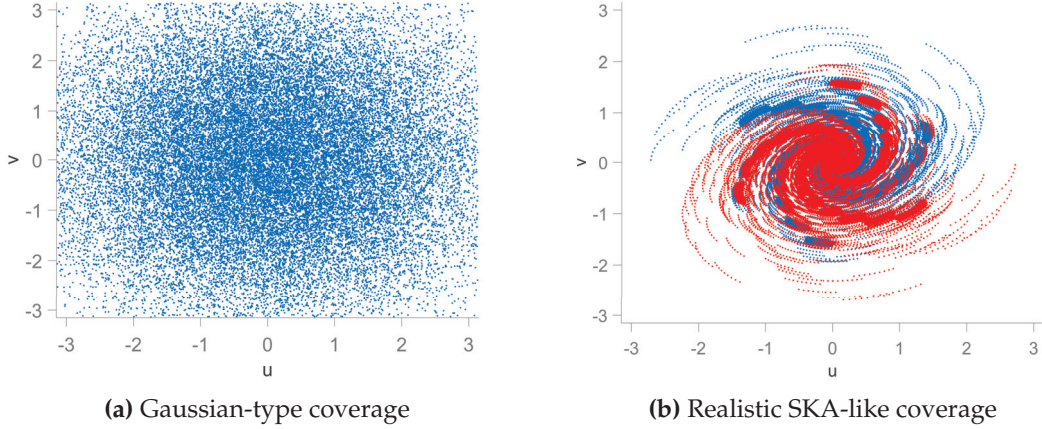
Figure 4.3 – Examples of a randomly selected density coverage in the Fourier domain for our experiments. **(a)** A Gaussian-type coverage with M = 32131, which can be created by a radio interferometer using 254 antennas. **(b)** A realistic SKA-like coverage of M = 578358 points from 254 antennas, with observation time 7200 sec and integration time 800 sec.

method identified by the elliptical tracks produced. Again, the points are framed in $[-\pi, \pi] \times [-\pi, \pi]$ and the tested versions include 128524 and 578358 points.

During the experiments, the chosen collections are constant per pattern to facilitate the comparison of the algorithms' performance. Hereafter, when we mention 'Gaussian-type coverage', we mean the one with 32131 points, and for 'realistic SKA-like', we mean the one with 578358 uv samples. The different sizes per pattern are used to test the impact of the coverage size change on the efficiency of our methods. The original simulated intensity data contain no noise. This is white Gaussian and is added after the generation of the visibilities and before the beginning of the main measurement process.

Here we have to mention that in our model, we are limited to the pixel size regarding the resolution that can be processed. Therefore, using the above-mentioned coverage patterns, objects that correspond to a size close to or smaller than the pixel size will lead to inaccurate measurements. Moreover, the models applying a fitting approach do not consider the effect of the PSF, and we do not proceed with a deconvolution, but we work directly with the input data.

4.3   Measurement error estimation

Knowing that the ellipticity, as we described in Section 1.3, is defined as a complex number, we assume that it is formed as $e = e_1 + ie_2$, where $e_1 = |e| \cos(2\theta)$ and $e_2 = |e| \sin(2\theta)$, where θ is the clockwise angle between the object's major axis and

xx'. Therefore, we can define the following three types of errors that we measure in our analysis:

$$\epsilon_{real} = \left| \text{Re}\left(e^t\right) - \text{Re}\left(e^m\right) \right|, \tag{4.6}$$

$$\epsilon_{imag} = \left| \text{Im}\left(e^t\right) - \text{Im}\left(e^m\right) \right|, \tag{4.7}$$

$$\epsilon_{abs} = \left| e^t - e^m \right|, \tag{4.8}$$

where e^t and e^m stand for the true and the measured ellipticity, respectively. Additionally, $|\cdot|$ for the first two equations calculates the absolute value and for the third one the complex magnitude. The error results presented in the tables of this chapter correspond to the mean value of the aforementioned quantities.

In the next step, following Heymans et al., 2006 and Patel et al., 2015, we fit a linear model to the measured ellipticity values and form three types of fitting, similarly to the error measurement:

$$\text{Re}\left(e^m\right) = \left(1 + m_{real}\right)\text{Re}\left(e^t\right) + c_{real}, \tag{4.9}$$

$$\text{Im}\left(e^m\right) = \left(1 + m_{imag}\right)\text{Im}\left(e^t\right) + c_{imag}, \tag{4.10}$$

$$\left|e^m\right| = \left(1 + m_{abs}\right)\left|e^t\right| + c_{abs}. \tag{4.11}$$

Here m_{real}, m_{imag} and m_{abs} are the multiplicative biases that occur in our fitting, and c_{real}, c_{imag} and c_{abs} are the additive biases for each case. Ideally, the values of these parameters must be equal to zero. When the multiplicative bias is not zero, it indicates calibration errors due to effects such as weight biases or noise. On the contrary, a non-zero additive bias can be a consequence of a systematic effect where even circular objects may appear elliptical. Furthermore, for every value of error or bias in each of the experiments, we give a confidence interval, calculated after dividing the whole image dataset used for each experiment into several not overlapping subsets and making the measurements again. The size of the interval represents the difference between the minimum and maximum values and the mean error or bias along the experiment.

4.4   LS-simple fitting

The baseline least-squares model variants are considered in the first part of our experiments, where each image processing is done independently until the shape measurement step. A very crucial choice for the performance of the whole model is for the values of the shapelet parameters, so we tested several approaches that aim to approximate them. We evaluate the algorithm's performance regarding both the accuracy and precision of the measurement and the computation time. The whole dataset with 10000 of Gaussian-profile objects was used for this part of the

(a) LS-simple

(b) LS n_{max} = 4

Figure 4.4 – Plots that present the difference between the measured and the true ellipticity as a function of the true values for the least-squares model variants, where: **(a) LS-simple** - the parameter approximation equations 2.21 and 2.22 are used. **(b) LS n_{max}** = 4 - the parameter approximation equation 2.21 for β is used and n_{max} remains constantly equal to 4. In all cases, the difference of the measured and the true values is presented as a function of the true ones (blue dots) and is linearly fitted with the red line. From left to right, we show the cases for the real part of the ellipticity, the imaginary part of the ellipticity, and its complex magnitude. The studied dataset contains simulated images with a single Gaussian-profiled object in each of them, and the corresponding visibilities were generated using a Gaussian-type coverage.

experiments. In constructing the visibilities, we use the Gaussian-type coverage, and we add white Gaussian noise with SNR = 20 dB.

In the first variant, we compute the values for β and n_{max} using the equations 2.21 and 2.22 presented in Section 2.3. Figure 4.4a depicts the fitting of the linear model for the three measurements, and the respective measurement errors and biases are given in Table 4.1.

Our next experiment (**LS n_{max}** = 4) focuses on the fact that the decomposition coefficients used for the shape measurement correspond to basis functions with order up to 4. Assuming that the shapelet functions of higher order are not necessary for a precise estimation, we adopt the minimum possible, according to the specifications, value for n_{max} and we decompose the same dataset, using a constant n_{max} = 4. Figure 4.4b shows the difference of the true and the measured ellipticity as a function of the true values, accompanied by the fitting we made to the data. The errors and biases that we got for this case are presented in Table 4.2.

Table 4.1 – Measurement errors and fitting biases, using the method **LS-simple** to Gaussian-profile data with Gaussian-type coverage and SNR = 20 dB

Measure	$\square_{real}$*	$\square_{imag}$	$\square_{abs}$
ϵ	0.0012 ± 10^{-4}	$0.0017 \pm 1.4 \cdot 10^{-4}$	$0.0043 \pm 1.6 \cdot 10^{-4}$
m	$0.0055 \pm 1.2 \cdot 10^{-3}$	$0.0099 \pm 1.8 \cdot 10^{-3}$	$0.0152 \pm 4.1 \cdot 10^{-3}$
c	$0.0003 \pm 1.1 \cdot 10^{-4}$	$0.0044 \pm 1.5 \cdot 10^{-4}$	$-0.0031 \pm 9.4 \cdot 10^{-4}$

Table 4.2 – Measurement errors and fitting biases, using the method **LS** $n_{max} = 4$ to Gaussian-profile data with Gaussian-type coverage and SNR = 20 dB

Measure	$\square_{real}$	$\square_{imag}$	$\square_{abs}$
ϵ	0.0023 ± 10^{-4}	$0.0037 \pm 1.4 \cdot 10^{-4}$	$0.0067 \pm 1.4 \cdot 10^{-4}$
m	$-0.0042 \pm 1.1 \cdot 10^{-3}$	$0.0098 \pm 1.4 \cdot 10^{-3}$	$0.0108 \pm 3.1 \cdot 10^{-3}$
c	$0.0011 \pm 1.2 \cdot 10^{-4}$	$0.0057 \pm 1.5 \cdot 10^{-4}$	$-0.0036 \pm 6.4 \cdot 10^{-4}$

We can observe that the model's performance is slightly decreased in terms of precision in most of the indexes. Only the multiplicative biases for the real part and the complex magnitude have smaller values. From the experiments, we noticed that the fitting with higher-order basis functions achieves a decrease of the shape measurement error by reducing most of the biases, which may be dependent on the noise in the visibilities or other systematic effects. Although, **LS** n_{max} = 4 in comparison to **LS-simple** can achieve measurements of the same quality while it preserves simultaneously computational resources.

Table 4.3 – Measurement errors and fitting biases, using the method **LS noise** to Gaussian-profile data with Gaussian-type coverage, SNR = 8 dB and $n_{max} = 4$

Measure	$\square_{real}$	$\square_{imag}$	$\square_{abs}$
ϵ	0.0023 ± 10^{-4}	$0.0030 \pm 1.5 \cdot 10^{-4}$	$0.0065 \pm 1.6 \cdot 10^{-4}$
m	$0.0059 \pm 1.2 \cdot 10^{-3}$	$0.0097 \pm 1.5 \cdot 10^{-3}$	$0.0158 \pm 1.9 \cdot 10^{-3}$
c	$0.0003 \pm 1.4 \cdot 10^{-4}$	$0.0042 \pm 1.7 \cdot 10^{-4}$	$-0.0035 \pm 3.8 \cdot 10^{-4}$

---

*Here and for the rest of the chapter, $\square$ represents either ϵ, m, or c, depending on the line of the table.

(a) Gaussian coverage

(b) Realistic SKA-like coverage
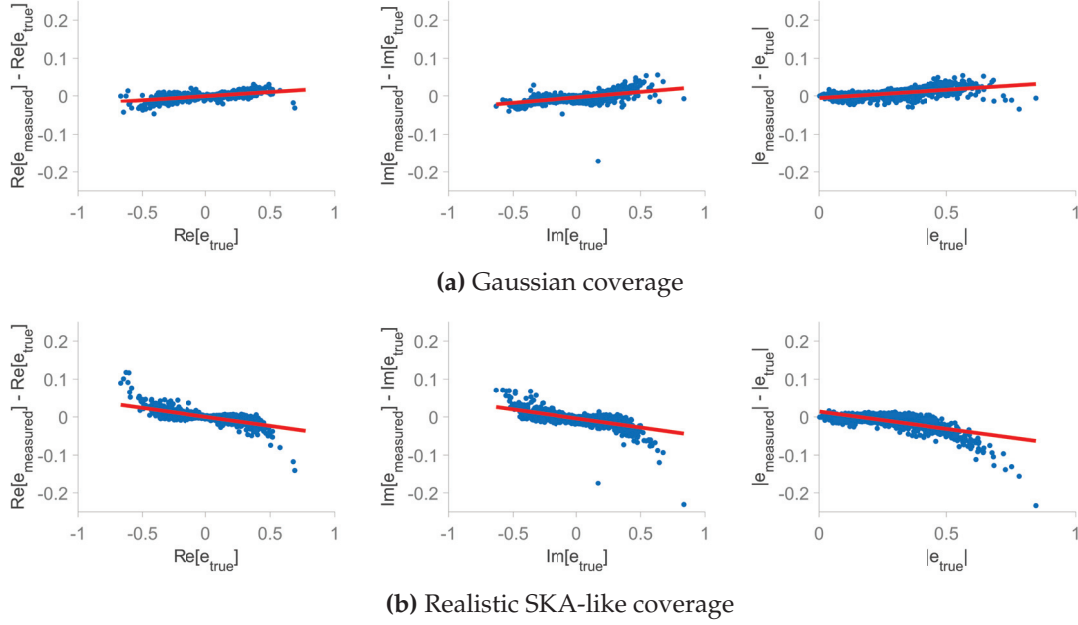
Figure 4.5 – Plots that present the difference between the measured and the true ellipticity as a function of the true values for the **LS-simple** model, where the parameter approximation equations 2.21 and 2.22 are used, applied to multi-profile data. **(a)** The results when the coverage is Gaussian type. **(b)** The results when we use a realistic SKA-like coverage. In all cases, the difference of the measured and the true values is presented as a function of the true ones (blue dots) and is linearly fitted with the red line. From left to right, we show the cases for the real part of the ellipticity, the imaginary part of the ellipticity, and its complex magnitude. The studied dataset contains simulated images of multi-profiled objects with a single entity in each of them.

Table 4.4 – Measurement errors and fitting biases, using the method **LS-simple** to multi-profile data with Gaussian-type coverage and SNR = 20 dB.

Measure	$\square_{real}$	$\square_{imag}$	$\square_{abs}$
ϵ	$0.0045 \pm 6.3 \cdot 10^{-5}$	$0.0068 \pm 6.0 \cdot 10^{-5}$	$0.0091 \pm 1.1 \cdot 10^{-4}$
m	$0.0215 \pm 3.2 \cdot 10^{-3}$	$0.0287 \pm 2.8 \cdot 10^{-3}$	$0.0436 \pm 5.0 \cdot 10^{-3}$
c	$-0.0008 \pm 2.2 \cdot 10^{-5}$	$-0.0040 \pm 1.8 \cdot 10^{-4}$	$-0.0059 \pm 3.1 \cdot 10^{-4}$

Finally, we tested the variant with $n_{max} = 4$ when for more intense additive noise, modifying the SNR value from 20 to 8 (hereafter, referred to as 'LS noise'). The measurements and the fitting of a linear model gave us, respectively, for each case the errors and multiplicative and additive biases shown in Table 4.3.

Table 4.5 – Measurement errors and fitting biases, using the method **LS-simple** to multi-profile data with realistic SKA-like coverage and SNR = 20 dB.

Measure	$\Box_{real}$	$\Box_{imag}$	$\Box_{abs}$
ϵ	$0.0068 \pm 1.7 \cdot 10^{-4}$	$0.0083 \pm 1.5 \cdot 10^{-4}$	$0.0119 \pm 2.99 \cdot 10^{-4}$
m	$0.0473 \pm 2.7 \cdot 10^{-3}$	$0.0475 \pm 4.1 \cdot 10^{-3}$	$0.0912 \pm 5.4 \cdot 10^{-3}$
c	$-0.00003 \pm 1.9 \cdot 10^{-5}$	$-0.0041 \pm 1.9 \cdot 10^{-4}$	$-0.0141 \pm 3.3 \cdot 10^{-4}$

We observe that the increase of the noise power gives a slight reduction in the efficiency of this method, which is logical because the source information is significantly more distorted than before. However, the most interesting conclusion is that with the considerably increased noise power, the shape measurement accuracy did not change dramatically, proving that this model variant is very resistant to noise.

After working with the Gaussian-profiled images, we apply the **LS-simple** method to the dataset consisting of 5000 multi-profiled objects put in separate images. We want to evaluate how the complexity of the input changes the efficiency of the measurements. In Table ,4.4 we show the results from these estimations, where we mark a small increase in the numbers, especially in the multiplicative biases. Similarly, we study the impact of the coverage type in the performance of **LS-simple** by modifying the previously used Gaussian-type pattern to the realistic SKA-like one. Although the results express a slight change, particularly the multiplicative biases, we notice a considerably precise shape estimation in most of the cases shown in Table 4.5. The conclusions of both experiments described above can also be verified by Figures 4.5a and 4.5b that provide plots of the difference between the measured and the true ellipticity as a function of the true values.

As a general observation, we deduce that the tested variations perform quite well in the presence of intense noise when we increase the data complexity or change the coverage pattern to a realistic one. The mean processing time per image for the presented experiments was about 0.6 seconds.

4.5   GLASSO regularization

This section is based on the work on the model that uses the GLASSO regularization strategy to provide an automatic or semi-automatic estimation of the β parameter. The motivation to study this model and its variants is the expectation that successfully finding a scale value that minimizes the object's fitting error using the shapelet dictionary with this scale will lead us to minimize the measurement error. To test this assumption, we used 101 scale values to create the same number of shapelet dictionaries and decompose 2200 images with multi-profile objects, resulting in 222200
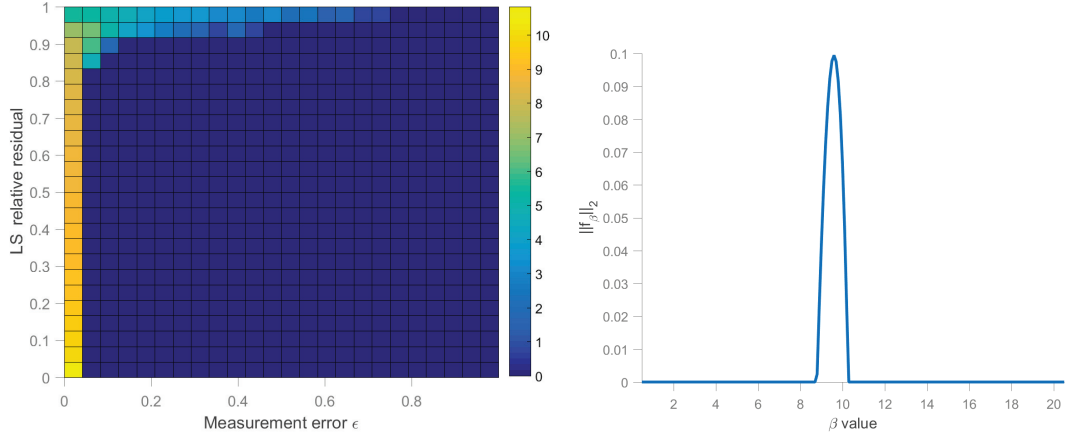
(a) Correlation histogram between ϵ_{abs} and LS relative residual.

(b) Coefficients example after GLASSO application.

Figure 4.6 – (a) A 3D histogram that shows the correlation between the complex magnitude of the measurement error and the relative residual from the solution of the least-squares problem. Values have been assigned to 25 bins per dimension after normalization, and the counts per bin are given on a logarithmic scale. **(b)** An example state after the end of GLASSO regularization procedure, where only a few groups of coefficients have a non-zero ℓ_2 norm.

measurements of the relative residual from the least-squares approximation, as defined in equation 3.3, and the respective measurement error, from equation 4.8. The goal was to investigate the correlation between these two quantities. In Figure 4.6a, we see a 3D histogram with the counts per bin in logarithmic scale that depicts this correlation, after normalizing their values from 0 to 1. As we may observe, the great majority of the cases confirm our assumption. Therefore, generally speaking, we can minimize the measurement error by minimizing the respective relative residual. Nonetheless, we expect our algorithm to give inaccurate estimations in individual cases where the above-mentioned correlation is not true. Moreover, we must clarify that the vice-versa assumption is, in general, not valid.

Due to the complexity of this problem, the execution of the algorithm is slower. For this reason, we have reduced the number of tested images while keeping the additive to the visibilities noise to the same level (SNR = 20) and the coverage to Gaussian-type.

To begin with, we evaluate the framework's variant that uses a constant long dictionary with 201 different β values ranging from 0.5 to 20.5 and n_{max} = 4. The resulting dictionary has 3015 columns representing the same number of shapelet functions. We work with 1000 images with Gaussian-profile objects, letting the sparse minimization algorithm select the scale value to optimize the object's shape measurement. Then, we follow the **LS-simple** approach with a short dictionary of a

single β and 15 columns. In Figure ,4.6b we can observe an example state after the end of this GLASSO regularization. There, only a few groups have a non-zero ℓ_2 norm of their decomposition coefficients. The scale value chosen for the following steps corresponds to the group with the maximum ℓ_2 norm. Figure 4.7a presents the ellipticity estimation for this case, where we see that large true ellipticity values produce the largest measurement errors. Table 4.6 captures the measurement errors in this experiment and the relevant biases from the linear fitting.

Table 4.6 – Measurement errors and fitting biases, using the method **GLASSO automatic** to Gaussian-profile data with Gaussian-type coverage and SNR = 20 dB

Measure	$\square_{real}$	$\square_{imag}$	$\square_{abs}$
ϵ	$0.0082 \pm 1.1 \cdot 10^{-3}$	0.0110 ± 10^{-3}	$0.0134 \pm 1.4 \cdot 10^{-3}$
$\mathfrak{m}$	0.062 ± 10^{-2}	$0.058 \pm 0.9 \cdot 10^{-2}$	$0.122 \pm 1.4 \cdot 10^{-2}$
$\mathfrak{c}$	$0.0014 \pm 0.9 \cdot 10^{-3}$	$0.0042 \pm 0.8 \cdot 10^{-3}$	$-0.0244 \pm 3.4 \cdot 10^{-3}$

We can form another version of the algorithm by using the information from the approximation equation 2.21 for the scale parameter. Here, we will again use the GLASSO regularization approach to β value, but this time we create an individual long dictionary for each object with scale values in the range of this approximation. Consequently, we form a dictionary with 3015 shapelet columns using 201 scale values on a higher resolution grid than the experiment above. We process the same 1000 images with the same coverage type, and we get the selected value from the GLASSO regularization, proceeding as before. The shape measurement results from this study are given in Figure 4.7b, whereas the errors and the respective biases are included in Table 4.7.

Table 4.7 – Measurement errors and fitting biases, using the method **GLASSO initial guess** to Gaussian-profile data with Gaussian-type coverage and SNR = 20 dB

Measure	$\square_{real}$	$\square_{imag}$	$\square_{abs}$
ϵ	$0.0051 \pm 6.7 \cdot 10^{-4}$	$0.0085 \pm 7.5 \cdot 10^{-4}$	$0.0119 \pm 9.3 \cdot 10^{-4}$
$\mathfrak{m}$	$0.0392 \pm 5.4 \cdot 10^{-3}$	$0.0429 \pm 5.9 \cdot 10^{-3}$	$0.1009 \pm 8.3 \cdot 10^{-3}$
$\mathfrak{c}$	$0.0008 \pm 5.9 \cdot 10^{-4}$	$0.0049 \pm 6.5 \cdot 10^{-4}$	$-0.0157 \pm 1.8 \cdot 10^{-3}$

Having a closer look at Figure 4.7b, we observe that the algorithm again tends to give measurements with larger errors when the object's true ellipticity is greater than 0.4.

This observation motivated us to develop one more variant of the model that mixes the sparse regularization, and the **LS-simple** approaches. In this sub-model, we let the algorithm itself select the cases to be processed using GLASSO regularization.

(a) GLASSO automatic

(b) GLASSO initial guess

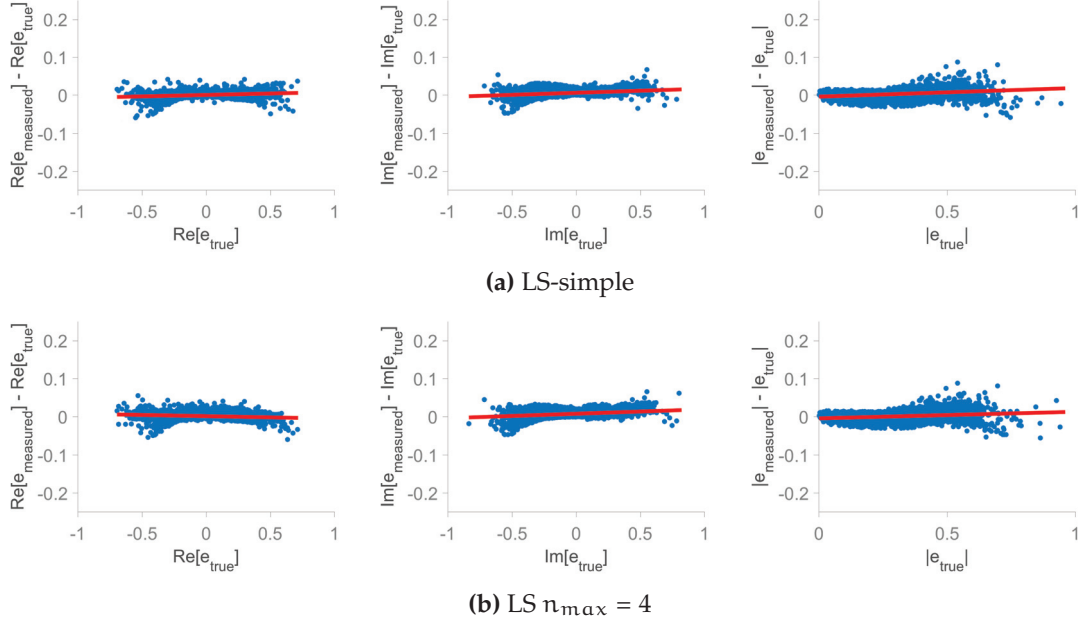**(c)** GLASSO mixed with LS

(d) GLASSO increasing precision

Figure 4.7 – Plots that present the difference between the measured and the true ellipticity as a function of the true values for the GLASSO model variants, where: **(a) GLASSO automatic** - the parameter is chosen from a constant for all images long dictionary with β range in $[0.5, 20.5]$. **(b) GLASSO initial guesss** - the parameter is chosen from a individual long dictionary per image with scale values based on the approximation equation 2.21. **(c) GLASSO mixed with LS** - all the images are processed using the baseline least-square framework (LS $n_{max} = 4$), and only these with large ellipticity in the first part are reprocessed using the GLASSO. **(d) GLASSO increasing precision** - β is chosen initially in $[0.5, 20.5]$ and, then, iteratively in dictionary of increasing precision in the range of the optimal value of the previous step. In all cases, the difference of the measured and the true values is presented as a function of the true ones (blue dots) and is linearly fitted with the red line. From left to right, we show the cases for the real part of the ellipticity, the imaginary part of the ellipticity, and its complex magnitude. The studied dataset contains simulated images with a single Gaussian-profiled object in each of them, and the corresponding visibilities were generated using a Gaussian-type coverage.

To achieve this goal, we use the criterion of the ellipticity magnitude, a value that we do not have in advance. Thus, in the first processing step, all images are analyzed with the baseline model. In the second step, we use the prior knowledge of this shape measurement, considering that the galaxies measured in the first part to have ellipticity complex magnitude greater than 0.5 tend to produce larger measurement errors. With this in mind, we make a second measurement for them using the GLASSO regularization. This variant is faster than the previously mentioned ones because only a part of the dataset is processed with the complex model. Therefore, in our experiments, we work with the whole database of the 10000 images of Gaussian-profile objects and the Gaussian-type coverage. The results are a mixture of those from the two sub-models, and the values of the β parameter in the long dictionary are chosen based on the approximation equation using the object's rms radius. The value of n_{max} remains constant and equal to 4 throughout the whole experiment. In Fig. ,4.7c one may see the difference between the measurements that we got and the true values as a function of the latter. According to 4.3, we got the estimation errors and the fitting biases shown in Table 4.8. After analyzing these results, it is clear that the mixture of these approaches did not provide more accuracy on the measurements and, more specifically, for the objects with large ellipticity that has been the main reason to propose this solution.

Table 4.8 – Measurement errors and fitting biases, using the method **GLASSO mixed with LS** to Gaussian-profile data with Gaussian-type coverage and SNR = 20 dB

Measure	$\square_{real}$	$\square_{imag}$	$\square_{abs}$
ϵ	$0.0052 \pm 1.7 \cdot 10^{-4}$	$0.0099 \pm 2.1 \cdot 10^{-4}$	$0.0124 \pm 2.6 \cdot 10^{-4}$
m	$0.0129 \pm 2.1 \cdot 10^{-3}$	$0.0284 \pm 2.3 \cdot 10^{-3}$	$0.0649 \pm 4.2 \cdot 10^{-3}$
c	$0.0012 \pm 1.9 \cdot 10^{-4}$	$0.0078 \pm 2.1 \cdot 10^{-4}$	$-0.0144 \pm 8.5 \cdot 10^{-4}$

For the last variant tested, we apply an automatic selection of the scale parameter iteratively by starting with a dictionary of large range and rough precision and then focusing on the interval around the value of the best fit for the subsequent step in a similar fashion to a multi-resolution analysis. Therefore, in each further iteration, the constructed dictionary has decreased range and increased precision at the magnitude of 10^{-1}. We process 1000 images of Gaussian-profile objects, using at each of the 3 iterations a dictionary of 20 different scale values and n_{max} = 4, and Gaussian-type coverage. The measurement errors for this part of the experiments are presented in Table 4.9. In addition, there, we may see the resulting biases after the linear fitting of the measured ellipticity values.

As one may observe in Figure 4.7d, this technique has improved the produced results, forming a model that exceeds the performance of all previous that use GLASSO regularization, but it cannot outperform the baseline least-squares variant.
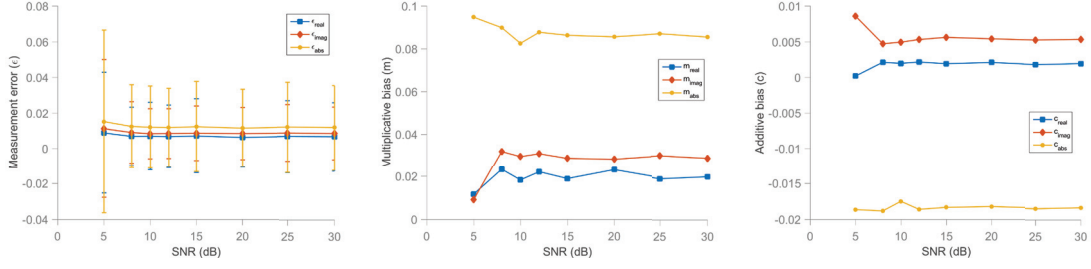
Figure 4.8 – Plots that present from left to right the measurement errors along with their variances, the multiplicative biases and the additive biases, as functions of the noise power in the model. The analysis is made using the **GLASSO increasing precision** variant. The values of the SNR used for the experiments were $5, 8, 10, 12, 15, 20, 25$ and 30 dB.

Although it provides a good alternative that gives very good results in comparison to the baseline model, and it is quite fast considering the complexity of the minimization problem. Furthermore, we have to mention that an important drawback of this version is the sensitivity in the calibration of the regularization parameter (see Section 3.2 for the used estimation), which becomes greater as the precision of the dictionary increases.

Table 4.9 – Measurement errors and fitting biases, using the method **GLASSO increasing precision** to Gaussian-profile data with Gaussian-type coverage and SNR = 20 dB

Measure	$\square_{real}$	$\square_{imag}$	$\square_{abs}$
ϵ	0.0048 ± 10^{-3}	$0.0089 \pm 0.9 \cdot 10^{-3}$	$0.0128 \pm 1.4 \cdot 10^{-3}$
m	0.042 ± 10^{-2}	$0.047 \pm 0.9 \cdot 10^{-2}$	$0.112 \pm 1.5 \cdot 10^{-2}$
c	$0.0011 \pm 0.9 \cdot 10^{-3}$	$0.0054 \pm 0.8 \cdot 10^{-3}$	$-0.0179 \pm 3.4 \cdot 10^{-3}$

Similarly to the baseline model, the importance of noise in the algorithm's performance has to be analyzed. Thus, we conducted a set of experiments where the best-in-performance GLASSO variant is used, and the noise power is varied from 5 dB to 30 dB. For comparison reasons, the set of 1000 images used in each of these experiments is the same.

In Figure 4.8, we can see that the method presents very robust characteristics, since with noise power up to 8 dB, the output remains almost unchanged. Only when the SNR is further increased, the variance of the measurement error also rises.

One more aspect that had to be investigated was the importance of the precision of the center localization in the measurement results. In our assumptions, we considered this parameter to be predefined accurately, as the object's centers have been set to be the origin. However, since this is not true in general, we have assumed that the algorithm in our possession has the ability to locate the center with an error

on the magnitude of 10% of the size of the object. Then, we used the same model (**GLASSO increasing precision**) to analyze the same images with the original case. Table 4.10 depicts the results after the measurements and the corresponding linear fitting.

Table 4.10 – Measurement errors and fitting biases, using the method **GLASSO missed centres** to Gaussian-profile data with Gaussian-type coverage and SNR = 20 dB

Measure	$\Box_{real}$	$\Box_{imag}$	$\Box_{abs}$
ϵ	$0.048 \pm 1.2 \cdot 10^{-2}$	$0.054 \pm 1.2 \cdot 10^{-2}$	$0.087 \pm 1.6 \cdot 10^{-2}$
$\mathfrak{m}$	$-0.118 \pm 8.9 \cdot 10^{-2}$	$-0.14 \pm 1.4 \cdot 10^{-1}$	$-0.24 \pm 1.2 \cdot 10^{-1}$
c	$-0.009 \pm 1.2 \cdot 10^{-2}$	$0.034 \pm 1.1 \cdot 10^{-2}$	$0.052 \pm 2.3 \cdot 10^{-2}$

It is apparent that all quantities have been increased significantly, with the measurement errors being even 7 times larger, meaning that the model is sensitive to the miscalculation of each object center. Hereafter, the results of the experiment will be referred to as **GLASSO missed centers**.

Studying the variants of the framework that employ sparse regularization, we notice that we can not outperform the outcomes of the baseline model. Therefore, we studied if and how much the GLASSO minimization approach can be improved. Having in mind that the main assumption which motivated us to develop this model was that the minimization of the fitting error would lead us in most of the cases to the best measurement, we decided to investigate what would be the influence to the result if we had found the β value that gives the minimum fitting error. Therefore, we tested 1000 images of Gaussian-profile data for 2000 scale values with precision 0.01, locating the value that minimizes the least-squares relative residual. We assumed that it was selected as optimal using the GLASSO technique, and we recorded the measurement errors and the respective fitting biases shown in Table 4.11.

Table 4.11 – Measurement errors and fitting biases, using the method **GLASSO best** to Gaussian-profile data with Gaussian-type coverage and SNR = 20 dB

Measure	$\Box_{real}$	$\Box_{imag}$	$\Box_{abs}$
ϵ	$0.0041 \pm 2.4 \cdot 10^{-4}$	$0.0051 \pm 3.4 \cdot 10^{-4}$	$0.0100 \pm 5.9 \cdot 10^{-4}$
$\mathfrak{m}$	$0.0029 \pm 4.8 \cdot 10^{-3}$	$0.0075 \pm 3.9 \cdot 10^{-3}$	$0.0199 \pm 5.2 \cdot 10^{-3}$
c	$0.00007 \pm 3.2 \cdot 10^{-5}$	$0.0071 \pm 1.1 \cdot 10^{-4}$	$-0.0115 \pm 2.6 \cdot 10^{-4}$

Here, the measurements are significantly better than all other cases in this section, but the performance is still a little inferior to these methods based only on the least-squares. This means that even if we achieve to optimize the performance

Table 4.12 – Resulting values $1 - R^2$ for the measurement of the badness of the fitting in GLASSO models, using Gaussian-type coverage in Gaussian-profile data, for both linear and higher polynomial fitting.

Model	Linear fitting			Higher order polynomial fitting		
	Real part	Imaginary part	Complex magnitude	Real part	Imaginary part	Complex magnitude
4	$48.1 \cdot 10^{-4}$	$43.5 \cdot 10^{-4}$	$107.4 \cdot 10^{-4}$	$12.6 \cdot 10^{-4}$	$15.4 \cdot 10^{-4}$	$32.5 \cdot 10^{-4}$
5	$18.9 \cdot 10^{-4}$	$24.2 \cdot 10^{-4}$	$54.5 \cdot 10^{-4}$	$8.2 \cdot 10^{-4}$	$11.4 \cdot 10^{-4}$	$34.6 \cdot 10^{-4}$
6	$23.1 \cdot 10^{-4}$	$25.8 \cdot 10^{-4}$	$85.4 \cdot 10^{-4}$	$12.9 \cdot 10^{-4}$	$16.3 \cdot 10^{-4}$	$60.8 \cdot 10^{-4}$
7	$44.7 \cdot 10^{-4}$	$38.7 \cdot 10^{-4}$	$103.2 \cdot 10^{-4}$	$11.9 \cdot 10^{-4}$	$14.9 \cdot 10^{-4}$	$40.4 \cdot 10^{-4}$

4: GLASSO automatic ; 5: GLASSO initial guess ; 6: GLASSO mixed with LS ; 7: GLASSO increasing precision.

of sparse regularization, it will not produce superior results than the baseline case. Nonetheless, it will be a very reliable model, with the major advantage of the ability to estimate the ellipticity of an object without any prior knowledge of its characteristics. The results of this case are presented in the rest of the text as **GLASSO best**.

As we can observe in Figure ,4.7 a good idea would be to try fitting a higher-order polynomial to the produced data instead of a linear one. Thus, we used a 3rd order polynomial for the fitting of the real and imaginary parts. There, we have suppressed the 2nd order of the polynomial. Similarly, we considered that the best choice for the absolute magnitude would be a 2nd order polynomial.

In Figure 4.9, we observe the resulting plots after the new fittings. There we can visually conclude that these models fit better our data than the simple linear ones. To evaluate this using a measure, we utilized the coefficient of determination R^2 that shows the goodness of a fitting. This is defined as:

$$R^2 \equiv 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \bar{y})^2},$$ (4.12)

where $\mathbf{y} = [y_1, ..., y_n]^T$ are the values that came after the measurement, $\mathbf{f} = [f_1, ..., f_n]^T$ are the ones that would be predicted using the fitted model, and $\bar{y}$ represents the mean of these values. We measured the coefficient values in all fittings using both linear and higher polynomial models to have a meaningful comparison. Table 4.12 shows for simplicity the values $1 - R^2$ that we got in all cases, which actually represent the badness of the corresponding fittings. There, we can verify our visual observations from the plots, since in all cases, the badness of the fitting has dramatically decreased when we use higher-order polynomials, with some cases being a quarter of these with the linear model. Although this phenomenon is influenced negatively due to the complex minimization procedure, which adds non-linearity to the problem, its main cause is likely related to the precision of the fitting with the shapelet basis.

(a) GLASSO automatic

(b) GLASSO initial guess

(c) GLASSO mixed with LS
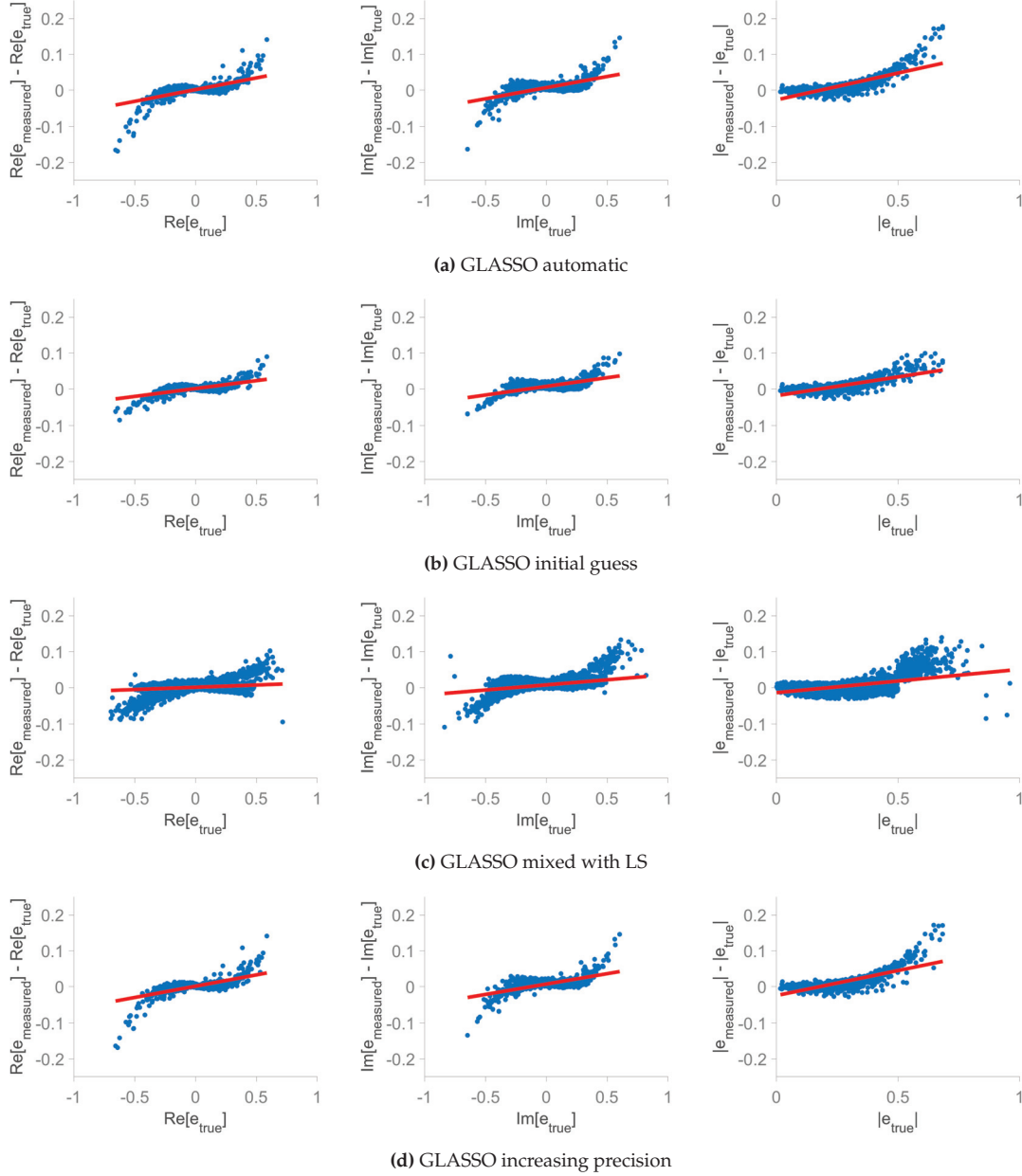

(d) GLASSO increasing precision
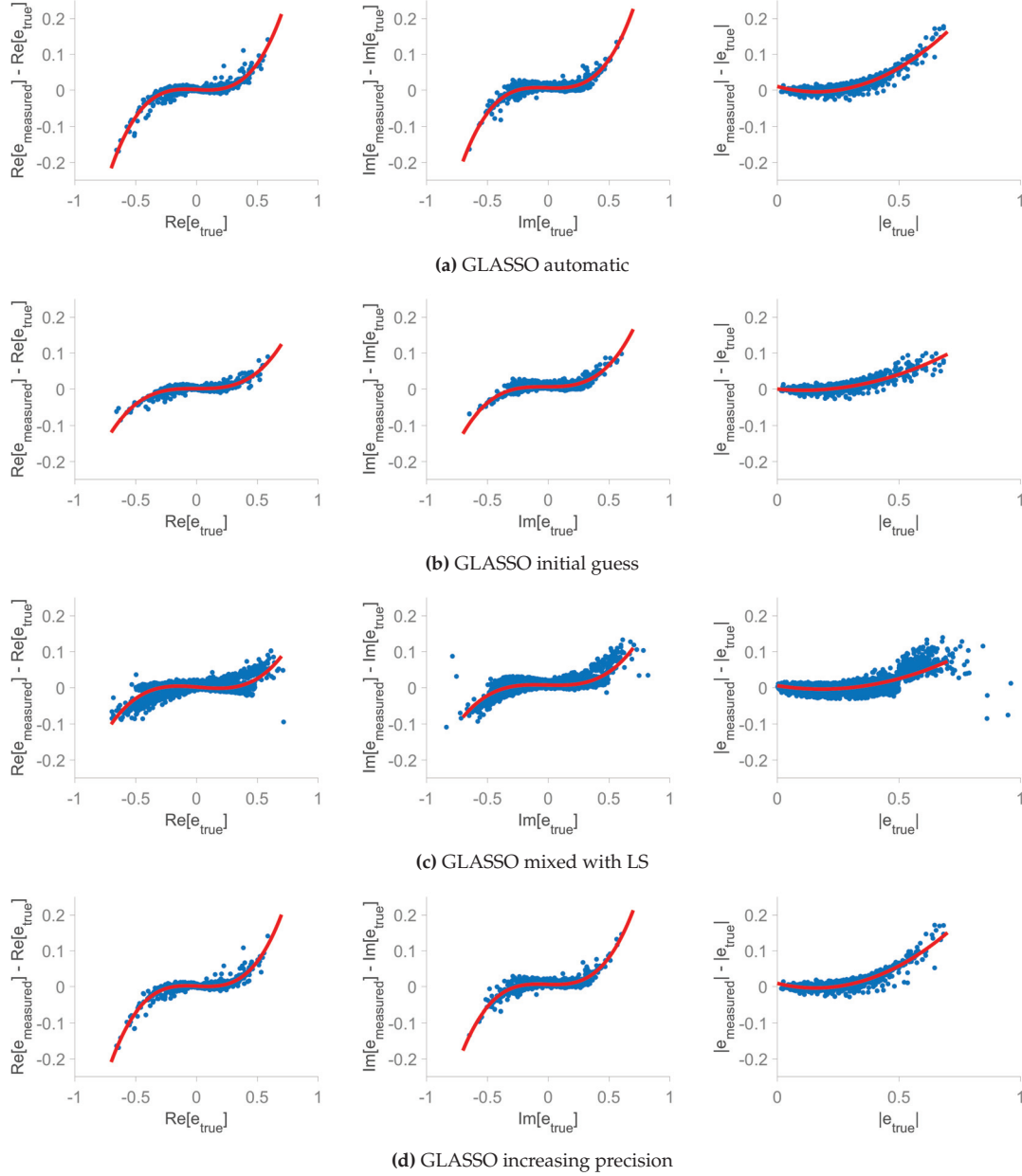
Figure 4.9 – Plots that present the difference between the measured and the true ellipticity as a function of the true values for the GLASSO model variants, where: **(a) GLASSO automatic** - the parameter is chosen from a constant for all images long dictionary with β range in $[0.5, 20.5]$. **(b) GLASSO initial guess** - the parameter is chosen from a individual long dictionary per image with scale values based on the approximation equation 2.21. **(c) GLASSO mixed with LS** - all the images are processed using the baseline least-square framework (LS $n_{max} = 4$), and only these with large ellipticity in the first part are reprocessed using the GLASSO. **(d) GLASSO increasing precision** - β is chosen initially in $[0.5, 20.5]$ and, then, iteratively in dictionary of increasing precision in the range of the optimal value of the previous step. In all cases, the difference of the measured and the true values is presented as a function of the true ones (blue dots) and is fitted with the red line using higher-order polynomials. From left to right, we show the cases for the real part of the ellipticity, the imaginary part of the ellipticity, and its complex magnitude. The studied dataset contains simulated images with a single Gaussian-profiled object in each of them, and the corresponding visibilities were generated using a Gaussian-type coverage.

After performing the tests described above, based on the dataset containing the Gaussian-profiled objects and the Gaussian-type coverage, we will investigate the influence of the data complexity on the resulting shape estimations. We conducted an experiment using the best of the model variants proposed in this section (i.e., **GLASSO increasing precision**), where we changed the input from images of Gaussian-profiled galaxies to these of multi-profiled entities. We performed this study using a collection of 2000 images from this set, and Table 4.13 presents the shape estimation errors we measured. There, one may also see the biases from the linear fitting made in the same results. The numbers show that the measurement quality has been decreased to some degree, but this can be expected after the addition of more complex profiles. This can be verified from Figure 4.10a, where we capture the difference of the measured and true ellipticity as a function of the latter.

Table 4.13 – Measurement errors and fitting biases, using the method **GLASSO increasing precision** to multi-profile data with Gaussian-type coverage and SNR = 20 dB

Measure	$\square_{real}$	$\square_{imag}$	$\square_{abs}$
ϵ	$0.0098 \pm 4.3 \cdot 10^{-4}$	$0.0155 \pm 4.5 \cdot 10^{-4}$	$0.0203 \pm 7.9 \cdot 10^{-4}$
m	$0.0468 \pm 6.7 \cdot 10^{-4}$	$0.0475 \pm 7.1 \cdot 10^{-4}$	$0.0995 \pm 9.1 \cdot 10^{-4}$
c	$-0.0049 \pm 1.6 \cdot 10^{-4}$	$-0.0101 \pm 6.9 \cdot 10^{-4}$	$-0.0167 \pm 7.2 \cdot 10^{-4}$

In the final step, we also modify the used coverage to give an even more realistic set-up in our experiments. For this purpose, we employ the SKA-like uv points collection, and we analyze again 2000 images in 2 precision levels. As explained earlier in the section, a significant disadvantage of the GLASSO regularization model is that the parameters of the group LASSO problem have to be well estimated before each analysis separately to avoid convergence to a wrong coefficient group or divergence at the end of the iterative procedure. When we work with Gaussian-type coverage, this problem is not as intense as using a SKA-like one. This very likely happens due to the extra complexity that we induce to the problem by evaluating the shapelet functions in uv points that have pair differences ranging from 10^{-8} to almost 2π. Considering this fact and the length of the dictionary, the regularization parameter and the gradient step values in the algorithm can become very sensitive to slight changes. The λ estimator in equation 3.6 gives a good parameter selection for our purposes, but we need to apply an additional estimation step for the gradient step.

Table 4.14 – Measurement errors and fitting biases, using the method **GLASSO increasing precision** to multi-profile data with realistic SKA-like coverage and SNR = 20 dB

Measure	$\square_{real}$	$\square_{imag}$	$\square_{abs}$
ϵ	$0.0255 \pm 2.5 \cdot 10^{-3}$	$0.0261 \pm 2.4 \cdot 10^{-3}$	$0.0424 \pm 4.4 \cdot 10^{-3}$
m	$-0.0082 \pm 1.7 \cdot 10^{-3}$	$-0.0190 \pm 5.7 \cdot 10^{-3}$	$0.0325 \pm 5.9 \cdot 10^{-3}$
c	$0.0001 \pm 1.2 \cdot 10^{-4}$	$-0.0077 \pm 4.3 \cdot 10^{-3}$	$-0.0116 \pm 6.5 \cdot 10^{-3}$

A very small selected value for the gradient step nullifies the coefficients vector, whereas a very large value causes the divergence of the method. Aiming to find a value that supports the convergence in most cases, we performed a search in the interval $[0.005, 2]$ applying iterative bisecting to a small subset of the dataset containing 100 images. After a series of iterations, we found that for the first precision level, a satisfying gradient step value is 0.0417124, and for the second iteration step, 0.0519836. Having identified this parameter, we use the acquired information to study the dataset of 2000 multi-profile objects with realistic SKA-like coverage. In Table ,4.14 we include the results from this part, where we can notice the increase in the measurement errors. However, the biases in the linear fitting of the errors have been reduced, which is a result of the more uniformly distributed larger errors than in the previous case, where the larger errors came from larger ellipticity values. Figure 4.10b helps to illustrate these conclusions.

4.6   MR-LS fitting

As we saw in Section ,4.5 the concept of minimizing the relative residual works well in terms of the given estimation errors. Although, the GLASSO minimization has certain drawbacks, with the most important of them being the computational complexity that translates in long execution times and several new parameters that appear in the model, transforming the problem of shapelets scale parameter estimation to a problem of group LASSO gradient step parameter estimation. This was the motivation to work on a new model that applies the same logic of residual minimization but in a simpler form.

In this context, we have built a solution that uses a multi-resolution grid search to identify the β value that achieves the best fitting regarding the relative residual. Starting with a set of values of step size 1 in the interval $[0.5, 20.5]$, and working in the same manner as the **GLASSO increasing precision** method from one precision level to the next, we analyze the image by solving only a series of least-squares problems that have a large computational difference comparing to the respective problems in GLASSO regularization. The n_{max} value is fixed to 4 for this method, which remains purely automatic, providing results in the same set of values with

(a) Gaussian coverage

(b) Realistic SKA-like coverage
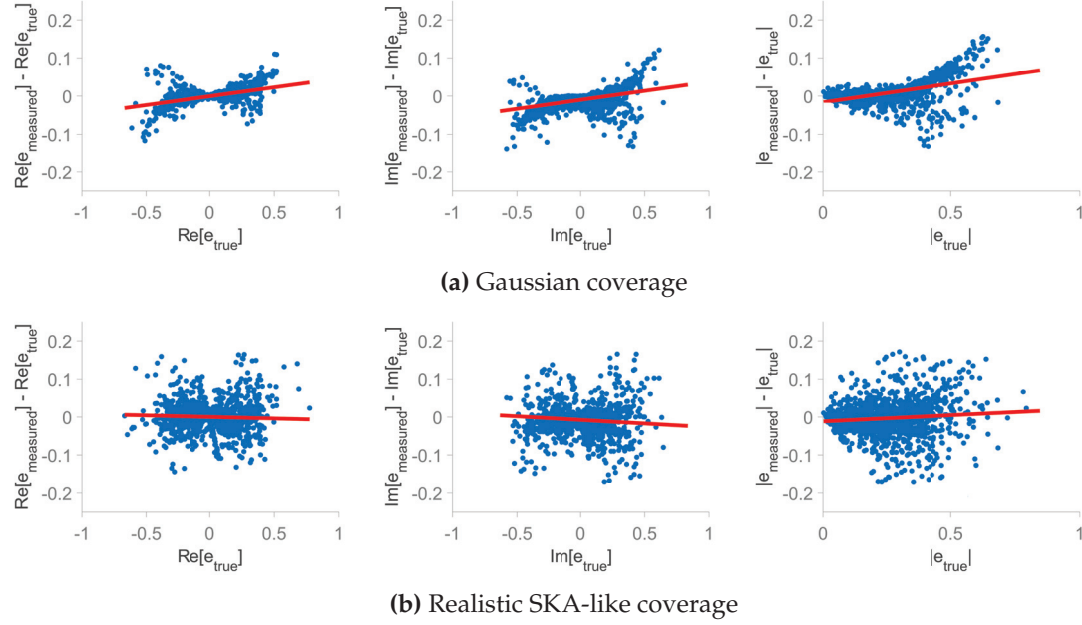
Figure 4.10 – Plots that present the difference between the measured and the true ellipticity as a function of the true values for the **GLASSO increasing precision** model, where the scale parameter is chosen initially in interval with a large range and, then, iteratively in a dictionary of increasing precision in the range of the optimal value of the previous step. **(a)** The results when the coverage is Gaussian type. **(b)** The results when we use a realistic SKA-like coverage. In all cases, the difference of the measured and the true values is presented as a function of the true ones (blue dots) and is linearly fitted with the red line. From left to right, we show the cases for the real part of the ellipticity, the imaginary part of the ellipticity, and its complex magnitude. The studied dataset contains simulated images of multi-profiled objects with a single entity in each of them.

the **GLASSO increasing precision**. All the presented experiments with this model use 3 precision iterations.

Table 4.15 – Measurement errors and fitting biases, using the method **MR-LS** to multi-profile data with Gaussian-type coverage and SNR = 20 dB

Measure	$\square_{real}$	$\square_{imag}$	$\square_{abs}$
ϵ	$0.0049 \pm 1.4 \cdot 10^{-4}$	$0.0084 \pm 3.1 \cdot 10^{-4}$	$0.0111 \pm 4.2 \cdot 10^{-4}$
m	$-0.0014 \pm 2.4 \cdot 10^{-4}$	$-0.0005 \pm 8.2 \cdot 10^{-5}$	$0.0141 \pm 8.6 \cdot 10^{-4}$
c	$-0.0006 \pm 2.4 \cdot 10^{-4}$	$-0.0099 \pm 5.5 \cdot 10^{-4}$	$-0.0045 \pm 6.1 \cdot 10^{-4}$

(a) Gaussian coverage

(b) Realistic SKA-like coverage
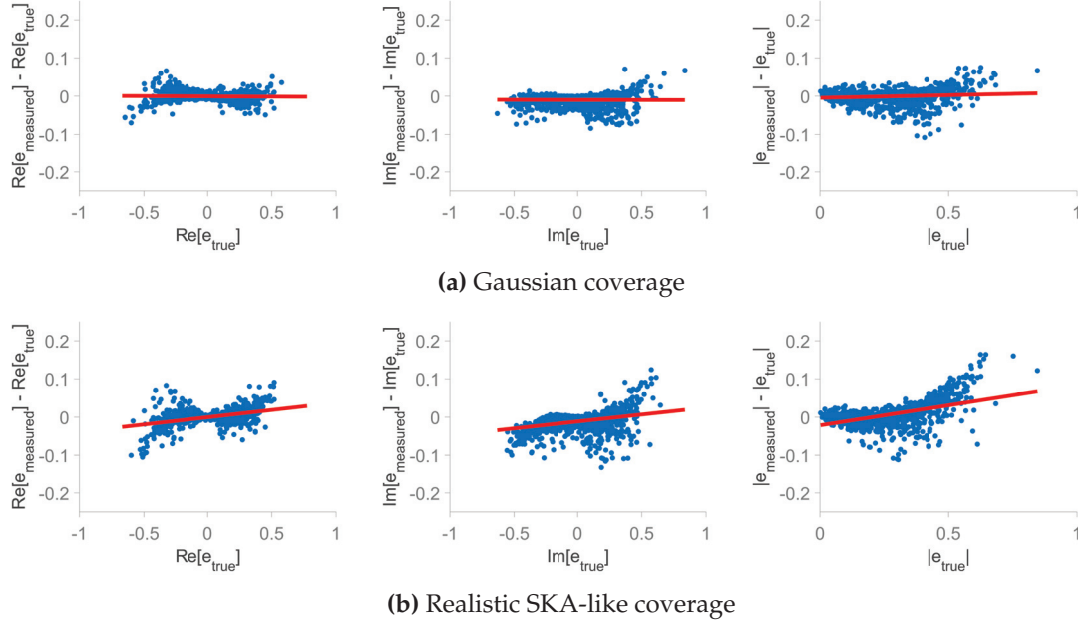
Figure 4.11 – Plots that present the difference between the measured and the true ellipticity as a function of the true values for the **MR-LS** model, where the scale parameter is chosen using grid search that minimizes the relative residual, initially in $[0.5, 20.5]$ and, then, iteratively in intervals of increased precision in the range of the value selected at the previous step. **(a)** The results when the coverage is Gaussian type. **(b)** The results when we use a realistic SKA-like coverage. In all cases, the difference of the measured and the true values is presented as a function of the true ones (blue dots) and is linearly fitted with the red line. From left to right, we show the cases for the real part of the ellipticity, the imaginary part of the ellipticity, and its complex magnitude. The studied dataset contains simulated images of multi-profiled objects with a single entity in each of them.

Table 4.16 – Measurement errors and fitting biases, using the method **MR-LS** to multi-profile data with realistic SKA-like coverage and SNR = 20 dB

Measure	$\square_{real}$	$\square_{imag}$	$\square_{abs}$
ϵ	$0.0079 \pm 3.7 \cdot 10^{-4}$	$0.0120 \pm 5.4 \cdot 10^{-4}$	$0.0178 \pm 8.5 \cdot 10^{-4}$
$\mathfrak{m}$	$0.0383 \pm 2.7 \cdot 10^{-3}$	$0.0365 \pm 2.9 \cdot 10^{-3}$	$0.0984 \pm 9.5 \cdot 10^{-3}$
$\mathfrak{c}$	$-0.0003 \pm 1.0 \cdot 10^{-4}$	$-0.0114 \pm 5.1 \cdot 10^{-3}$	$-0.0216 \pm 8.8 \cdot 10^{-3}$

Trying to evaluate the efficiency of this framework, we conduct an experiment with 2000 images of multi-profiled objects and visibilities generated using a Gaussian-type coverage, after adding Gaussian noise with SNR = 20 dB. Table 4.15 contains the errors and biases obtained from this study. The measurement has become more accurate compared to the previous section's method, decreasing the linear fitting biases simultaneously. We can verify these observations in Figure 4.11a.

(a) Gaussian coverage

(b) Realistic SKA-like coverage
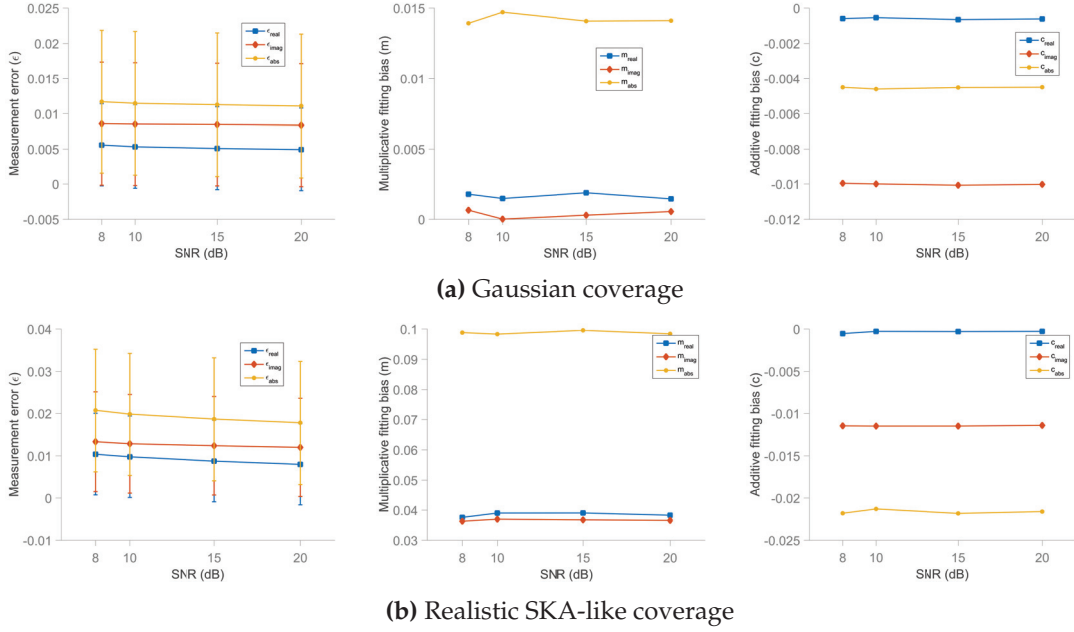
Figure 4.12 – Plots that present from left to right the measurement errors along with their variances, the multiplicative biases and the additive biases, as functions of the noise power in the model. The analysis is made using the **MR-LS** model, and: **(a)** a Gaussian type coverage, **(b)** a realistic SKA-like coverage. The values of the SNR used for the experiments were $8, 10, 15$ and 20 dB.

In the next stage, we repeat the same experiment after changing the coverage type. This time we use the realistic SKA-like one, and we test the influence of this change on the precision of the model. Again, the number of processed objects remains constant, and the signal-to-noise ratio is set to 20 dB. Figure 4.11b and Table 4.16 display the outcomes of this analysis. We can notice a small increase in the numbers, particularly in the multiplicative biases caused by larger measurement errors for objects with large true ellipticity.

In the final step, we study the robustness of the **MR-LS** approach to noise presence. More specifically, we test the same dataset with both coverage types in a series of experiments that we modify the noise power gradually. Then, we measure the same performance characteristics and check the importance of the noise level in the measurement accuracy. The signal-to-noise levels that we test are $8, 10, 15$, and 20. As we may observe in Figures 4.12a and 4.12b, the impact of the noise presence in this method is minimal. The numbers considering the errors and biases in the results have been almost unchanged, even when we had an SNR of 8 dB.

Finally, it is worth mentioning that the model described in this section is the practical realization of the theoretical **GLASSO best** of Section 4.5 since the implementation successfully locates the scale value that minimizes the relative residual in the fitting.

This fact can be confirmed by comparing the results from the two methods, where we see very similar numbers. The change can be explained by the different complexity of the data in the input between the two cases.

4.7   SARA-QMM solution

In the previous sections, we worked with models trying to make a precise shape measurement directly from the radio interferometric observations, applying a fitting procedure, and determining the missing parameter(s). Starting from this section on, we consider the shape estimation problem as a two-step task. In the first step, the intensity image is reconstructed using the visibilities and the coverage of the telescope. In the second one, the shape estimation is made from the outputs of the first part.

For this purpose, a good image reconstruction solution is necessary to get an intensity object with the minimum distortion in the shape characteristics. In our case, we adopted the SARA algorithm proved to provide a precise restoration of the intensity information, as explained in more detail in Sections 2.6 and 3.4.

Table 4.17 – Measurement errors and fitting biases, using the method **SARA-QMM** to multi-profile data with Gaussian-type coverage and SNR = 20 dB

Measure	$\square_{real}$	$\square_{imag}$	$\square_{abs}$
ϵ	$0.0039 \pm 7.9 \cdot 10^{-5}$	$0.0053 \pm 9.9 \cdot 10^{-5}$	$0.0073 \pm 1.7 \cdot 10^{-4}$
m	$-0.0092 \pm 3.9 \cdot 10^{-4}$	$-0.0256 \pm 7.2 \cdot 10^{-4}$	$-0.0204 \pm 1.4 \cdot 10^{-3}$
c	$-0.0003 \pm 2.5 \cdot 10^{-5}$	$0.0003 \pm 1.9 \cdot 10^{-5}$	$0.0010 \pm 2.9 \cdot 10^{-4}$

Table 4.18 – Measurement errors and fitting biases, using the method **SARA-QMM** to multi-profile data with realistic SKA-like coverage and SNR = 20 dB

Measure	$\square_{real}$	$\square_{imag}$	$\square_{abs}$
ϵ	$0.0027 \pm 1.9 \cdot 10^{-5}$	$0.0027 \pm 5.4 \cdot 10^{-5}$	$0.0042 \pm 6.9 \cdot 10^{-5}$
m	$-0.0028 \pm 1.8 \cdot 10^{-4}$	$-0.0095 \pm 2.2 \cdot 10^{-4}$	$-0.0071 \pm 5.6 \cdot 10^{-4}$
c	$0.0001 \pm 1.8 \cdot 10^{-5}$	$0.00007 \pm 2.1 \cdot 10^{-5}$	$0.0003 \pm 2.4 \cdot 10^{-5}$

The motivation for this development was a search for an algorithm that has practically no parameters to estimate or precisely calculate before proceeding to the main processing part. This framework includes a fully implemented model with all necessary parameters preset and ready to measure the shape of any input type. It is built so that it can work with a single object in any position of the image, or, unlike shapelet-based models, with the whole field of view containing several instances.

(a) Gaussian coverage

(b) Realistic SKA-like coverage

Figure 4.13 – Plots that present the difference between the measured and the true ellipticity as a function of the true values for the **SARA-QMM** model, where the intensity image is recovered using the SARA algorithm and the measurements are made in the output with the quadruple moments' technique. **(a)** The results when the coverage is Gaussian type. **(b)** The results when we use a realistic SKA-like coverage. In all cases, the difference of the measured and the true values is presented as a function of the true ones (blue dots) and is linearly fitted with the red line. From left to right, we show the cases for the real part of the ellipticity, the imaginary part of the ellipticity, and its complex magnitude. The studied dataset contains simulated images of multi-profiled objects with multiple entities in each of them.

Particularly for multi-object images, this solution gives a considerable advantage because the large majority of the processing time is dedicated to the intensity information restoration, and the reconstruction of more objects in the field of view decreases the processing time per object.

The experiments for this approach were based on the subset of images with multiple objects of various profiles. This subset contains 500 items of 20 entities per item, and the errors are measured per object. Table 4.17 introduces the measurement errors and the fitting biases, using this method with SNR = 20 dB. From this analysis, we have achieved a very accurate shape estimation of the 10000 objects that we studied, with fitting biases better than the ones of the **LS-simple**. Additionally, from Figure 4.13a, we deduce that this approach is not biased to larger ellipticity objects since the distribution of the errors is closer to uniform than in the other models. This confirms the small multiplicative bias produced in this experiment.

(a) Gaussian coverage

(b) Realistic SKA-like coverage

Figure 4.14 – Plots that present from left to right the measurement errors along with their variances, the multiplicative biases and the additive biases, as functions of the noise power in the model. The analysis is made using the **SARA-QMM** model, and: **(a)** a Gaussian type coverage, **(b)** a realistic SKA-like coverage. The values of the SNR used for the experiments were 8, 10, 15 and 20 dB.

Similarly to the studies in the previous sections, we repeat this experiment with only modification in the coverage used for the visibilities generation. As we can notice in Table 4.18, opposing to the results of **LS-simple**, **GLASSO increasing precision**, and **MR-LS**, this change helps in the improvement of the accuracy of this solution. Furthermore, the errors and biases have further decreased, denoting that the only factor influencing the model's efficiency is the size of the uv coverage and not its type.

The robustness to noise is also a question to help evaluate the performance of the method. For this reason, we designed a collection of experiments similar to the others presented in this section. We study again 500 images with 20 objects with various profiles per image, using both types of coverage. The parameter that changes in these experiments is the SNR level ranging between 8 and 20 dB. Figure 4.14a gives the corresponding results for Gaussian-type coverage, whereas Figure 4.14b shows these for the realistic SKA-like one. We can see that the quality of the measurement has been declined, although considering the already small numbers achieved for SNR = 20 dB, the reduction is not big.

Finally, we would like to examine the correlation of the measurement error with the size of the used coverage. This is the impulse to construct a new series of analyses

(a) Gaussian coverage

(b) Realistic SKA-like coverage

Figure 4.15 – Plots that demonstrate from left to right the measurement errors along with their variances, the multiplicative biases and the additive biases corresponding to the absolute magnitude of the ellipticity. The presentation is done as functions of the noise power in the model, for different coverage sizes. The analysis is made using the **SARA-QMM** model, and: **(a)** a Gaussian type coverage, **(b)** a realistic SKA-like coverage. The values of the SNR used for the experiments were 8, 10, 15 and 20 dB.

that repeat the above-described experiments using not only different coverage types but also sizes. We reform our experiments with the same 500 images and a ranging noise level (SNR = 8, 10, 15, 20 dB), modifying the collection of the uv points used for the visibilities generation. For this goal, we employ all sets introduced in Section 4.2, and we capture the outputs in Figure 4.15, where for clarity, only the numbers regarding the complex magnitude are considered. The performance boost with the increase of the visibility points is clear, describing the potential of **SARA-QMM** method for the new telescope generation that brings larger coverage sizes.

4.8 CLEAN-QMM solution

As explained in Section 3.5, this section's method focuses on assessing the importance of the SARA algorithm as a part of a shape measurement framework. Therefore, we present a solution similar to the previous section. The only difference is that we substitute the SARA reconstruction part with the alternative MS-CLEAN deconvolution algorithm.

Table 4.19 – Measurement errors and fitting biases, using the method **CLEAN-QMM** to multi-profile data with Gaussian-type coverage and SNR = 20 dB

Measure	$\square_{real}$	$\square_{imag}$	$\square_{abs}$
ϵ	$0.0346 \pm 3.7 \cdot 10^{-3}$	$0.0373 \pm 4.0 \cdot 10^{-3}$	$0.0566 \pm 7.1 \cdot 10^{-3}$
$\mathfrak{m}$	$-0.0399 \pm 5.8 \cdot 10^{-3}$	$-0.0159 \pm 6.4 \cdot 10^{-3}$	$-0.0660 \pm 9.1 \cdot 10^{-3}$
c	$0.0052 \pm 5.5 \cdot 10^{-3}$	$0.0019 \pm 4.8 \cdot 10^{-3}$	$0.0257 \pm 8.8 \cdot 10^{-3}$

In the beginning, the framework is applied to a set of 500 images containing multiple objects of various profiles. The tests include additive noise with SNR = 20 dB, and the corresponding visibilities are constructed using the Gaussian-type coverage. The outcomes of this experiment are shown in Table 4.19, where a significant increase in all indexes comparing to **SARA-QMM** is distinct. As we can observe in Figure 4.16a, the method gives larger errors for all ellipticity values. The difference in the number of points with large error and small ellipticity can be justified because these objects are significantly more in the dataset, as shown in Section 4.2.

Table 4.20 – Measurement errors and fitting biases, using the method **CLEAN-QMM** to multi-profile data with realistic SKA-like coverage and SNR = 20 dB

Measure	$\square_{real}$	$\square_{imag}$	$\square_{abs}$
ϵ	$0.0234 \pm 2.7 \cdot 10^{-3}$	$0.0233 \pm 2.6 \cdot 10^{-3}$	$0.0367 \pm 5.0 \cdot 10^{-3}$
$\mathfrak{m}$	$-0.0365 \pm 5.1 \cdot 10^{-3}$	$-0.0334 \pm 5.0 \cdot 10^{-3}$	$-0.0400 \pm 7.2 \cdot 10^{-3}$
c	$0.0021 \pm 6.3 \cdot 10^{-3}$	$0.0005 \pm 6.5 \cdot 10^{-3}$	$0.0109 \pm 7.9 \cdot 10^{-3}$

In the next step, we apply the same method to radio interferometric data generated using a realistic SKA-like coverage. The number of processed images remains 500 and the type of them, too. Additionally, the imposed noise creates signal-to-noise ratio of 20 dB. The numbers obtained from this analysis are given in Table 4.20, where we mark an improvement in the quality of the estimations and the respective linear fit biases resulting from the bigger coverage that provides more information for the image reconstruction. As depicted in Figure 4.16b, there is a better concentration of the points closer to the $x'x$ axis.

The next series of experiments aims to test the importance of the noise presence in the measurement quality using **CLEAN-QMM**. Once more, we make a sequence of analyses using the mentioned configurations and noise ranging between 8 and 20 dB. The study includes both the Gaussian-type and the realistic SKA-like coverage profiles. After working with 500 multi-object images, we got the results presented in Figure 4.17. We note that the method reveals robust characteristics regarding the added noise power and shows an increase in the indexes after moving below the limit of 10 dB.

(a) Gaussian coverage
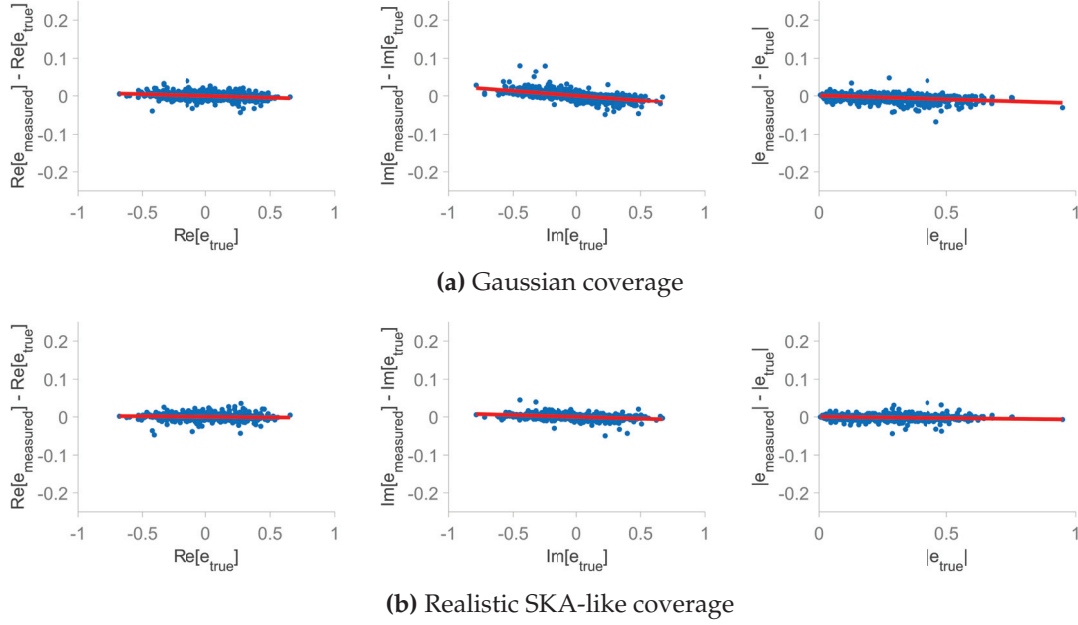

(b) Realistic SKA-like coverage

Figure 4.16 – Plots that present the difference between the measured and the true ellipticity as a function of the true values for the **CLEAN-QMM** model, where the intensity image is recovered using the MS-CLEAN deconvolution method, and the measurements are made in the output with the quadruple moments' technique. **(a)** The results when the coverage is Gaussian type. **(b)** The results when we use a realistic SKA-like coverage. In all cases, the difference of the measured and the true values is presented as a function of the true ones (blue dots) and is linearly fitted with the red line. From left to right, we show the cases for the real part of the ellipticity, the imaginary part of the ellipticity, and its complex magnitude. The studied dataset contains simulated images of multi-profiled objects with multiple entities in each of them.

The last question we attempt to answer is the influence of the coverage size on the shape estimation accuracy. A series of experiments was organized where, except for the change in the noise level (considered levels between 8 and 20 dB), we apply a different coverage size in each run. Therefore, we utilize the extra sets of uv points created for this scope, and we repeat the analysis of the 500 multi-object images with various profiles. Figure 4.18a captures the outputs of this study for the Gaussian coverage type, while Figure 4.18b is dedicated to the realistic SKA-like type. In both cases, for simplicity, only the outcomes related to the complex magnitude are presented. The results show a similar pattern to the experiment described before, where the method is robust to noise power up to 10 dB. Moreover, the increase in the number of uv points available for the analysis reduces the relevant errors and biases, independently of the coverage type used.

(a) Gaussian coverage

(b) Realistic SKA-like coverage
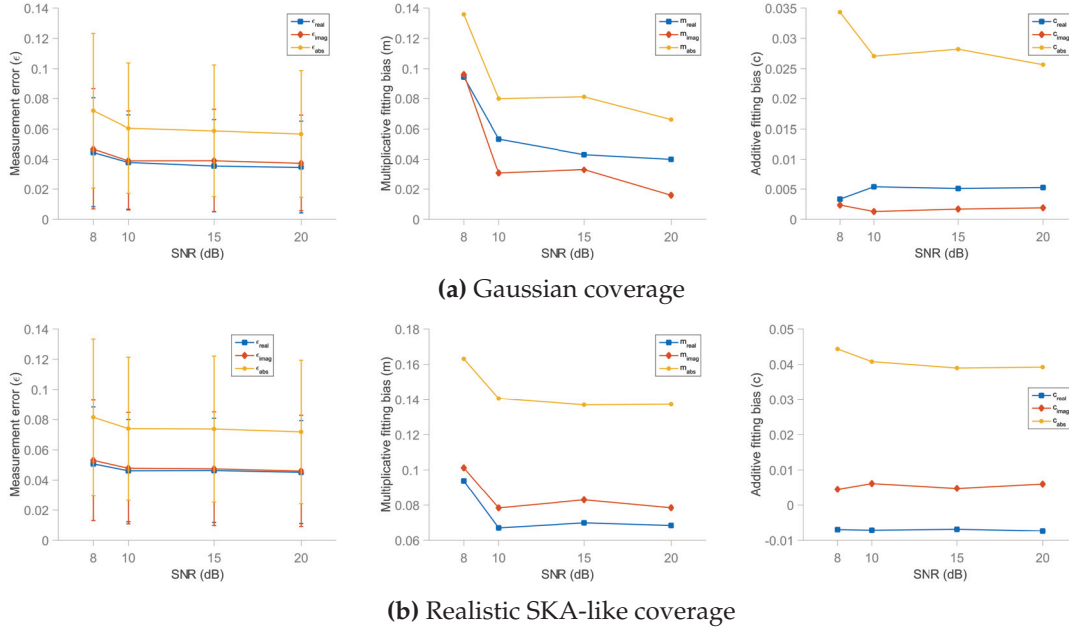
Figure 4.17 – Plots that present from left to right the measurement errors along with their variances, the multiplicative biases and the additive biases, as functions of the noise power in the model. The analysis is made using the **CLEAN-QMM** model, and: **(a)** a Gaussian type coverage, **(b)** a realistic SKA-like coverage. The values of the SNR used for the experiments were 8, 10, 15 and 20 dB.

4.9 Direct measurement from optical data

This section focuses on a model that does not consider the problem in the radio domain and makes the measurements directly from the optical data, whether these are simulated or real. Using the method, we do not aim to propose a new solution, alternative to the others presented in this chapter, but to use it as a tool for a deeper evaluation of these frameworks. More specifically, we would like to test the limits of the quadruple moments technique for shape measurement in optical data, which is a very commonly used solution among astronomers for this type of problem. Hence, we will be able to access the efficiency of this tool and the effect of it in the frameworks, like **SARA-QMM** and **CLEAN-QMM**, that use it as a part of their algorithm.

(a) Gaussian coverage
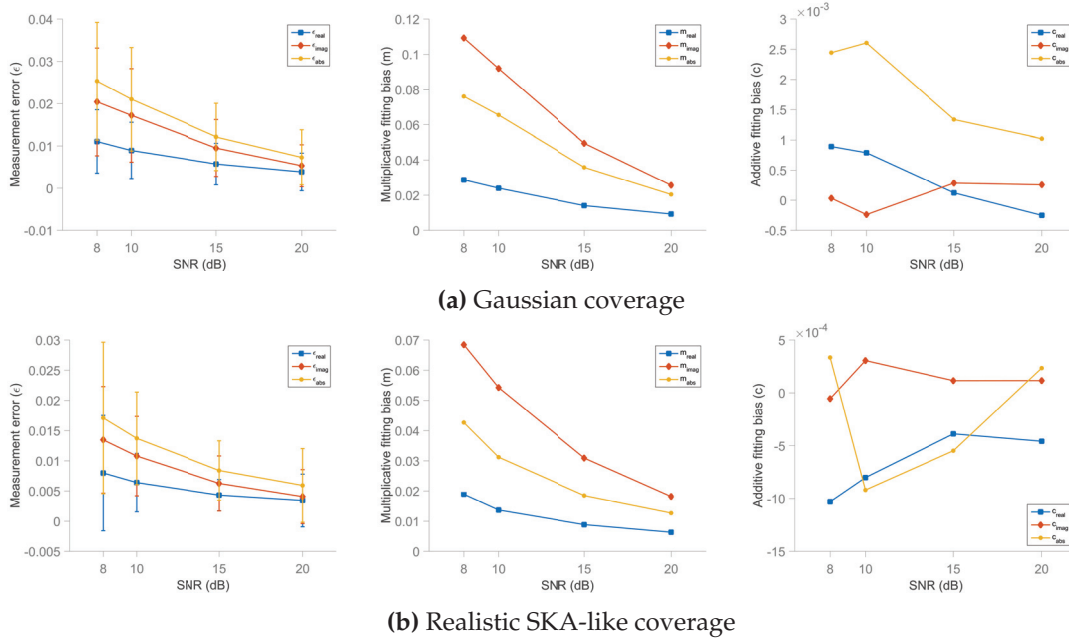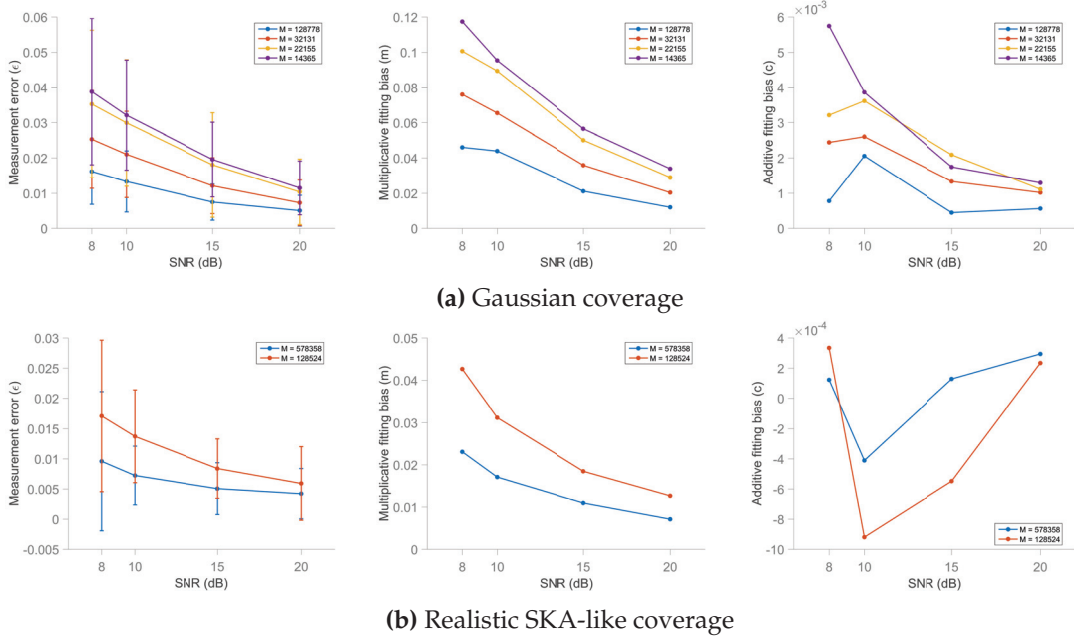

(b) Realistic SKA-like coverage
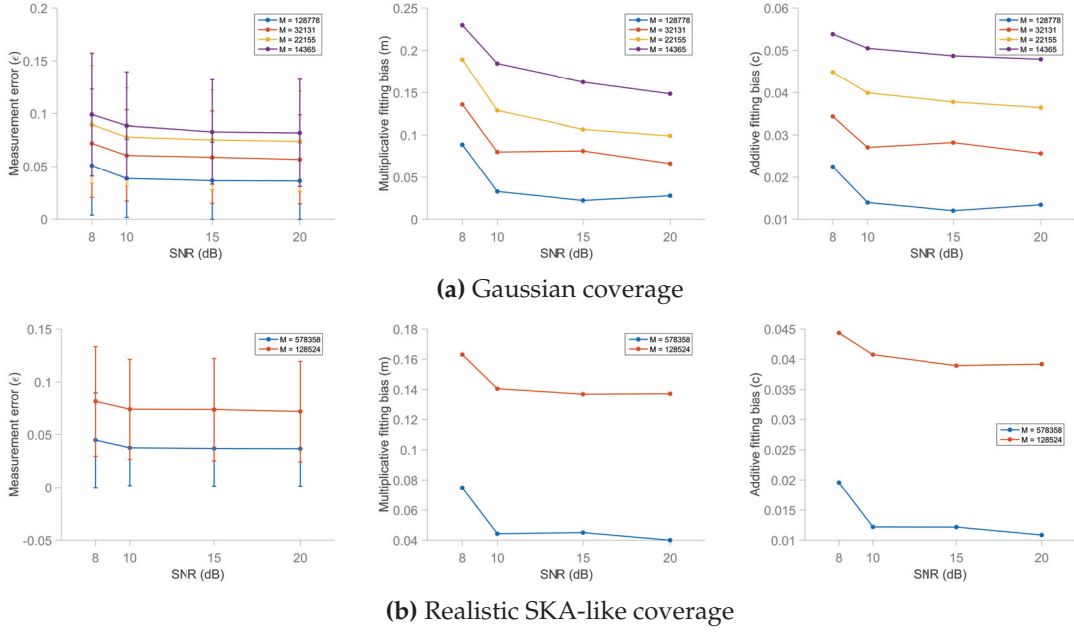
Figure 4.18 – Plots that demonstrate from left to right the measurement errors along with their variances, the multiplicative biases and the additive biases corresponding to the absolute magnitude of the ellipticity. The presentation is done as functions of the noise power in the model, for different coverage sizes. The analysis is made using the **CLEAN-QMM** model, and: **(a)** a Gaussian type coverage, **(b)** a realistic SKA-like coverage. The values of the SNR used for the experiments were 8, 10, 15 and 20 dB.

Table 4.21 – Measurement errors and fitting biases, using the quadruple moments for direct measurement to optical multi-profile data with SNR = 20 dB

Measure	$\square_{real}$	$\square_{imag}$	$\square_{abs}$
ϵ	$0.0035 \pm 2.3 \cdot 10^{-5}$	$0.0037 \pm 6.5 \cdot 10^{-5}$	$0.0056 \pm 8.1 \cdot 10^{-5}$
m	$-0.0014 \pm 1.2 \cdot 10^{-4}$	$-0.0041 \pm 1.9 \cdot 10^{-4}$	$-0.0052 \pm 2.5 \cdot 10^{-4}$
c	$0.00003 \pm 1.8 \cdot 10^{-5}$	$0.00004 \pm 2.1 \cdot 10^{-5}$	$0.0008 \pm 4.5 \cdot 10^{-5}$

With this in mind, we perform a direct measurement of the same 500 multi-object images with entities of various profiles. The noise intensity is always set at SNR = 20 dB. Table 4.21 and Figure 4.19 provide the processing results, where we notice the very small error and bias values. Additionally, through the plots, we observe a similar behavior of the direct optical measurement with the measurement using the **SARA-QMM** approach for the same dataset, a fact that can be verified by the error values got in the table.

In the next step, the study is concentrated on the impact of the noise level on the procedure's outcome. Thus, we rerun the above analysis with several SNR values to test the efficiency of the quadruple moments when the noise power rises. In
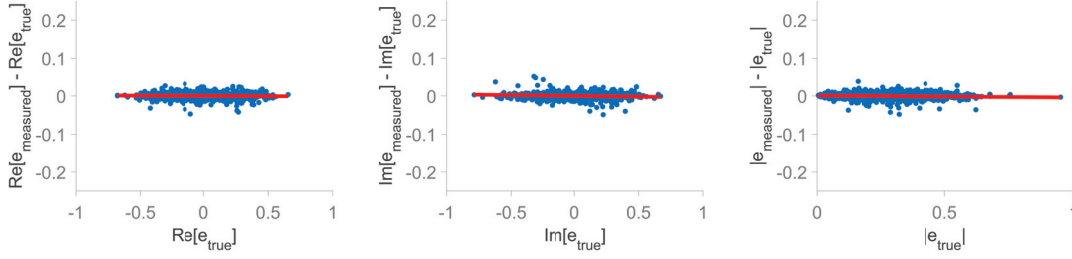
Figure 4.19 – Plots that present the difference between the measured and the true ellipticity as a function of the true values for the **Direct optical** model, where the measurements are done directly in the intensity image with the quadruple moments' technique. In all cases, the difference of the measured and the true values is presented as a function of the true ones (blue dots) and is linearly fitted with the red line. From left to right, we show the cases for the real part of the ellipticity, the imaginary part of the ellipticity, and its complex magnitude. The studied dataset contains simulated images of multi-profiled objects with multiple entities in each of them.
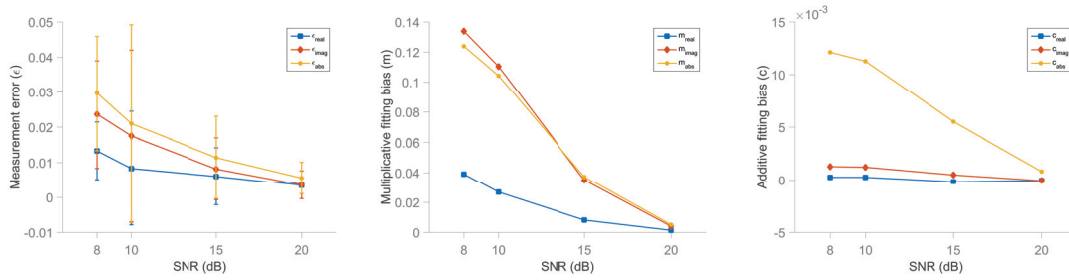
Figure 4.20 – Plots that present from left to right the measurement errors along with their variances, the multiplicative biases and the additive biases, as functions of the noise power in the model. The analysis is made using the **Direct optical** model, and the values of the SNR used for the experiments were $8, 10, 15$ and 20 dB.

Figure 4.20 where the results are depicted, we note a growth in the respective numbers when SNR reaches 8 dB that has a pattern similar to the one we met in Section 4.7, revealing a large correlation of the error of the **SARA-QMM** method with the performance of the quadruple moments technique.

4.10 Preliminary results from real data processing

After exploring the capabilities of all our models under different conditions and using several datasets and coverage types, we will investigate the potential of these frameworks on real data. For this purpose, we designed an environment that comes as close as possible to real radio interferometric data analysis.

At first, we adopt a set of real optical data coming from the COSMOS survey[†]. The Cosmic Evolution Survey (COSMOS) aims to study the galaxy evolution and
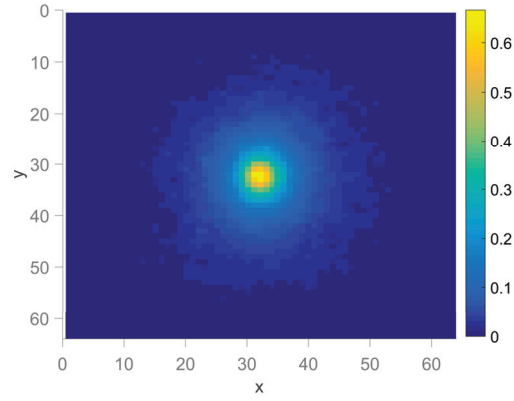
(a) A real object
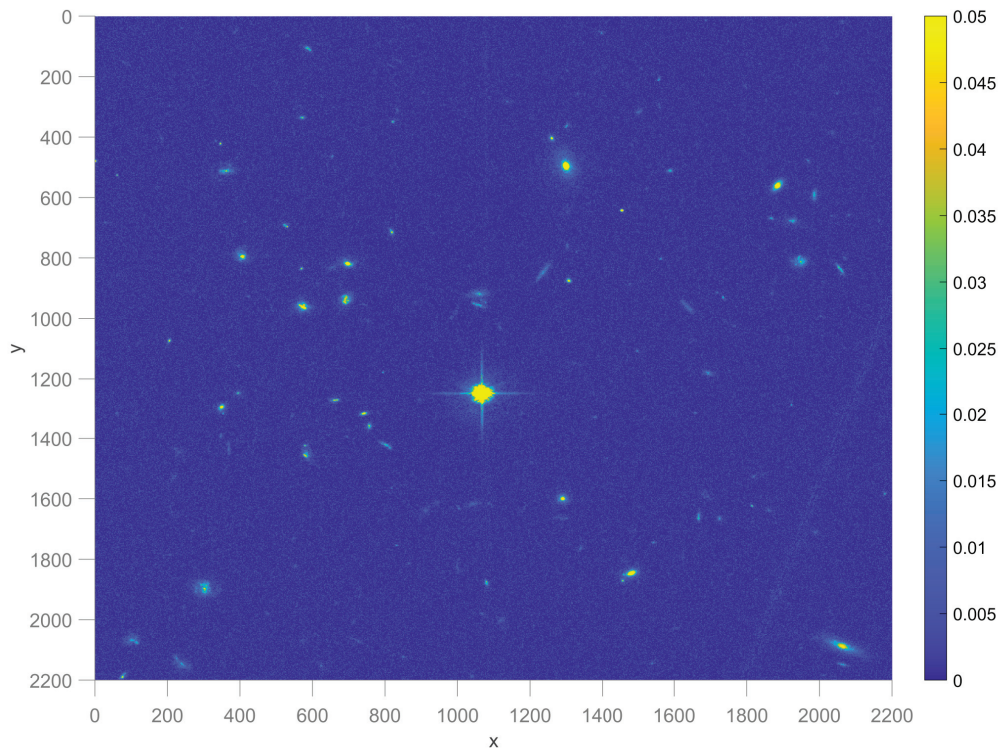
(b) A real part of the sky

Figure 4.21 – Examples of images originate from the COSMOS optical survey used for the experiment with real data, with: **(a)** an image with a real object, and **(b)** a part of the sky with multiple objects.

Table 4.22 – Measurement errors obtained in the experiments with real data from the COSMOS survey, studying all main models with realistic SKA-like coverage.

Model	ϵ_{real}	ϵ_{imag}	ϵ_{abs}
1	$0.0346 \pm 2.9 \cdot 10^{-3}$	$0.0299 \pm 1.4 \cdot 10^{-3}$	$0.0525 \pm 3.7 \cdot 10^{-3}$
7	$0.0375 \pm 8.4 \cdot 10^{-3}$	$0.0341 \pm 3.6 \cdot 10^{-3}$	$0.0598 \pm 9.8 \cdot 10^{-3}$
8	$0.0351 \pm 3.1 \cdot 10^{-3}$	$0.0302 \pm 1.5 \cdot 10^{-3}$	$0.0528 \pm 3.9 \cdot 10^{-3}$
9	$0.00007 \pm 1.6 \cdot 10^{-6}$	$0.00008 \pm 2.9 \cdot 10^{-6}$	$0.00011 \pm 4.5 \cdot 10^{-6}$
10	$0.0474 \pm 6.5 \cdot 10^{-3}$	$0.0389 \pm 3.9 \cdot 10^{-3}$	$0.0685 \pm 9.5 \cdot 10^{-3}$

1: LS-simple; 7: GLASSO increasing precision; 8: MR-LS; 9: SARA-QMM; 10: CLEAN-QMM.

Table 4.23 – Fitting biases obtained in the experiments with real data from the COSMOS survey, studying all main models with realistic SKA-like coverage.

Model	m_{real}	c_{real}	m_{imag}	c_{imag}	m_{abs}	c_{abs}
1	$-0.1410 \pm 9.2 \cdot 10^{-3}$	$0.0049 \pm 5.2 \cdot 10^{-4}$	$-0.0961 \pm 5.3 \cdot 10^{-3}$	$-0.0080 \pm 7.2 \cdot 10^{-4}$	$-0.1774 \pm 1.8 \cdot 10^{-2}$	$0.0286 \pm 2.4 \cdot 10^{-3}$
7	$0.1899 \pm 8.4 \cdot 10^{-3}$	$0.0158 \pm 2.0 \cdot 10^{-3}$	$0.0726 \pm 6.3 \cdot 10^{-3}$	$-0.0119 \pm 8.5 \cdot 10^{-4}$	$0.1489 \pm 2.4 \cdot 10^{-2}$	$0.0005 \pm 5.9 \cdot 10^{-4}$
8	$-0.1256 \pm 8.4 \cdot 10^{-3}$	$0.0052 \pm 4.9 \cdot 10^{-4}$	$-0.0989 \pm 5.8 \cdot 10^{-3}$	$-0.0077 \pm 6.3 \cdot 10^{-4}$	$-0.1655 \pm 1.1 \cdot 10^{-2}$	$0.0257 \pm 3.8 \cdot 10^{-3}$
9	$-0.0001 \pm 7.1 \cdot 10^{-5}$	$0.00002 \pm 1.1 \cdot 10^{-4}$	$-0.0001 \pm 8.9 \cdot 10^{-5}$	$0.00004 \pm 2.9 \cdot 10^{-4}$	$-0.00002 \pm 4.0 \cdot 10^{-5}$	$-0.00006 \pm 4.4 \cdot 10^{-4}$
10	$-0.0944 \pm 5.3 \cdot 10^{-3}$	$-0.0092 \pm 8.7 \cdot 10^{-4}$	$-0.0639 \pm 3.1 \cdot 10^{-3}$	$0.00008 \pm 1.2 \cdot 10^{-4}$	$-0.0571 \pm 2.8 \cdot 10^{-2}$	$0.0019 \pm 9.1 \cdot 10^{-4}$

1: LS-simple; 7: GLASSO increasing precision; 8: MR-LS; 9: SARA-QMM; 10: CLEAN-QMM.

formation in correlation with the local galaxy environment and the redshift. It covers approximately a 2 degree² equatorial field in multiple wavelengths. The data collection we utilize in this work comes from the HST Treasury Project centered on RA=10:00:28.6 and DEC=+02:12:21.0, using the ACS camera with resolution 0.03 arcsec/pixel (Koekemoer et al., 2007; Massey et al., 2010). In our experiments we used some of the images provided and extracted the objects for our measurements, excluding astronomical objects other than galaxies or cases with overlaps. Figure „„„4.21a captures an example of a galaxy from this new set with real objects, which contains 1200 items with sizes 64×64, 128×128 or 256×256. In addition, in Figure 4.21b, the reader can observe a sub-image from the original data form before extracting the objects.

In the second step, we use the direct measurement in the optical data approach to measure the objects separately from this dataset, recording the results. These numbers are used in the performance evaluation of the other models as the ground truth shape information in the absence of such knowledge. Considering the efficiency of the quadruple moment technique in achieving a good shape measurement, as presented in Section 4.9, this is a close alternative to the unknown true ellipticity values.

In the final stage, we assume the realistic SKA-like coverage to make the generation of the visibilities. This part will make the experiment one step more natural because it will produce radio interferometric data from the observation of real objects in the

---

†All COSMOS data can be found in the public COSMOS Archive: irsa.ipac.caltech.edu/data/ COSMOS/overview.html

(a) LS-simple
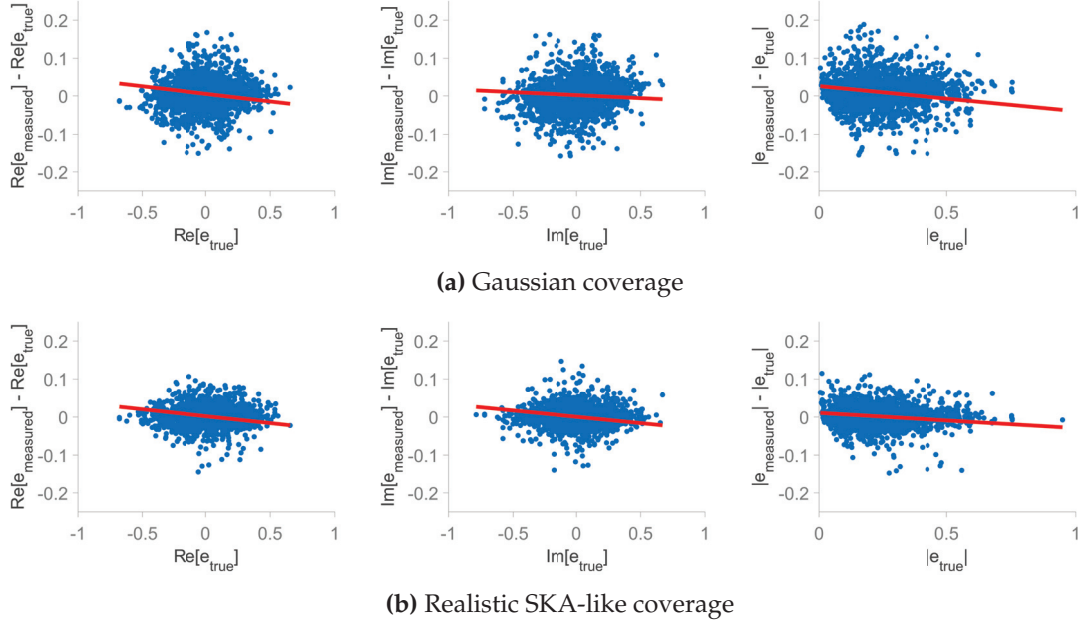

(b) GLASSO increasing precision

(c) MR-LS
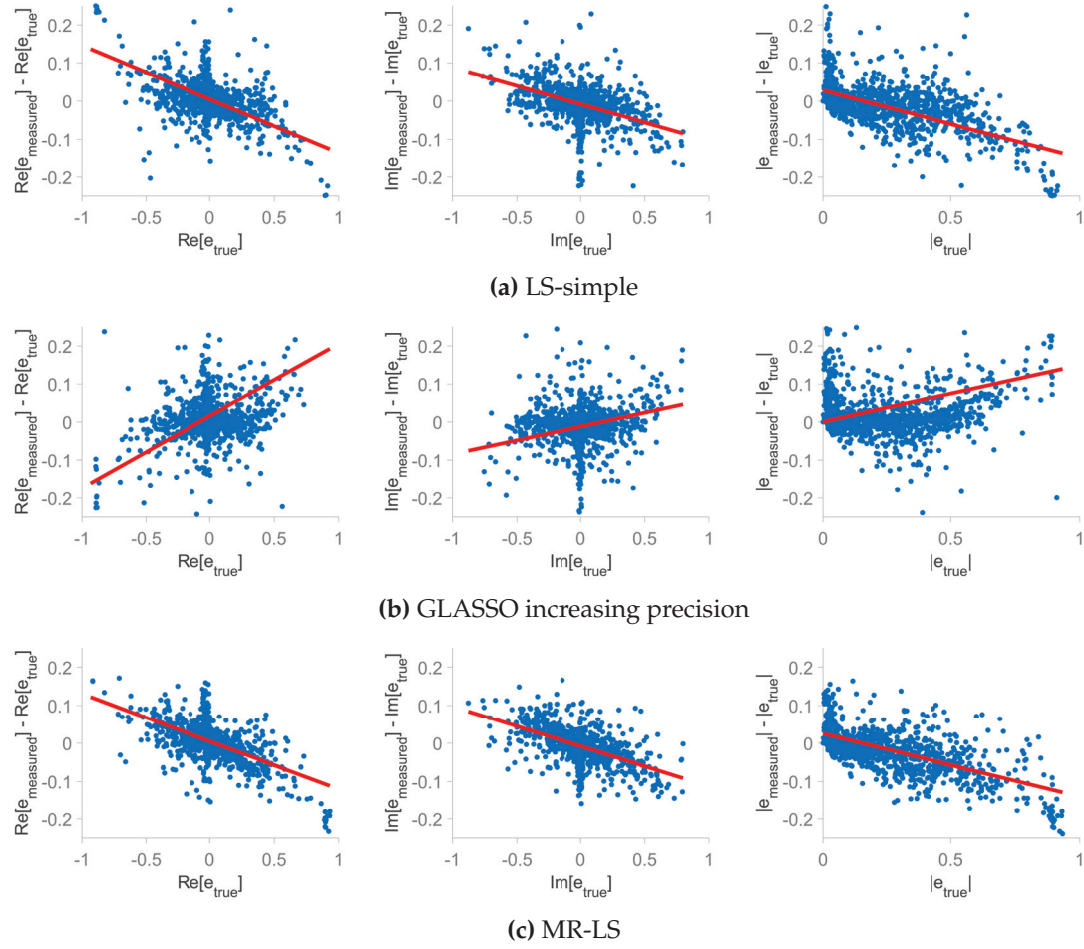
Figure 4.22 – Plots that present the difference between the measured ellipticity and the true ellipticity as a function of the true values when the presented models are applied to the data originating from the COSMOS optical survey with a realistic SKA-like coverage, where: **(a) LS-simple** - the scale parameter approximation is made using the equations 2.21 and 2.22. **(b) GLASSO increasing precision** - β is chosen initially in $[0.5, 20.5]$ and, then, iteratively in dictionary of increasing precision in the range of the optimal value of the previous step. **(c) MR-LS** - the scale parameter is chosen using grid search that minimizes the relative residual, initially in $[0.5, 20.5]$ and, then, iteratively in intervals of increased precision in the range of the value selected at the previous step. In all cases, the true values for the evaluation come from the direct measurement of the same data in the optical images with the quadruple moments. The difference of the measured and the simulated true values is presented as a function of the second ones (blue dots) and is linearly fitted with the red line. From left to right, we show the cases for the real part of the ellipticity, the imaginary part of the ellipticity, and its complex magnitude.

(a) SARA-QMM

(b) CLEAN-QMM

Figure 4.23 – Plots that present the difference between the measured ellipticity and the true ellipticity as a function of the true values when the presented models are applied to the data originating from the COSMOS optical survey with a realistic SKA-like coverage, where: **(a) SARA-QMM** - the intensity image is recovered using the SARA algorithm, and the measurements are made in the output with the quadruple moments' technique. **(b) CLEAN-QMM** - the intensity image is recovered using the MS-CLEAN deconvolution method, and the measurements are made in the output with the quadruple moments. In all cases, the true values for the evaluation come from the direct measurement of the same data in the optical images with the quadruple moments. The difference of the measured and the simulated true values is presented as a function of the second ones (blue dots) and is linearly fitted with the red line. From left to right, we show the cases for the real part of the ellipticity, the imaginary part of the ellipticity, and its complex magnitude.

field of view using a real radio telescope configuration.

After making the set-up for our experiments, we proceed to analyze the 1200 images in the dataset using all proposed methods. This study does not include extra noise in the visibilities because the original data are already noisy. Tables 4.22 and 4.23 give an overall view of the measurement errors and the respective biases after the linear fitting for all models. Furthermore, Figures 4.22 and 4.23 depict the difference between the measured and the true ellipticity as a function of the true values for the models based on fitting using a shapelet dictionary, and these that restore the sky brightness information, respectively.

As one may see, the estimations in all cases are adequately satisfying, acknowledging the complexity of the data. We also note the similar numbers given by the shapelet-fitting methods, in particular, these of **LS-simple** and **MR-LS**. Both

present comparable abilities regarding both the resulting errors and biases, with the only difference that the latter introduces an automatic way for the measurements. We also mark the extremely accurate estimations accomplished by the **SARA-QMM** approach, which offers precise ellipticity values. This fact confirms one of the conclusions in Section 4.9 about the close dependence of the precision of the aforementioned method on the accuracy of the quadruple moment measurement. Practically, the substantial majority of the errors of this model come from the errors brought in by the moment technique. However, this is not true for the **CLEAN-QMM** solution, which in most cases is inferior to **SARA-QMM** by 3 orders of magnitude, which proves the superiority of the second in the accurate solution of this problem.

4.11   Overall evaluation and discussion

Considering the experimental observations in this chapter, we may conclude that all the proposed algorithms produced satisfactory results. The measurements in the significant majority of the cases were made with high accuracy and precision.

The models built with the shapelet-fitting logic tend to produce larger errors and measurements that diverge from the true values slightly more than the other cases when the tested objects have large ellipticity. The most important stage in these algorithms is the selection of the shapelet parameters' values, where, in particular, the scale parameter determines the precision of the ellipticity estimation. Additionally, these methods face greater difficulty in making accurate measurements when the telescope coverage used is a realistic SKA-like one.

The **LS-simple** approach is a relatively straightforward version of the shapelet-fitting framework, being considerably fast in the processing with an average rate of 0.6 seconds per image. However, the quality of the shape estimations is not reduced by this simplicity. The equations used for the approximation of the shapelet parameters in the model is a convenient initial guess, but a supplementary study must be done to increase the method's efficiency by increasing the accuracy of the estimated β parameter. Moreover, regarding the n_{max} parameter, the experiments proved that a considerable reduction in the method's complexity could be achieved by using the minimum, according to the specification, acceptable value without changing the quality of the offered estimations. Finally, this model has been demonstrated to be very resistant to the introduction of large noise, giving outputs of similar precision to those with considerably higher SNR. Here, we have to mention that the estimation equations applied in this approach use information from the intensity images, which, in general, is not possible during the processing of interferometric data. In that case, it would require the reconstruction of the intensity data from

the radio ones, and the measurements directly from the radio data would not be necessary.

The **GLASSO regularization** variants were developed aiming to find an approach that approximates one of the three basic parameters of the shapelet fitting idea without manual guidance or any prior knowledge on the data characteristics. The variant using a constant dictionary to process the whole dataset is purely automatic, although it lacks precision. To work with objects of various sizes, we must employ a wide range of scale values to cover all object possible sizes, while this prevents the use of high precision in β values. When the scale parameter is not approximated with high precision, we produce a less accurate measurement. To cover all cases and have high precision simultaneously, a substantially large dictionary with millions of columns would be required, an analysis that is not computationally supported by most of the machines in this period. A more realistic plan is constructing an individually formed dictionary per image due to the reduction in the required computational resources. An issue in this approach is the tendency of the sparse regularization to converge to values near the edges of our selected interval, indicating that the optimal value is out of it. In this context, an alternative way suggests working in multiple precision levels, reprocessing the images several times after starting from a low-precision dictionary. Each time we use a dictionary with scale values selected in the range of the β value chosen at the previous iteration. This is a purely automatic approach that avoids at the same time the issue of converging in values at the edges of our selection. According to our experiments, the latter variant gave estimations with high accuracy regarding the produced results, comparable to those of the **LS-simple**. Additionally, an important disadvantage of the GLASSO regularization is its computational complexity with an average processing time of almost 60 seconds per image, which is approximately a hundred times[‡] slower than the case with **LS-simple**. This issue could be tackled when the measurements are done in large computer clusters. Moreover, the obtained results show that an improvement in fitting is possible if a higher-order polynomial is used for this purpose to fit the produced data (Fig. 4.9) instead of the linear one (Fig. 4.7), with a significant increase in the goodness measure.

Two more tested issues are the large increase in the noise of our data and the miscalculation of the object center. For the first one, we observed that the proposed methods are very robust to noise up to the level of 8 dB. On the other hand, if the center is not localized with very high precision, the measurement error can be highly increased. Therefore, we conclude that a point that has to be improved to make this model more independent of any prior knowledge about the characteristics of the objects for processing is the precise calculation of each center before making the

---

[‡]The analysis for this work was conducted in MS Windows 10 - 64-bit system, CPU - Intel Core i7-4720HQ @ 2.60 GHz, RAM - 16 GB.

fitting. Furthermore, we experimented with the potential of sparse regularization in comparison with the simple least-squares framework, and we observed that it might be a purely automatic method, but it slightly lacks the precision of the baseline model. A general remark that concerns both the basic and the GLASSO models is the tendency for small objects to produce larger measurement errors because their corresponding size is close to or smaller than the pixel size, in which our model resolution is limited. Finally, it is worth considering that this approach has difficulties in processing when we apply a realistic SKA-like coverage to generate the visibilities. This modification complicates the problem of advanced regularization that we attempt to solve using group sparsity due to the many data points in the uv coverage that have considerably large variance in their pair value difference. This makes the estimation of the group LASSO parameters very crucial for the appropriate convergence of the algorithm. We approximate quite successfully the regularization parameter using an estimation form, but the gradient step needs to be more accurately set individually per image to help the precision of the model. A more sophisticated approach in this direction would benefit the initial goal.

The minimization of the relative residual of the fitting to the shapelet dictionary was an interesting concept that worked well in the **GLASSO regularization** model variants. However, the complexity of the advanced minimization added an extra burden to the procedure. This was the motivation to propose the **MR-LS** model that follows the same main principle, approaching the problem more straightforwardly. This solution dropped the processing time from almost 60 seconds per image of the previous model to approximately 2.4 seconds per image. Additionally, we have guaranteed convergence to the optimal value without the complication of the estimation of supplementary parameters introduced in the model. On top of that, the measurement accuracy of the framework has surpassed these of the **GLASSO regularization** variants, with all performance measures being smaller and close to the **LS-simple** method that assumes the knowledge on data characteristics. This confirms the study of Section 4.5, which states that a precise localization of the scale value that minimizes the relative residual will give high precision measurements. The drawbacks of the approach are the common ones for all shapelet-fitting-based methods. The object must be well centered before making the decomposition to maximize the precision in the shape estimation. Moreover, objects with larger ellipticity tend to be measured with less accuracy, a characteristic that is suspected to be induced due to the shapelet dictionary and its difficulty to fit the data exactly. Finally, the efficiency of this choice in processing data containing noise with a higher power was remarkable, keeping the errors and biases almost untouched after the increase of the noise level.

In the other proposed processing direction, we analyze the data in a two-step algorithm. The first step focuses on the intensity information restoration from the

(a) LS-simple

(b) GLASSO increasing precision
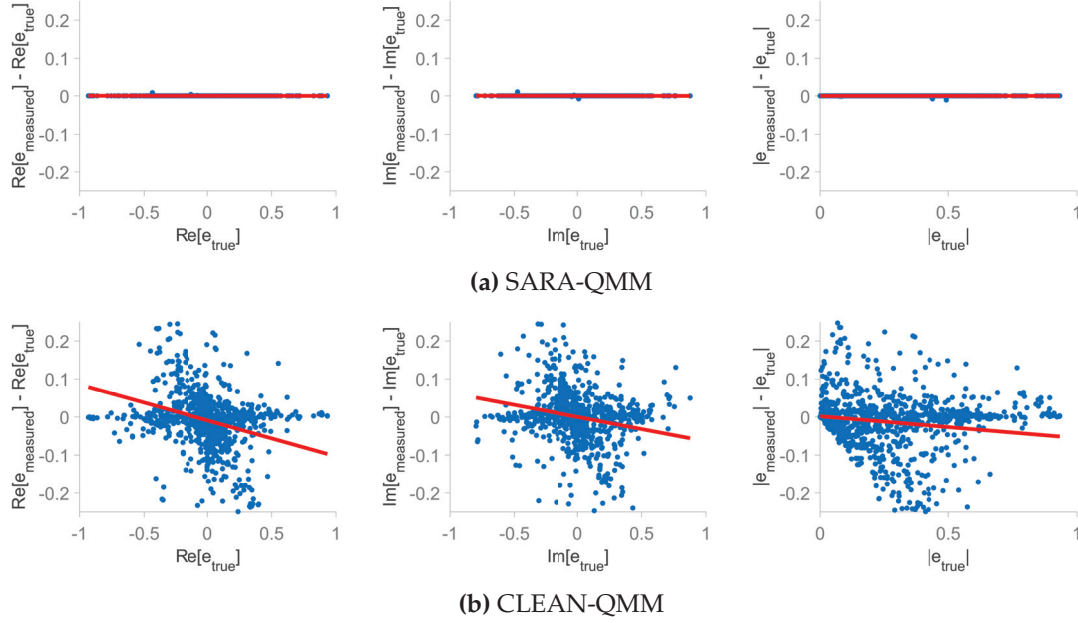
(c) MR-LS
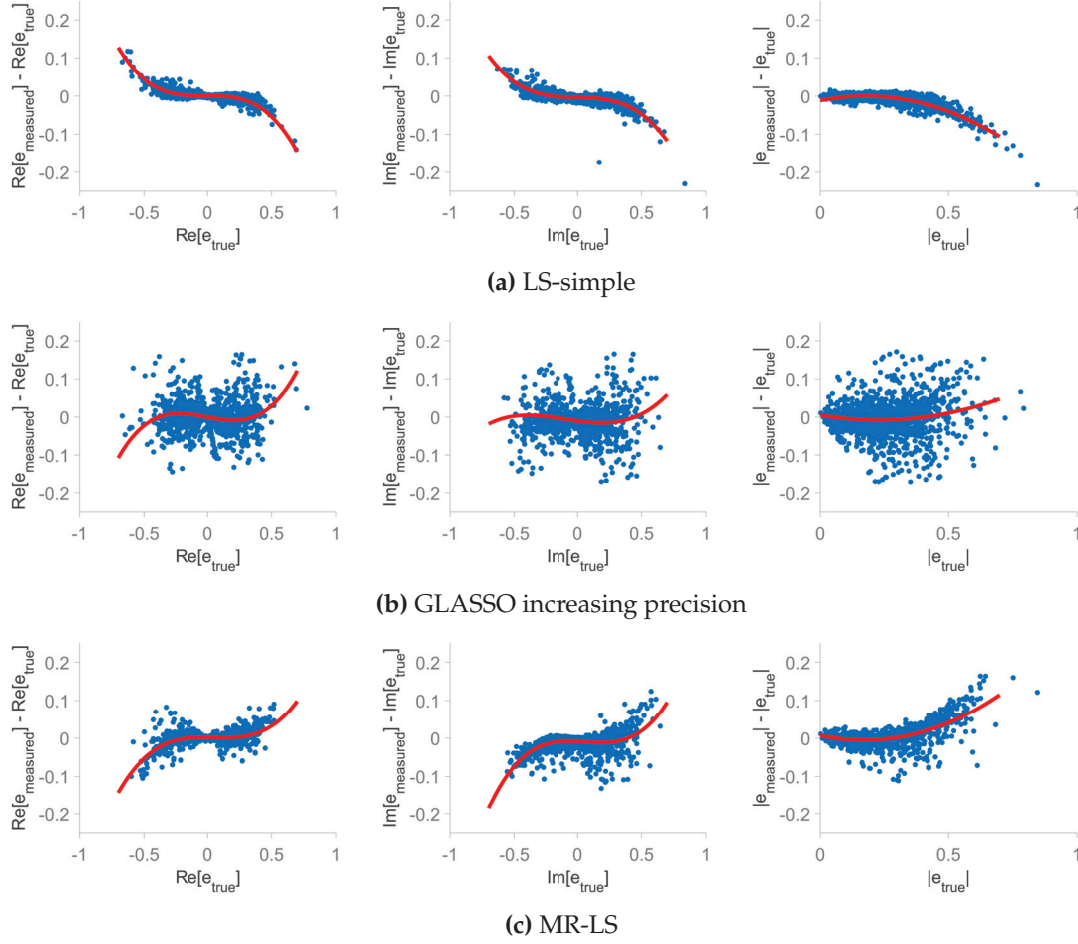
Figure 4.24 – Plots that present the difference between the measured ellipticity and the true ellipticity as a function of the true values when the presented models are applied to multi-profiled simulated data with a realistic SKA-like coverage, where: **(a) LS-simple** - the scale parameter approximation is made using the equations 2.21 and 2.22. **(b) GLASSO increasing precision** - β is chosen initially in $[0.5, 20.5]$ and, then, iteratively in dictionary of increasing precision in the range of the optimal value of the previous step. **(c) MR-LS** - the scale parameter is chosen using grid search that minimizes the relative residual, initially in $[0.5, 20.5]$ and, then, iteratively in intervals of increased precision in the range of the value selected at the previous step. In all cases, the difference of the measured and the true values is presented as a function of the true ones (blue dots) and is fitted with the red line using higher-order polynomials. From left to right, we show the cases for the real part of the ellipticity, the imaginary part of the ellipticity, and its complex magnitude.

(a) SARA-QMM

(b) CLEAN-QMM

(c) Direct optical

Figure 4.25 – Plots that present the difference between the measured ellipticity and the true ellipticity as a function of the true values when the presented models are applied to multi-profiled simulated data with a realistic SKA-like coverage, where: **(a) SARA-QMM** - the intensity image is recovered using the SARA algorithm, and the measurements are made in the output with the quadruple moments' technique. **(b) CLEAN-QMM** - the intensity image is recovered using the MS-CLEAN deconvolution method, and the measurements are made in the output with the quadruple moments. **(c) Direct optical** - the measurements are done directly in the intensity image using the quadruple moments. In all cases, the difference of the measured and the true values is presented as a function of the true ones (blue dots) and is fitted with the red line using higher-order polynomials. From left to right, we show the cases for the real part of the ellipticity, the imaginary part of the ellipticity, and its complex magnitude.

interferometric data, while in the second one, we measure the shape of the objects in the outputs of the first stage. For this purpose, we have presented two solutions that differ in how the sky's brightness is reconstructed. One of them is implemented with the SARA algorithm (**SARA-QMM**) that employs the compressed sensing approach and complex optimization tools to retrieve the intensity images. Then, the shape is measured from the reconstructed data using the quadruple moments. This framework has the advantage coming initially from its design, where all parameters are preset, and the analysis is made for any object and any coverage pattern used without further adaptation steps. Besides that, the model has the benefit of the ability to study multiple objects in the field of view, being able to reconstruct the brightness information independently of the entities presented in the image, and isolate objects for measurement utilizing an embedded object extractor. This can speed up the processing time by a large factor when the provided field of view is denser. The resulting mean execution time for an image of 256×256 is approximately 42 seconds per image. Although assuming that our images contain 20 entities each, this time can be translated to approximately 2.1 seconds per object, which makes its performance very competitive compared to the previously mentioned cases. In terms of measurement accuracy **SARA-QMM** gives minimal error values and linear fitting biases and outperforms all previously discussed models when compared to the same dataset and coverage. In particular, for real data, the algorithm produces measurements with precision in the order of magnitude of 10^{-5}. Additionally, the method's robustness in the presence of intense noise is very satisfactory, while the increase in the size of the coverage can offer a supplementary boost in the framework's efficiency. From our tests, the latter is proved to be the only factor that affects the performance of the measurement, unlike the other methods. This can be considered as an additional advantage of the proposal assuming the capabilities of modern radio interferometric arrays.

The alternative solution (**CLEAN-QMM**) that follows the image reconstruction path is CLEAN-based and has a similar algorithm with **SARA-QMM**, differing only in the reconstruction stage. It can be considered state-of-the-art in the domain, although it has not been used as a whole framework for the specific problem. The MS-CLEAN deconvolution algorithm is very commonly used among scientists in the area to restore the intensity information from the visibilities. In our study, we want to compare the efficiency of a model including MS-CLEAN to the results we accomplished using the **SARA-QMM** approach. Regarding the computational complexity of the model, this is significantly faster than the one of the previous paragraph, performing the measurement with a mean processing time ≈ 14.5 seconds per image that corresponds to approximately 0.75 seconds per object. However, the method is proved to give a much lower efficiency level in the main target, which is the measurement task. The estimation errors and the relevant biases from the linear fitting are significantly inferior in quality than those of **SARA-QMM** but in

Table 4.24 – Resulting values $1 - R^2$ for the measurement of the badness of the fitting in the main models, using a realistic SKA-coverage in multi-profiled simulated data, for both linear and higher polynomial fitting.

Model	Linear fitting			Higher order polynomial fitting		
	Real part	Imaginary part	Complex magnitude	Real part	Imaginary part	Complex magnitude
1	$34.1 \cdot 10^{-4}$	$30.2 \cdot 10^{-4}$	$128.7 \cdot 10^{-4}$	$14.8 \cdot 10^{-4}$	$18.1 \cdot 10^{-4}$	$66.8 \cdot 10^{-4}$
7	$544.5 \cdot 10^{-4}$	$527.9 \cdot 10^{-4}$	$1410.8 \cdot 10^{-4}$	$522.5 \cdot 10^{-4}$	$518.4 \cdot 10^{-4}$	$1208.2 \cdot 10^{-4}$
8	$89.9 \cdot 10^{-4}$	$137.4 \cdot 10^{-4}$	$380.1 \cdot 10^{-4}$	$74.6 \cdot 10^{-4}$	$116.4 \cdot 10^{-4}$	$330.0 \cdot 10^{-4}$
9	$6.9 \cdot 10^{-4}$	$12.9 \cdot 10^{-4}$	$33.8 \cdot 10^{-4}$	$6.8 \cdot 10^{-4}$	$12.9 \cdot 10^{-4}$	$33.7 \cdot 10^{-4}$
10	$812.5 \cdot 10^{-4}$	$707.9 \cdot 10^{-4}$	$918.3 \cdot 10^{-4}$	$808.7 \cdot 10^{-4}$	$706.2 \cdot 10^{-4}$	$917.1 \cdot 10^{-4}$
11	$9.0 \cdot 10^{-4}$	$18.0 \cdot 10^{-4}$	$43.2 \cdot 10^{-4}$	$8.9 \cdot 10^{-4}$	$18.0 \cdot 10^{-4}$	$43.1 \cdot 10^{-4}$

1: LS-simple ; 7: GLASSO increasing precision ; 8: MR-LS ; 9: SARA-QMM ; 10: CLEAN-QMM ; 11: Direct optical.

the same order of magnitude with the shapelet-fitting models, a fact that is clearer when we process real data. Furthermore, **CLEAN-QMM** has a good performance regarding its strength to adapt to data with higher noise levels, whereas the output improves when we provide a larger coverage, a point that is common to both reconstruction-based solutions.

The last experiments we presented were dedicated to the direct measurement of the objects' shape in the optical data. Motivated to analyze the impact of the quadruple moments' technique on the errors produced by the two models using it, we measured the tested objects precisely. The estimation deviations produced were very close to these came from the **SARA-QMM** framework, while the efficiency of the direct optical measurement in the presence of noise approximates this of the same method. These conclusions were confirmed during the measurement of real data with the **SARA-QMM** algorithm, where we used the direct optical measurement information as the ground truth for the performance analysis, and we achieved substantially low numbers for the corresponding indexes.

As a final evaluation step, we worked on fitting the results using all main models with multi-profile objects and a realistic SKA-like coverage. We substituted the linear fitting of the above cases with a higher-polynomial one, similarly to the study included in Section 4.5 for the methods built on **GLASSO regularization**. Figures 4.24, and 4.25 capture the outcomes of the procedure, while Table 4.24 records the badness of the fitting using the index of equation 4.12 for both linear and higher-polynomial cases.

(a) Gaussian coverage

(b) Realistic SKA-like coverage
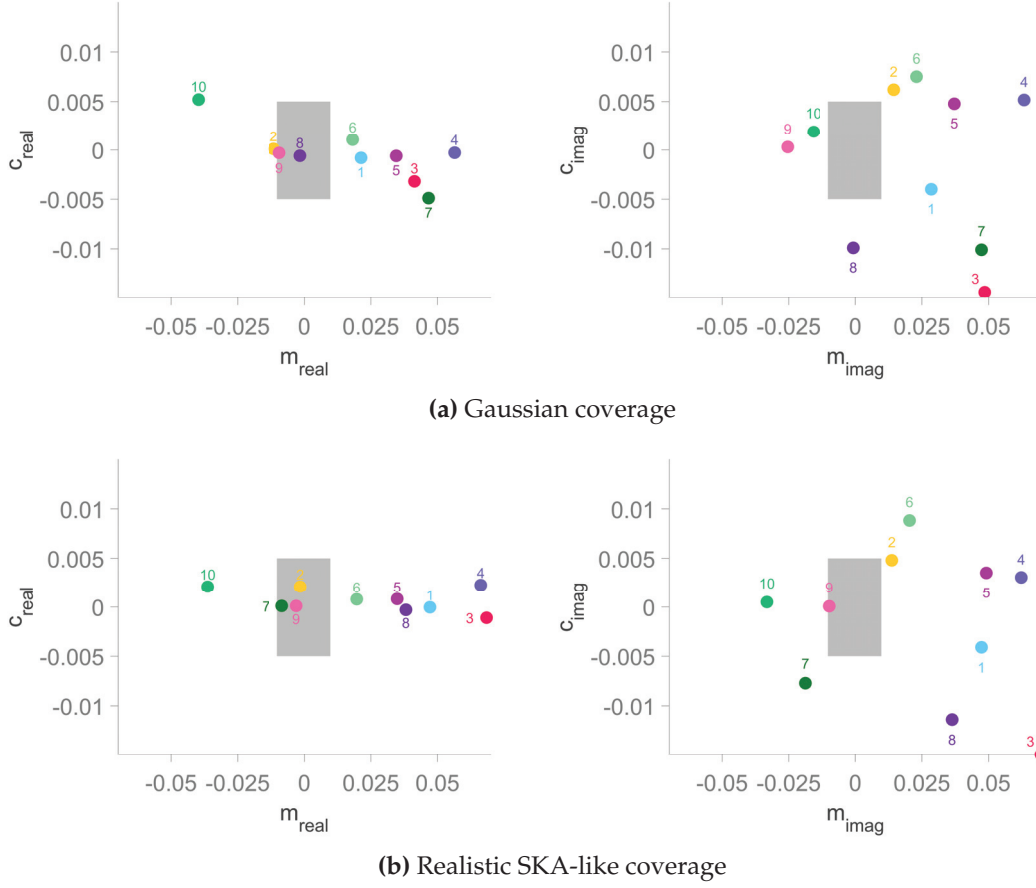
1: LS-simple;   2: LS $n_{max} = 4$;   3: LS noise;   4: GLASSO automatic;   5: GLASSO initial guess;

6: GLASSO mixed with LS;   7: GLASSO increasing precision;   8: MR-LS;   9: SARA-QMM;   10: CLEAN-QMM.

Figure 4.26 – Plots depicting the resulting multiplicative and additive biases after our experiments for all the implemented models and their variants with multi-profiled data, when fitting the measurement results to the true values for the real (left) and the imaginary (right) parts. We present the numbers after working on different coverage types, with: **(a)** Gaussian profile, and **(b)** realistic SKA-like profile. Following similar notation as in Heymans et al., 2006, the gray area represents the ideal zone that the biases should be located, for a model to be considered as very reliable, with $|m| \leqslant 10^{-2}$ and $|c| \leqslant 5 \cdot 10^{-3}$.

As we can observe, the new fitting option has improved the results in the shapelet-based frameworks, while those of the reconstruction-based ones are practically unchanged. This is expected because of the above-explained sensitivity of the shapelet fitting to objects with larger ellipticity, which leads to a pattern of errors closer to higher polynomials than to a linear model.

(a) Processing time and fitting quality correlation

(b) Measurement error and fitting quality correlation

(c) Processing time and measurement error correlation
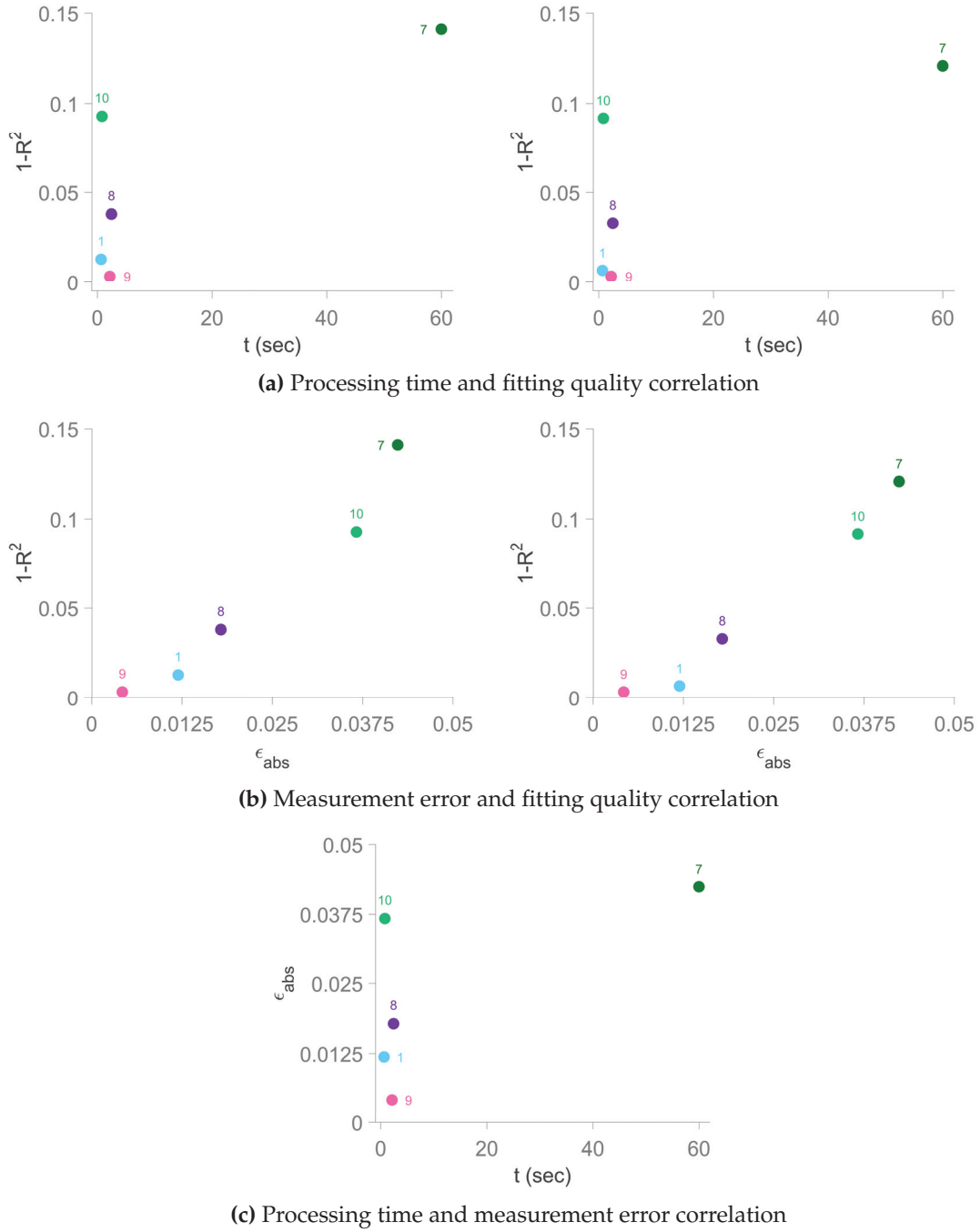
1: LS-simple ; 7: GLASSO increasing precision ; 8: MR-LS ; 9: SARA-QMM ; 10: CLEAN-QMM.

Figure 4.27 – Plots depicting the correlation among several measured quantities after our experiments for all the main implemented models with multi-profiled data and a realistic SKA-like coverage, when we perform linear (left) and polynomial (right) fitting of the measurement results to the true values. The pair correlations explored are: **(a)** processing time and fitting quality correlation, **(b)** measurement error and fitting quality correlation, and **(c)** processing time and measurement error correlation. The processing time in each case is expressed in seconds, and the measurement error concerns the complex magnitude.

Table 4.25 – Summary of the measurement errors as a result of studying all main models, using multi-profiled simulated data and a realistic SKA-like coverage§with SNR = 20 dB.

Model	ϵ_{real}	ϵ_{imag}	ϵ_{abs}
1	$0.0068 \pm 1.7 \cdot 10^{-4}$	$0.0083 \pm 1.5 \cdot 10^{-4}$	$0.0119 \pm 2.99 \cdot 10^{-4}$
7	$0.0255 \pm 2.5 \cdot 10^{-3}$	$0.0261 \pm 2.4 \cdot 10^{-3}$	$0.0424 \pm 4.4 \cdot 10^{-3}$
8	$0.0079 \pm 3.7 \cdot 10^{-4}$	$0.0120 \pm 5.4 \cdot 10^{-4}$	$0.0178 \pm 8.5 \cdot 10^{-4}$
9	$0.0027 \pm 1.9 \cdot 10^{-5}$	$0.0027 \pm 5.4 \cdot 10^{-5}$	$0.0042 \pm 6.9 \cdot 10^{-5}$
10	$0.0234 \pm 2.7 \cdot 10^{-3}$	$0.0233 \pm 2.6 \cdot 10^{-3}$	$0.0367 \pm 5.0 \cdot 10^{-3}$
11	$0.0035 \pm 2.3 \cdot 10^{-5}$	$0.0037 \pm 6.5 \cdot 10^{-5}$	$0.0056 \pm 8.1 \cdot 10^{-5}$

1: LS-simple; 7: GLASSO increasing precision; 8: MR-LS; 9: SARA-QMM; 10: CLEAN-QMM;

11: Direct optical.

Table 4.26 – Summary of the multiplicative fitting biases as a result of studying all main models, using multi-profiled simulated data and a realistic SKA-like coverage with SNR = 20 dB.

Model	m_{real}	m_{imag}	m_{abs}
1	$0.0473 \pm 2.7 \cdot 10^{-3}$	$0.0475 \pm 4.1 \cdot 10^{-3}$	$0.0912 \pm 5.4 \cdot 10^{-3}$
7	$-0.0082 \pm 1.7 \cdot 10^{-3}$	$-0.0190 \pm 5.7 \cdot 10^{-3}$	$0.0325 \pm 5.9 \cdot 10^{-3}$
8	$0.0383 \pm 2.7 \cdot 10^{-3}$	$0.0365 \pm 2.9 \cdot 10^{-3}$	$0.0984 \pm 9.5 \cdot 10^{-3}$
9	$-0.0028 \pm 1.8 \cdot 10^{-4}$	$-0.0095 \pm 2.2 \cdot 10^{-4}$	$-0.0071 \pm 5.6 \cdot 10^{-4}$
10	$-0.0365 \pm 5.1 \cdot 10^{-3}$	$-0.0334 \pm 5.0 \cdot 10^{-3}$	$-0.0400 \pm 7.2 \cdot 10^{-3}$
11	$-0.0014 \pm 1.2 \cdot 10^{-4}$	$-0.0041 \pm 1.9 \cdot 10^{-4}$	$-0.0052 \pm 2.5 \cdot 10^{-4}$

1: LS-simple; 7: GLASSO increasing precision; 8: MR-LS; 9: SARA-QMM; 10: CLEAN-QMM;

11: Direct optical.

Table 4.27 – Summary of the additive fitting biases as a result of studying all main models, using multi-profiled simulated data and a realistic SKA-like coverage with SNR = 20 dB.

Model	c_{real}	c_{imag}	c_{abs}
1	$-0.00003 \pm 1.9 \cdot 10^{-5}$	$-0.0041 \pm 1.9 \cdot 10^{-4}$	$-0.0141 \pm 3.3 \cdot 10^{-4}$
7	$0.0001 \pm 1.2 \cdot 10^{-4}$	$-0.0077 \pm 4.3 \cdot 10^{-3}$	$-0.0116 \pm 6.5 \cdot 10^{-3}$
8	$-0.0003 \pm 1.0 \cdot 10^{-4}$	$-0.0114 \pm 5.1 \cdot 10^{-3}$	$-0.0216 \pm 8.8 \cdot 10^{-3}$
9	$0.0001 \pm 1.8 \cdot 10^{-5}$	$0.00007 \pm 2.1 \cdot 10^{-5}$	$0.0003 \pm 2.4 \cdot 10^{-5}$
10	$0.0021 \pm 6.3 \cdot 10^{-3}$	$0.0005 \pm 6.5 \cdot 10^{-3}$	$0.0109 \pm 7.9 \cdot 10^{-3}$
11	$0.00003 \pm 1.8 \cdot 10^{-5}$	$0.00004 \pm 2.1 \cdot 10^{-5}$	$0.0008 \pm 4.5 \cdot 10^{-5}$

1: LS-simple; 7: GLASSO increasing precision; 8: MR-LS; 9: SARA-QMM; 10: CLEAN-QMM;

11: Direct optical.

---

§Regarding the results in Tables 4.25, 4.26, and 4.27, the use of the SKA-like coverage does not apply to the model making a direct measurement in the optical data, where we do not generate visibilities.

Figure 4.26 depicts the biases we got from the fitting procedure after the experiments with all implemented model variants, applied to multi-profile simulated objects with a realistic SKA-like uv coverage. Moreover, in Figure ,4.27 we demonstrate the correlation in pairs among the processing time, the measurement error for the complex magnitude, and the fitting quality measured in Table 4.24 both for linear and polynomial fit. Finally, Table 4.25 summarizes the error estimations from the analysis using the main models studied in this chapter under the above-mentioned conditions. Similarly, Tables 4.26 and 4.27 give a summary of the relevant multiplicative and additive biases from the fitting of linear models to the measured values.

Summary and Next Steps

$\mathcal{F}$ OLLOWING the analysis made in this Thesis, this chapter summarizes the most important points included in the study, with Section 5.1 being dedicated to this synopsis. Moreover, Section 5.2 discusses the potential next steps that could continue the research in the topic.

5.1   Synopsis of the main points

This study developed and evaluated a collection of frameworks for galaxy shape measurements in astrophysical images. For this purpose, we moved in two different implementation directions. In the first one, we attempted to solve the problem after fitting the data to the appropriate dictionary. The choice of the parameters of this dictionary is the focal point of these methods. In the second direction, we make the measurement after reconstructing the sky brightness using algorithms that will optimize this task.

For the fitting-based approaches, the basic concept is built on the direct processing of radio data. Our simulations generate them using datasets of simulated intensity images containing elliptical astronomical objects of random ellipticity, profile, size, orientation, and position. In the first phase, we extract the object from the data, and we construct the visibilities with the help of a predefined set of points in the uv plane. Then, we decompose the data employing the Cartesian shapelets to build a dictionary of basis functions. After we transform the coefficients of the object's decomposition into their polar version, an unbiased estimator that utilizes this information is applied to perform the ellipticity measurement. Each image is processed separately up to the ellipticity estimation stage, where the information from the whole dataset is used.

In the **LS-simple** model, we apply the approximation equations used in other works on the subject. Each image has its own dictionary for decomposition based on the parameter values selected in the previous step. In the next step, we try to fit the visibilities to the dictionary we constructed, estimating the coefficients that we use for the shape measurement.

In the **GLASSO minimization** we aim for the automatic selection of the optimal value of the β parameter, and we construct a sparse regularization solution. This method finds the group of coefficients giving the minimum relative residual in the fitting of the object in the dictionary, using the group sparsity theory to a long shapelet dictionary that contains basis functions evaluated in a wide range of scale values. For this task, we developed and tested four variants of this algorithm. One of them constructs a constant dictionary for all images and analyzes a large variety of scale values, a tactic that suffers from the lack of precision in the result. Another variant uses the same approximation equation with the **LS-simple** model and determines the range of values around the estimated scale value from the equations, constructing an individual dictionary for each image. In one more version, we develop a mixed method combining **LS-simple** and sparse regularization. There, only the objects measured with large ellipticity with **LS-simple** are re-processed using a constant long dictionary and sparse regularization, aiming to reduce the produced error that is often large in those cases. Further optimization of the algorithm may be done using repeated sparse regularization using dictionaries that increase their precision in every iteration of the algorithm, which is a modification included in the last proposed variant of the model.

Motivated by the idea of the minimization of the fitting error, we developed the **MR-LS** method, which keeps the strong points of both previously mentioned cases. We have a model built on a straightforward context that automatically selects the optimal value for the β parameter, providing a measurement without prior knowledge of the object characteristics. We start similarly to the other two methods, and we construct many short dictionaries for a single scale value, chosen uniformly from a predefined interval. Then, we exhaustively analyze the data using the aforementioned dictionaries until we find the one that minimizes the fitting relative residual. From this grid search optimization solution, we locate the parameter value that best suits our data. Then, the measurement is made using this value, a procedure that makes the whole execution much faster with an algorithm that needs 4% of the processing time of the **GLASSO minimization**.

In the second implementation direction, we start from the same intensity information, and we construct the interferometric data similarly to the fitting-based methods. Then, we consider the shape measurement problem as a two-step task. The first part focuses on the precise reconstruction of the sky's brightness information from

the visibilities. The second part measures the shape of the objects in the restored intensity data after extracting them. The measurement is done using the quadruple moment technique, widely popular among scientists in the field to perform shape estimation in optical data.

In our proposed **SARA-QMM** algorithm, we apply an existing approach for image reconstruction to perform the first step's goal. SARA utilizes compressed sensing and complex optimization techniques to restore with high accuracy the intensity information from visibilities that correspond to the limited coverage of the uv plane. As an alternative approach, and to compare the efficiency of the **SARA-QMM** model, we developed a model that uses the MS-CLEAN deconvolution method to restore the sky brightness (**CLEAN-QMM**). This is considered a model that includes state-of-the-art tools that are very popular in astrophysics to solve this type of problem.

During our analysis of the performance of the models, we deduced that all shapelet-fitting models present larger measurement errors for objects with large ellipticity. **LS-simple** gives accurate ellipticity estimation in a short time due to its low computational complexity. However, it assumes a certain knowledge on the object characteristics, usually obtained from the intensity data. On the other hand, the automatic **GLASSO minimization** gives very precise results in most cases, but the size of the dictionary used in the models increases the computational complexity of the problem and makes the solution significantly slower. Furthermore, the parameters of the group sparsity regularization need to be well approximated to increase its efficiency. With the **MR-LS** algorithm, we can accomplish better precision in the measurements with results comparable to the simple approach, although the method performs purely automatically. Additionally, we reduced the computational weight of the problem to almost 4% of this in the group sparsity solution. The introduction of **SARA-QMM** gave a remarkable precision to the shape measurement, applying a different direction. It avoids large errors on objects with large ellipticity and works well independently of the type and characteristics of the data, showing increased performance in larger coverage sizes. Finally, **CLEAN-QMM** is a satisfactory model that measures with adequate accuracy the shape of the objects, but its efficiency is inferior to the one of the **SARA-QMM**. All models have negligible sensitivity in the presence of noisy data, proving that considerably good measurements can be achieved even with $SNR = 8$ dB. In our way to make the conclusions, we tested all methods in different types of data and coverage patterns with several sizes. We also presented a study of the same solutions on real data from the COSMOS survey.

Concerning the main goal of this Thesis, we accomplished to develop and evaluate a series of frameworks able to measure the shape of objects in radio interferometric data with random features. These measurements are made with high precision and

accuracy levels and without demanding any prior knowledge on the properties of the objects. The methods were analyzed in detail for their characteristics, detecting the advantages and disadvantages of their application in future research projects.

5.2 Proposed next steps

As we saw in our experimental analysis, a few points in the developed solution could be considered for their further improvement, aiming at better accuracy in the measurements. In this section, we will provide some ideas on how this can be achieved in each method.

A general issue that could be the topic of research in fitting-based methods would be their ability to simultaneously process multiple objects in the field-of-view. The current implementation offers a shape estimation for objects located in separate images because the visibilities vector must include information for a single entity. At the moment, this is done by extracting the objects in the intensity images before generating the interferometric data. Although, for more effective implementation, this must be done directly in the visibilities, which would require the adoption of a preprocessing step that will apply information unmixing techniques.

Regarding the specific models that use shapelet decomposition, **GLASSO regularization** is the one that has a certain space for enhancement with extra features. The first element to consider is the significantly large processing time in comparison to the other methods. This could be faced by an alternative implementation of the algorithm that will remain the same in the core but will include a parallelized strategy in the execution to minimize the time consumed in steps that can be processed independently and simultaneously. The second aspect for consideration is the convergence of the method and its optimization. We show that when the data are generated using a realistic SKA-like coverage, the convergence becomes more difficult. To achieve it, we must precisely set the regularization and step gradient parameter values. The regularization parameter has an efficient estimator, but for the other parameter, we have not adopted such a solution. A different solution for this, except for using an estimator, would be to employ a machine learning technique for parameter estimation called K-fold cross-validation. We can quantify how well a method generalizes using the generalization error that computes the expected prediction error for unseen data. Ideally, we must choose the gradient step to minimize the mistakes that will be made in the future. Since we do not have the future data, we can simulate the future by using the data we have, which is accomplished by cross-validation. To attain this objective, we split the data into two sets: the train set and the validation one. There is not a strict rule regarding the proportions in each set, but a good balance is usually 80%-20%, respectively.

Then, we pretend that the validation set is our unknown data, and we train our model using the train set. An estimate of the generalization error can be achieved by computing the measurement error in the validation data after the end of the training phase. We perform tests using this setup and several gradient step values. Ideally, we also repeat the procedure as many times as possible by partitioning the data in different ways, giving us confidence in our results. Likely the most efficient way to do this is the K-fold version of the cross-validation setup. In this case, instead of creating sets randomly each time, we split our data into K groups (typically K = 5). At each iteration, we choose one group to be the validation test, while the rest constitute the training set. The iterations continue until we have used all groups as validation sets, and the results are given by the mean of the results at each iteration. As we mentioned in Chapter 4, the gradient step must ideally be selected per image. However, an analysis, as described above, could give a very good approximate value that will minimize the errors for all measurements.

The most important point of concern in **SARA-QMM** solution is the processing time. As explained in more detail earlier, this can be optimized by providing an image for the analysis that contains several objects, where the resulting processing time will be reduced as many times as the number of objects in the image. Nonetheless, we can offer an extra boost in the computational simplicity of the model by introducing parallelization in its phases. There are steps in the algorithm that can be implemented to process the data parallelly, expecting to decrease the execution time to 25% of the non-parallel version.

Except for the already developed approaches, solutions can be proposed by utilizing tools from different domains. One of them could include applying deep learning to this problem, a path that has not been investigated in the past. A potential plan would be to feed the reconstruction to a deep neural network that would identify the parameters of interest (e.g., the ellipticity field at some appropriate resolution in the field of view). This can also lead to a comparative analysis with a model that will consist of a neural network working directly on the visibilities.

Massive Open Online Courses on Radio Astronomy

A substantial part of this Thesis was dedicated to the development of a series of courses in radio astronomy. This chapter presents these courses, starting by introducing the motivations for their creation in Section 6.1. Sections 6.2 and 6.3 describe the structure and content of each course, while Section 6.4 gives the details on the contribution of the Thesis author to this development.

6.1 Motivations

As demonstrated in Section ,1.1 radio astronomy is a very intriguing scientific area over the past decades, with many new pieces of research having their concentration in developing this subject or using it to answer fundamental cosmological questions. The interest in the field is equally intense among both professional astronomers and amateurs, and a series of courses that will introduce the basic theoretical and practical aspects of radio astronomy was necessary.

It is expected that radio astronomy will become a powerful and complementary to the other way to perform observations. For this reason, bachelor or master students that are motivated to pursue a Ph.D. in astrophysics will be able to take the courses and get help in building a solid basis on how to manipulate and analyze radio interferometer data and develop new imaging data analysis tools. In addition, professionals, like engineers and scientists interested in applying their expertise to new projects related to the implementation of radio telescopes, such as the SKA project in South Africa and Australia, will have the opportunity to expand their knowledge helped by our courses.
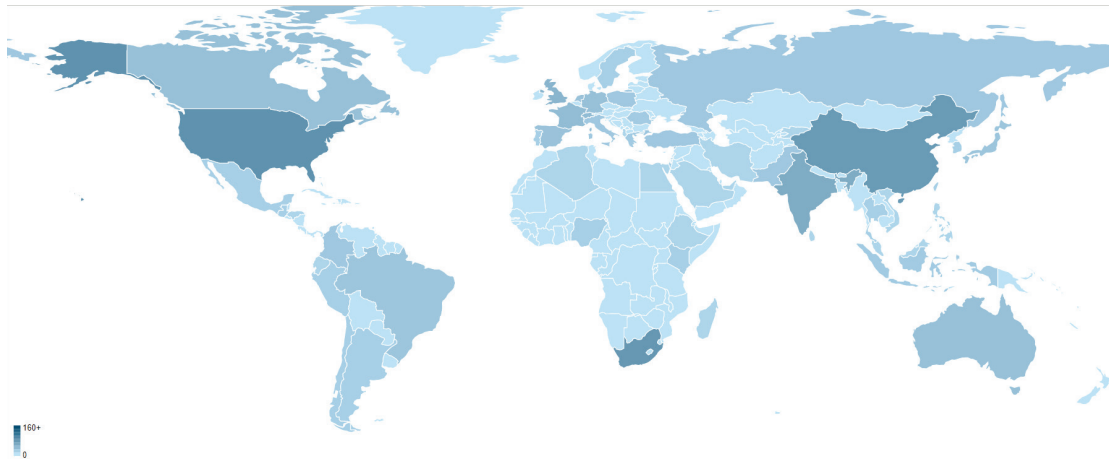
(a) Geographical distribution
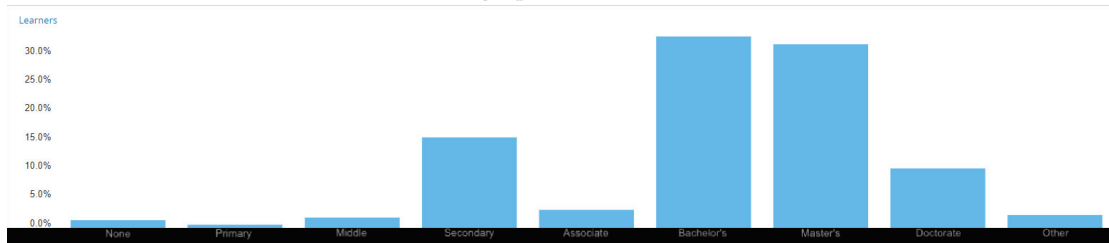
(b) Educational distribution
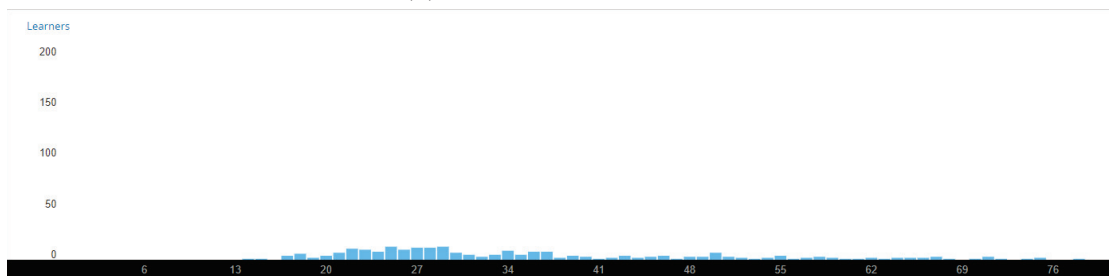
(c) Age distribution

Figure 6.1 – Statistics regarding the participation to the premier run of the first course, where we present: **(a)** the geographical, **(b)** the educational, and **(c)** the age distributions.

Through their interaction with the course, they will be able to get a better understanding of the scope of these projects through the mathematical and astronomical concepts included in the courses. Furthermore, they are expected to learn the techniques of correct manipulation of the radio interferometer data using modern software developed for this purpose.

These MOOCs are planned to be used to train professionals involved in the SKA projects and students studying in areas related to large radio telescope projects in technical universities across South Africa.
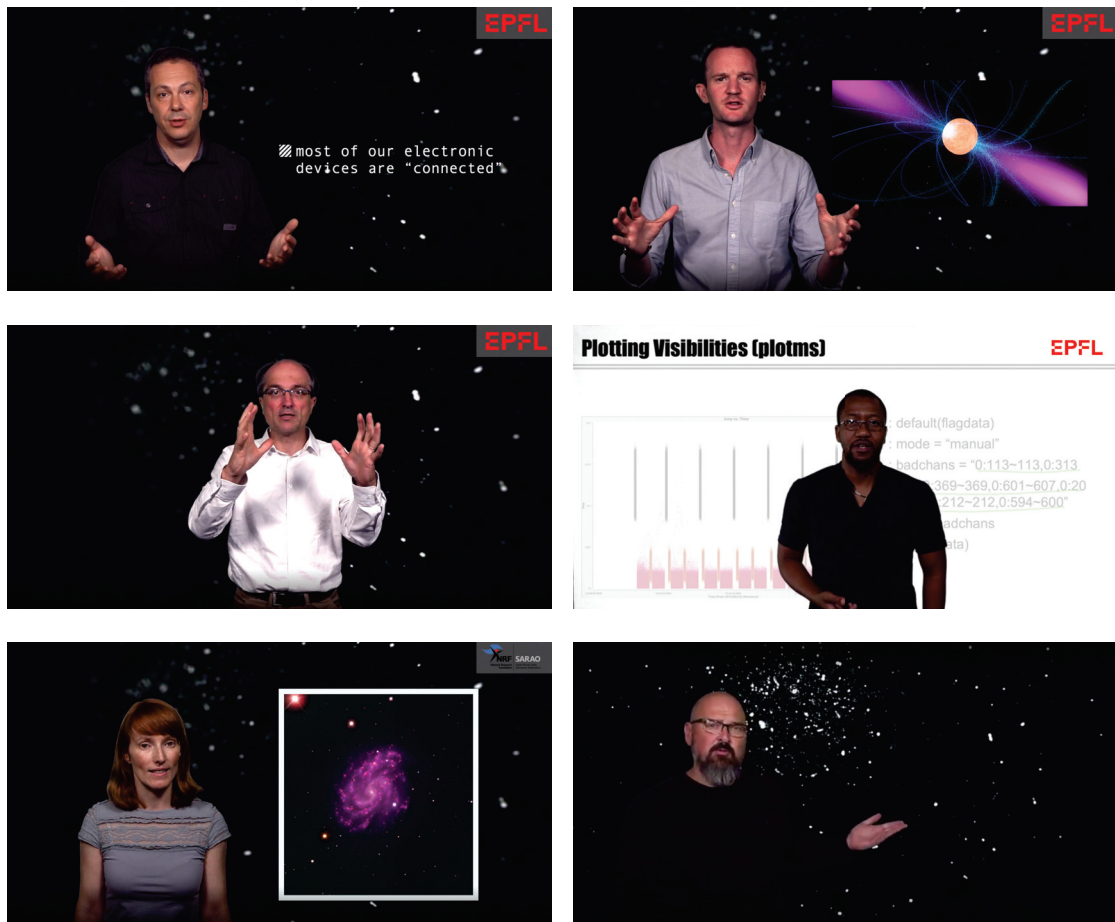
Figure 6.2 – Representative parts of the video lectures including the lecturers presenting the material of the courses.

One more factor that motivated the development of these courses was the absence of a large variety of material in the form of Massive Open Online Courses in related topics. There were a few cases of courses in astronomy and astrophysics, including a general background of the whole domain that were quite descriptive. Although, there was no material on radio astronomy, which increased the importance of this work.

Finally, the material was considered the best idea to be available in the form of MOOCs to give the advantage of the large potential audience. A typical university course is limited to the students enrolled in a program or, in the best case, in the whole university. On the other hand, a MOOC would attract people from every part of the world, with or without scientific expertise, with a diverse educational background or age. As shown in Figure 6.1 that comes from the inaugural run of the first course, the geographical (6.1a), the educational (6.1b), and the age (6.1c) distributions confirm this consideration.

The series contains two courses of 6 weeks syllabus in English language and is supported by the *edX* MOOC platform. They are freely accessible to anyone, with the possibility for the participant to receive a verified certificate after the successful completion for a fee. The level is adapted for students at the end of their bachelor's or the beginning of their master's, assuming only basic knowledge in mathematics, physics, and computer science concepts. The courses' team consists of (alphabetically):

Vasileios ANGELOPOULOS	Dr. in Image Processing & Radio Astronomy	EPFL
Frédéric COURBIN	Prof. in Astrophysics	EPFL
Griffin FOSTER	Dr. in Radio Astronomy	Zendar
Jean-Paul KNEIB	Prof. in Astrophysics	EPFL
Sphesihle MAKHATHINI	Dr. in Radio Astronomy	Wits University
Kim MCALPINE	Dr. in Radio Astronomy	SARAO
Oleg SMIRNOV	Prof. in Astronomy & Astrophysics	Rhodes University & SARAO

6.2   The Radio Sky I: Science and Observations

In this course, participants will learn about the most important objects that emit in the radio regime, both within and beyond our own galaxy. Covering topics from the birth and death of stars to the detection of gravitational waves, it will be explained how radio observations contribute to a richer understanding of the Universe we inhabit. We introduce the basics of radio detectors and the technologies we use to observe at radio wavelengths. In this section, we analyze the structure and the content of this course[*].

Week 1: General Introduction

Description
This week participants will learn why radio observations are essential for astrophysics and cosmology, the early history of radio observations, and some initial technical principles.

---

[*]The course is permanently offered in self-paced learning mode, accessible through its landing page: www.edx.org/course/radio-sky-1

Objectives

- Understand how we can probe the Universe. Why we do radio astronomy. What types of radio signals exist and their origin (man-made vs. physical processes).

- The nature of happenstance discoveries and their role in driving progress. Discovery drives technology and vice versa.

- Recall/recite the overall history of the Universe from the Big Bang to galaxy formation to modern times.

- Understand the different scales of the Universe and which structures they correspond to.

- Recall the history of radio astronomy and radio observations.

Section 1: Introduction

1. Text Page: Introduction and Objectives

2. Video Lecture: Why Radio?

3. Discussion Prompt: Radio Facilities that You Know

4. Discussion Prompt: The Observable Universe

Section 2: Historical Perspective

1. Text Page: Radio Observation Historical Events

2. Video Lecture: Early History (Before 1950)

3. Practice Quiz: Early History (Before 1950)

4. Video Lecture: History (After 1950)

5. Practice Quiz: History (After 1950)

6. Discussion Prompt: Recent Radio Astronomy History

Section 3: Radio Observations

1. Text Page: List of Key Radio Observations

2. Video Lecture: Key Radio Observations

3. Practice Quiz: Key Radio Observations

4. Discussion Prompt: Radio Observations

Section 4: Basic Properties of the Universe

1. Text Page: Cosmological and Astrophysical Contexts

2. Video Lecture: Cosmological Context

3. Practice Quiz: Cosmological Context

4. Video Lecture: Galaxies and Galaxy Clusters

5. Practice Quiz: Galaxies and Galaxy Clusters

6. Video Lecture: Star Formation

7. Practice Quiz: Star Formation

Section 5: Graded Weekly Assignment 1

Week 2: The Photon Messenger and the Blackbody Emission

Description
This week participants will learn about light and its properties. Additionally, we will study how photons propagate in the mediums.

Objectives
- Know that light can be seen as a particle or a wave.
- Recall all properties of photons (emission, absorption).
- Understand the frequency-energy relationships.
- Understand what polarization is and what physical processes can give rise to it. Give concrete examples to support this.
- Analyze the spectra and use the features in the spectrum to identify the emission mechanism.

Section 1: Introduction

1. Text Page: Introduction and Objectives

2. Video Lecture: Properties of Photon

3. Practice Quiz: Photon Properties

4. Video Lecture: Photon Propagation

5. Practice Quiz: Photon Propagation

Section 2: Photons

1. Text Page: Blackbody Emission

2. Video Lecture: Blackbody Emission and Radiation Transport (5 Video Parts)

3. Practice Quiz: Emission and Absorption of Photons

4. Video Lecture: Spectra in Different Units

Section 3: Graded Weekly Assignment 2

Week 3: Continuum Photons and Line Physics

Description
This week participants will learn about the continuum emission mechanisms and physics behind the emission lines.

Objectives
- Understand and describe how each emission mechanism works.

- Understand how the spectral lines are produced.

- Analyze the spectra and use the features in the spectrum to identify the emission mechanism.

- Calculate the temperature of a blackbody from its spectrum.

Section 1: Introduction

1. Text Page: Introduction and Objectives

Section 2: Other Emission Types

1. Text Page: Continuum Emission and Line Physics

2. Video Lecture: Production of Continuum Photons (4 Video Parts)

3. Practice Quiz: Production of Photon Continuum

4. Video Lecture: Line Physics (5 Video Parts)

5. Practice Quiz: Line Physics

Section 3: Graded Weekly Assignment 3

Week 4: Photon Detection and Detectors

Description
This week participants will learn about how photons are detected. We will also investigate the type of detectors and their properties.

Objectives
 - Recall the different types of detectors and how they relate to different wave-lengths.

 - Evaluate the signal-noise ratio of detection as a function of time depending on the type of noise.

 - Calculate the energy/wavelength/frequency of light.

 - Express the above quantities in different units.

Section 1: Introduction

  1. Text Page: Introduction and Objectives

Section 2: Detecting Photons

  1. Text Page: Photon Detection

  2. Video Lecture: Basic Principles of Photon Detection

  3. Practice Quiz: Photon Detection

  4. Discussion Prompt: Detecting the Radio Wave from a Cell Phone

Section 3: Real Detectors

  1. Text Page: Real Detectors

  2. Video Lecture: High Energy Photon Detection

  3. Practice Quiz: High Energy Photon Detection

  4. Video Lecture: Radio Detectors

  5. Practice Quiz: Radio Detectors

Section 4: Graded Weekly Assignment 4

Week 5: Science in the Radio - Part A

Description
This week participants will learn about how to perform science in the radio domain.

Objectives
 - Recognize different types of astronomical objects in the Radio and identify their main morphological features.

 - Identify the emission mechanisms responsible/present for each type of object and relate them to their counterparts at other wavelengths.

 - Explain how to diagnose the emission mechanism from a spectrum.

Section 1: Introduction

   1. Text Page: Introduction and Objectives

Section 2: Sources in the Milky Way, Pulsars & Supernovae

   1. Text Page: Sky at Radio Wavelengths

   2. Video Lecture: Supernovae and Supernovae Remnants (3 Video Parts)

   3. Practice Quiz: Supernovae and Supernovae Remnants

   4. Video Lecture: Pulsars & Transients

   5. Practice Quiz: Pulsars & Transients

Section 3: AGN & Galaxies

   1. Text Page: Active Galactic Nucleus

   2. Video Lecture: AGN

   3. Practice Quiz: AGN

   4. Video Lecture: Galaxies - Continuum Emission (4 Video Parts)

   5. Practice Quiz: Galaxies - Continuum Emission

Section 4: Graded Weekly Assignment 5

Week 6: Science in the Radio - Part B

Description

This week we will continue our investigation on how to perform science in the radio domain.

Objectives

- Use a spectral line to determine the velocity, learn how to time pulsars.

- Determine which frequency ranges are appropriate for different types of objects.

- Calculate the effect of Doppler Boosting on a jet flux and calculate the age of radio objects.

- Analyze the regimes and types of phenomena where radio observations have an advantage over other wavelengths.

Section 1: Introduction

1. Text Page: Introduction and Objectives

Section 2: Galaxies & Clusters

1. Text Page: Galaxies & Clusters

2. Video Lecture: Galaxies - HI Emission

3. Practice Quiz: Galaxies - HI Emission

4. Video Lecture: Clusters of Galaxies

Section 3: Cosmic Microwave Background

1. Text Page: Introduction to CMB

2. Video Lecture: Cosmic Microwave Background

3. Practice Quiz: Cosmic Microwave Background

Section 4: Radio Surveys

1. Text Page: Introduction to Sky Surveys

2. Video Lecture: Radio Surveys

3. Practice Quiz: Radio Surveys

Section 5: Graded Weekly Assignment 6

Graded Final Assignment

Feedback and other useful information

This course was launched for the first time on October 1st, 2020, as instructor-paced and ended 7 weeks later (including the week allocated to the Final Assignment). In instructor-paced mode, the program is followed strictly by all participants with specific deadlines per task, allowing everyone to start and finish with the course run. After some updates in the content and small corrections, 'The Radio Sky I' was relaunched on March 15th, 2021, in a permanent run, following the self-paced logic, where each student set their own plan on how fast or slow wants to participate in the course.

During its inaugural run, the course accomplished a large number of participants, reaching 467 active students (462 audits and 5 paid users) and showing a continuously increasing direction. This trend was maintained even after the end date until the second launch in March 2021, when the course was archived, but interested learners could still enroll and watch the video lectures as audits. Figure 6.3a gives a diagram of the enrollment rate as it is provided by the *edX* system.

After finishing each week and at the end of the whole course, learners were invited to provide feedback describing their thoughts. The reaction from all people who participated in these surveys was overall very positive for their experience with the course, achieving an average rate for the weekly feedback of 8.2 out of 10. Additionally, at the final survey, 77% of the learners said that the presentation of the content was clear, 85% of the people stated that the difficulty of the course was just right, while the relevant index for the difficulty of the exercises was 92%. Finally, in more specific comments, some minor typographical errors were reported, and a couple of proposed changes, which were considered in the corrections we made before the second run.
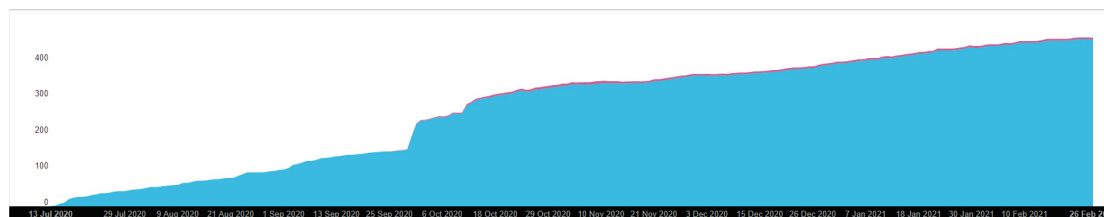
(a) Course enrollment rate

(b) Radio facilities map

Figure 6.3 – Information on the run of 'The Radio Sky I'. **(a)** The enrollment rate for the first run of the course. **(b)** A map generated from the responses in one of the course's activities, where the participants give information on radio facilities that they know.

One more factor that showed active participation was the willingness of the students to give their responses and take part in discussions on topics related to the material of the week. Several discussion prompts were distributed along the course, and learners gave their points of view and interacted with their peers. One of them is depicted in Figure 6.3b, where a map was created from the people's response regarding radio facilities that they know. In this task, they were asked to give as much information as possible for a radio facility of their choice in any part of the world. Then, using the coordinates given by them, we could generate a map accessible to everyone containing all this information.

6.3   The Radio Sky II: Observational Radio Astronomy

This course covers the principles and practices of radio astronomical observations, in particular with modern interferometers. Topics range from radio telescope technology to the measurement equation for radio interferometric calibration and imaging. The course is expected to have its premiere launch in October 2021, and it is under development at the moment this Thesis is written. Here, we present the content and structure of 'The Radio Sky II', noting that its final form might slightly differ.

Week 1: Overview of Telescope Technology - Part A

Description
This week we will learn about telescope technology, with a particular focus on optical and radio telescopes.

Objectives

- Recall the different wavelength regimes and how these correspond to different telescope technologies.

- Understand the principles of telescope optics and EM radiation collection.

- Understand the concepts of the field of view, primary beam, spectral resolution, frequency coverage, time resolution, single-pixel versus matrix.

Section 1: Introduction

1. Text Page: Introduction and Objectives

Section 2: Optical & Near Infrared Telescopes

1. Text Page: Optical Telescope

2. Video Lecture: Concept of a Telescope

3. Practice Quiz: Performance of a Telescope

4. Video Lecture: Telescope Sites

5. Discussion Prompt: Telescope Sites and Designs

6. Practice Quiz: Telescope Sites

Section 3: Radio Telescope

1. Text Page: Radio Telescope

2. Video Lecture: Single Dish

3. Discussion Prompt: Current and Planned Telescopes

4. Video Lecture: Radio Interferometer

5. Practice Quiz: Radio Interferometer

Section 4: UV, X-ray and Gamma-ray Telescopes

1. Text Page: Other Telescopes

2. Video Lecture: UV, X-ray and Gamma-ray Telescope

3. Practice Quiz: UV, X-ray and Gamma-ray Telescope

Section 5: Graded Weekly Assignment 1

Week 2: Overview of Telescope Technology - Part B

Description
This week we will continue our study on telescope technology, with a particular focus on optical and radio telescopes.

Objectives
- Calculate the resolution of any given telescope.

- Compare the sensitivities of different telescopes.

- Be able to connect different astrophysical phenomena to the wavelength regimes that they can be observed in and to the corresponding telescopes.

- Understand the basic concept of the SKA project and the technologies applied to it.

Section 1: Introduction

1. Text Page: Introduction and Objectives

Section 2: Other Telescopes

1. Text Page: Gravitational-wave Astronomy

2. Video Lecture: Neutrino and Gravitational-wave Telescope

3. Practice Quiz: Neutrino and Gravitational-wave Telescope

4. Video Lecture: Radio Counterpart of Gravitational Waves

Section 3: Radio Intelligent Signals

1. Text Page: Intelligent Signal Detection

2. Video Lecture: Signal from Intelligent Life

3. Discussion Prompt: Detection of an Arecibo-like RADAR

Section 4: The SKA Project

1. Text Page: Description of the SKA Project

2. Video Lecture: The Main Goals of SKA

3. Developing an Observing Proposal: Science Justification

Section 5: Graded Weekly Assignment 2

Week 3: Observational Radio Astronomy

Description
This week we will learn about observational radio astronomy and some key concepts related to this field.

Objectives
- The learner will be able to recite the following definitions: flux, luminosity, brightness temperature, system temperature, sensitivity, resolution, radiometer equation, spectral lines, single radio dish.
- The learner will be able to explain the different units of measure (flux, luminosity, etc.) used in radio astronomy in their own words.
- The learner will be able to explain the difference between continuous and line emission.

- The learner will be able to give examples of different single-dish telescopes.

- The learner will be able to contrast the resolution of a single dish radio telescope with an optical telescope.

- The learner will be able to discuss in their own words the two fundamental attributes of a single-dish radio telescope, namely, its sensitivity and its resolution.

- The learner will be able to determine the flux of a source given its brightness distribution.

- The learner will be able to calculate the resolution of a single dish.

- The learner will be able to use the radiometer equation to determine the sensitivity of a single dish telescope.

- The student will be able to analyze which factors influence the sensitivity and resolution of a single dish telescope.

- The student will be able to justify the types of objects that different single-dish telescopes can observe based on their characteristics.

- The student would be able to design a theoretical telescope given a list of properties it should have.

Section 1: Introduction

1. Text Page: Introduction and Objectives

Section 2: Flux, Luminosity & Brightness Temperature

1. Text Page: Flux, Luminosity & Brightness Temperature

2. Video Lecture: Flux, Luminosity & Brightness Temperature Explained

3. Jupyter Notebook: Flux, Luminosity & Brightness Temperature

4. Practice Quiz: Flux, Luminosity & Brightness Temperature

Section 3: Radio Telescope Principles

1. Text Page: Radio Telescope Types and Capabilities

2. Video Lecture: Radio Telescope Principles & Radiometer Equation

3. Jupyter Notebook: Radiometer Equation

4. Practice Quiz: Radio Telescope Principles

Section 4: Single Dish Observations

1. Text Page: Single Dish Observations

2. Video Lecture: Diversity of Telescope Technologies (2 Video Parts)

3. Developing an Observing Proposal: Choosing Radio Facility, Sensitivity Calculations

4. Developing an Observing Proposal: Technical Aspects

5. Practice Quiz: Telescope Diversity

Section 5: Time-domain Radio Astronomy

1. Text Page: Time-domain Radio Astronomy

2. Video Lecture: Observing Pulsars

3. Video Lecture: Transients & FRBs

4. Practice Quiz: Observing Pulsars, Transients & FRBs

Section 6: Graded Weekly Assignment 3

Week 4: Imaging with an Interferometric Array

Description
This week we will learn about imaging science and strategies using an interferometric array.

Objectives
- The learner will be able to recall and recite the following concepts: Fourier Transform, uv-coverage, imaging, deconvolution, visibilities, and point-spread function.
- The learner will be able to explain how the imaging (lm-plane) and visibility domain (uv-plane/spatial frequencies) are related by the Fourier Transform. They would be able to explain why an interferometer acts as a spatial filter by using the VLA as an example (which includes an understanding of what the point-spread function is).
- They would be able to describe what is meant by the word imaging. They would be able to summarize the most basic deconvolution algorithm. They would be able to explain what the calibration entails.

- They would be able to calculate a few basic Fourier Transforms. They would be able to calculate the basic visibility function associated with a very basic sky distribution (one or two-point sources, etc.). Calculate the basic resolution of an interferometer. Determine if a source will be resolved or unresolved by a specific telescope.

- They would be able to associate different *uv* coverage and point-spread functions with each other. They would be able to predict how the *uv* coverage would affect the point-spread function and the final image.

- They would be able to judge which array or configuration of an array is more appropriate to observe specific sources and conduct particular experiments.

- They would be able to design an array layout for a particular use case (e.g., play with Ivan Marti-Vidal's APSYN interferometry simulator).

Section 1: Introduction

1. Text Page: Introduction and Objectives

Section 2: Fourier Transform and Spatial Frequencies

1. Text Page: Fourier Transform

2. Video Lecture: Fourier Transform and Spatial Frequencies

3. Jupyter Notebook: Fourier Transform and Spatial Frequencies

4. Practice Quiz: Fourier Transform and Spatial Frequencies

Section 3: Sampling and Point-Spread Functions

1. Text Page: Sampling & PSFs

2. Video Lecture: Sampling and Point Spread Functions

3. Jupyter Notebook: Sampling and Point Spread Functions

4. Practice Quiz: Sampling and Point Spread Functions

Section 4: Visibilities of a Celestial Source

1. Text Page: Visibilities of a Celestial Source

2. Video Lecture: A van Cittert-Zernike Interlude

3. Jupyter Notebook: Point Source Visibility Function

4. Practice Quiz: Visibility Function

Section 5: APSYNSIM

1. Text Page: APSYNSIM Documentation

2. Video Lecture: APSYNSIM Demo

3. Peer-graded Assignment: APSYNSIM Telescope Assignment

Section 6: Graded Weekly Assignment 4

Week 5: From Visibilities to Images - Part A

Description
This week we will learn about the cycle of radio interferometric data.

Objectives
- Understand the cycle of radio interferometric data from visibilities to images.
- Remember the range of currently operating interferometric arrays and their general capabilities.
- Be able to compute the survey speed, given the parameters of a radio interferometer.
- Data reduction tutorial: be able to create an image from a small set of observations.
- Derive some scientific parameters (flux, spectral index, brightness temperature) from the results of the data reduction tutorial.

Section 1: Introduction

1. Text Page: Introduction and Objectives

Section 2: RIME Introduction

1. Text Page: RIME Introduction

2. Video Lecture: RIME I - Stokes Parameters

3. Jupyter Notebook: Stokes Parameters

4. Video Lecture: RIME II - Jones Calculus

5. Jupyter Notebook: Jones Matrices

6. Practice Quiz: RIME I & II

Section 3: All-sky Radio Interferometry

1. Text Page: All-sky Radio Interferometry

2. Video Lecture: RIME III - All-sky

3. Video Lecture: RIME IV

4. Practice Quiz: RIME III & IV

5. Video Lecture: Data Inspection & Flagging

6. Data Notebook

Section 4: Graded Weekly Assignment 5

Week 6: From Visibilities to Images - Part B

Description
This week we will continue our investigation into the cycle of radio interferometric data.

Objectives
- Understand the cycle of radio interferometric data from visibilities to images.
- Remember the range of currently operating interferometric arrays and their general capabilities.
- Be able to compute the survey speed, given the parameters of a radio interferometer.
- Data reduction tutorial: be able to create an image from a small set of observations.
- Derive some scientific parameters (flux, spectral index, brightness temperature) from the results of the data reduction tutorial.

Section 1: Introduction

1. Text Page: Introduction and Objectives

Section 2: Calibration

1. Text Page: Calibration

2. Video Lecture: Calibration & Self-calibration (1GC & 2GC)

3. Practice Quiz: Calibration & Self-calibration (1GC & 2GC)

4. Video Lecture: Cross-calibration (1GC)

5. Developing an Observing Proposal: Observational Strategy

Section 3: Deconvolution

1. Text Page: Deconvolution

2. Video Lecture: CLEAN

3. Practice Quiz: Deconvolution

4. Video Lecture: CLEAN Demo

Section 4: Self-calibration

1. Text Page: Self-calibration

2. Video Lecture: Self-calibration (2GC)

3. Practice Quiz: Self-calibration (2GC)

Section 5: Graded Weekly Assignment 6

Graded Final Assignment

6.4   Personal contribution

This section includes the Thesis author's contribution to the development and maintenance of these MOOCs, giving a brief list of the main activities undertaken for this purpose. More specifically, these tasks include the following:

- Participate in weekly meetings with other team members.

- Organize and participate in individual meetings with other people assisting the development (mainly, people from CEDE team).

- Compute the total duration of the weeks in the original plan and check the topics of the videos to propose an efficient split into two different courses.

- Develop the practice quizzes for Radio Sky I and Radio Sky II and name them appropriately.

- Develop the graded assignments for Radio Sky I.

- Write the text pages for Radio Sky I and Radio Sky II and name them appropriately.

- Develop the discussion prompts for Radio Sky I and Radio Sky II.

- Organize the original blueprint.

- Develop new blueprints for Radio Sky I and Radio Sky II after splitting the course.

- Maintain and update the blueprints regularly.

- Develop a material status spreadsheet including all details of the course material.

- Develop new material status spreadsheets for Radio Sky I and Radio Sky II after splitting the course.

- Maintain and update the material status document regularly.

- Organize the whole course material in Google Drive into weeks and lectures with dedicated numbering for each element.

- Re-organize the material in Google Drive after splitting the course. Re-number everything and appropriately rename weeks, sections, lectures, and other elements.

- Re-number the videos/lectures in the Trello board after the first changes in the original blueprint.

- Re-number, re-organize, and update everything in all above documents/platforms in each modification in the course elements.

- Prepare the keywords list for Radio Sky I.

- Re-organize the team members' biography document to make it clearer.

- Re-organize the landing page document for both courses to make it clearer.

- Change the writing format of the whole course material so that it will be possible to upload it to the *edX* platform.

- Prepare and review the pages for the 'Introduction' section, including material on the basics of the course.

- Upload the whole course material into the *edX* platform.

- Keep track of the open issues and regularly update the other team members on them.

- Follow the course discussion forum and interact with the learners whenever necessary.

- Update Radio Sky I to prepare it for the relaunch in self-paced mode.

- Setup several documents for the courses (e.g., copyright release agreements), distribute them for signature, collect and submit them.

Conclusions

*W*E presented a work that treats a fundamental problem in cosmology, and this is the matter distribution in the Universe. To accomplish this objective, we study galaxy shape measurement algorithms that can help reveal the underlying distribution in a local neighborhood of galaxies by estimating the shear caused by the weak gravitational lensing effect due to the gravitational fields of large-scale structures. The analysis is based on radio interferometric data, utilizing the substantial recent progress in the area of radio astronomy and developing solutions that would benefit from the data collected by modern radio telescope projects, such as SKA.

We presented two types of methods for object shape measurement, built on different approaches. In one of them, we consider a study that has been done in the past in a similar project, and we go one step forward. We decompose the studied objects using a collection of shapelet basis functions, and we estimate their ellipticity based on the decomposition coefficients. A critical aspect of achieving a precise result in this category of models is the reliable approximation of the parameters of the basis functions. Our models implement different ways to achieve this.

In the other type of proposed methods, we set up a two-step algorithm that reconstructs the original sky brightness information from the visibilities and measures the ellipticity of the objects from the produced intensity data. Towards that goal, we employ two different algorithms: the state-of-the-art MS-CLEAN deconvolution method that is well-known in the astrophysics community for this category of problems, and the SARA algorithm that is a modern image reconstruction solution applying complex optimization and compressed sensing means to restore the intensity information from the radio data accurately.

After analyzing each method separately and using simulated and real datasets containing galaxy entities of several properties, we evaluated the efficiency of the developed schemes. We characterized the models and identified their advantages and disadvantages that could lead to further studies aiming to improve them or new research based on the implemented algorithms to explore a problem in the domain.

Motivated by the strong interest in radio astronomy of people coming from diverse educational backgrounds, we also developed a series of Massive Open Online Courses (MOOCs) that introduce the basic knowledge in the field to learners around the world. People with professional or amateur experience in astronomy and astrophysics have the possibility to freely take the courses and benefit from online learning wherever they are and whenever they want to. According to the participants' feedback and evaluation, students are very satisfied by the material presented and feel excited by this opportunity. These conclusions are also verified by the increasing enrollment rate shown in the course statistics.

Bibliography

Bach, F. R. (2008). Consistency of the group lasso and multiple kernel learning. *J. Mach. Learn. Res.*, *9*(Jun), 1179–1225.

Bacon, D. J., Massey, R. J., Refregier, A. R., & Ellis, R. S. (2003). Joint cosmic shear measurements with the keck and william herschel telescopes. *MNRAS*, *344*(3), 673–685.

Bacon, D. J., Refregier, A. R., & Ellis, R. S. (2000). Detection of weak gravitational lensing by large-scale structure. *MNRAS*, *318*(2), 625–640.

Baraniuk, R. G. (2007). Compressive sensing [lecture notes]. *IEEE signal processing magazine*, *24*(4), 118–121.

Bartelmann, M., & Schneider, P. (2001). Weak gravitational lensing. *Physics Reports*, *340*(4-5), 291–472.

Bauschke, H. H., Combettes, P. L. et al. (2011). *Convex analysis and monotone operator theory in hilbert spaces* (Vol. 408). Springer.

Bertsekas, D. P. (1982). *Constrained optimization and lagrange multiplier methods*. Academic press.

Blandford, R., Saust, A., Brainerd, T., & Villumsen, J. (1991). The distortion of distant galaxy images by large-scale structure. *MNRAS*, *251*(4), 600–627.

Bobin, J., Starck, J.-L., Fadili, J. M., Moudden, Y., & Donoho, D. L. (2007). Morphological component analysis: an adaptive thresholding strategy. *IEEE transactions on image processing*, *16*(11), 2675–2681.

Boyd, S., Parikh, N., & Chu, E. (2011). *Distributed optimization and statistical learning via the alternating direction method of multipliers*. Now Publishers Inc.

Brainerd, T. G., Smail, I., & Mould, J. (1995). Evolution in the clustering of galaxies to r= 26. *Monthly Notices of the Royal Astronomical Society*, *275*(3), 781–789.

Brown, M., Bacon, D., Camera, S., Harrison, I., Joachimi, B., Metcalf, R., Pourtsidou, A., Takahashi, K., Zuntz, J., Abdalla, F., et al. (2015). Weak gravitational lensing with the square kilometre array. *arXiv preprint arXiv:1501.03828*.

Burke, B. F., Graham-Smith, F., & Wilkinson, P. N. (2019). *An introduction to radio astronomy*. Cambridge University Press.

Candes, E. J. (2008). The restricted isometry property and its implications for compressed sensing. *Comptes rendus mathematique*, *346*(9-10), 589–592.

Candès, E. J. et al. (2006a). Compressive sampling. *Proceedings of the international congress of mathematicians*, *3*, 1433–1452.

Candès, E. J., Romberg, J., & Tao, T. (2006b). Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on information theory*, *52*(2), 489–509.

Carrillo, R. E., McEwen, J. D., & Wiaux, Y. (2014). Purify: a new approach to radio-interferometric imaging. *MNRAS*, *439*(4), 3591–3604.

Carrillo, R., McEwen, J., & Wiaux, Y. (2012). Sparsity averaging reweighted analysis (sara): a novel algorithm for radio-interferometric imaging. *MNRAS*, *426*(2), 1223–1234.

Chang, T.-C., & Refregier, A. (2002). Shape reconstruction and weak lensing measurement with interferometers: a shapelet approach. *ApJ*, *570*(1), 447.

Chang, T.-C., Refregier, A., & Helfand, D. J. (2004). Weak lensing by large-scale structure with the first radio survey. *ApJ*, *617*(2), 794.

Combettes, P. L., & Pesquet, J.-C. (2011). Proximal splitting methods in signal processing, chapter fixed-point algorithms for inverse problems in science and engineering.

Combettes, P. L., & Pesquet, J.-C. (2012). Primal-dual splitting algorithm for solving inclusions with mixtures of composite, lipschitzian, and parallel-sum type monotone operators. *Set-Valued and variational analysis*, *20*(2), 307–330.

Condat, L. (2013). A primal–dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms. *Journal of optimization theory and applications*, *158*(2), 460–479.

Condon, J. J., & Ransom, S. M. (2016). *Essential radio astronomy* (Vol. 2). Princeton University Press.

Cornwell, T. J. (2008). Multiscale clean deconvolution of radio synthesis images. *IEEE Journal of selected topics in signal processing*, *2*(5), 793–801.

Daubechies, I. (1992). *Ten lectures on wavelets*. SIAM.

Donoho, D. L. (2006). Compressed sensing. *IEEE Transactions on information theory*, *52*(4), 1289–1306.

Eddington, A. S. (1920). *Space, time and gravitation: an outline of the general relativity theory*. University Press.

Fornasier M., R. H. (2011). In O. Scherzer (Ed.), *Handbook of mathematical methods in imaging* (pp. 187–228). Springer Science & Business Media.

Garsden, H., Girard, J., Starck, J.-L., Corbel, S., Tasse, C., Woiselle, A., McKean, J., Van Amesfoort, A. S., Anderson, J., Avruch, I., et al. (2015). Lofar sparse image reconstruction. *A&A*, *575*, A90.

Geisler, D., & Hodge, P. (1980). Ellipticities of globular clusters of the large magellanic cloud. *ApJ*, *242*, 66–73.

Goldberg, D. M., & Natarajan, P. (2002). The galaxy octopole moment as a probe of weak-lensing shear fields. *The Astrophysical Journal*, *564*(1), 65.

Gribonval, R., & Nielsen, M. (2003). Sparse representations in unions of bases. *IEEE transactions on Information theory*, *49*(12), 3320–3325.

Gunn, J. E. (1967). On the propagation of light in inhomogeneous cosmologies. i. mean effects. *ApJ*, *150*, 737.

Hellbourg, G. (2014). *Radio frequency interference spatial processing for modern radio telescopes* (Doctoral dissertation). Université d'Orléans.

Hey, J. (1946). Solar radiations in the 4–6 metre radio wave-length band. *Nature*, *157*(3976), 47–48.

Heymans, C., Van Waerbeke, L., Bacon, D., Berge, J., Bernstein, G., Bertin, E., Bridle, S., Brown, M. L., Clowe, D., Dahle, H., et al. (2006). The shear testing programme i. weak lensing analysis of simulated ground-based observations. *MNRAS*, *368*(3), 1323–1339.

Högbom, J. (1974). Aperture synthesis with a non-regular distribution of interferometer baselines. *A&AS*, *15*, 417.

Huang, J., Zhang, T. et al. (2010). The benefit of group sparsity. *Ann. Stat.*, *38*(4), 1978–2004.

Huang, J., Zhang, T., & Metaxas, D. (2011). Learning with structured sparsity. *J. Mach. Learn. Res.*, *12*(Nov), 3371–3412.

Jain, B., & Seljak, U. (1997). Cosmological model predictions for weak lensing: linear and nonlinear regimes. *ApJ*, *484*(2), 560.

Jansky, K. G. (1933). Electrical disturbances apparently of extraterrestrial origin. *Proceedings of the Institute of Radio Engineers*, *21*(10), 1387–1398.

Jarvis, M., Bernstein, G., Fischer, P., Smith, D., Jain, B., Tyson, J., & Wittman, D. (2003). Weak-lensing results from the 75 square degree cerro tololo inter-american observatory survey. *AJ*, *125*(3), 1014.

Kaiser, N. (1992). Weak gravitational lensing of distant galaxies. *ApJ*, *388*, 272–286.

Kaiser, N. (1998). Weak lensing and cosmology. *ApJ*, *498*(1), 26.

Kaiser, N., Squires, G., & Broadhurst, T. (1994). A method for weak lensing observations. *arXiv preprint astro-ph/9411005*.

Kaiser, N., Wilson, G., & Luppino, G. A. (2000). Large-scale cosmic shear measurements. *preprint (arXiv:astro-ph/0003338)*.

Kneib, J. P., Schober, J., Gheller, C., & Schaerer, D. (2020). *The square kilometer array (ska) swiss interests and contribution* (tech. rep.). École Polytechnique Fédérale de Lausanne (EPFL).

Kneib, J.-P., Ellis, R. S., Smail, I., Couch, W., & Sharples, R. (1996). Hubble space telescope observations of the lensing cluster abell 2218. *The Astrophysical Journal*, *471*(2), 643.

Koekemoer, A. M., Aussel, H., Calzetti, D., Capak, P., Giavalisco, M., Kneib, J.-P., Leauthaud, A., Le Fevre, O., McCracken, H., Massey, R., et al. (2007). The cosmos survey: hubble space telescope advanced camera for surveys ob-

servations and data processing. *The Astrophysical Journal Supplement Series*, *172*(1), 196.

Komodakis, N., & Pesquet, J.-C. (2015). Playing with duality: an overview of recent primal? dual approaches for solving large-scale optimization problems. *IEEE Signal Processing Magazine*, *32*(6), 31–54.

Mandelbaum, R., Rowe, B., Armstrong, R., Bard, D., Bertin, E., Bosch, J., Boutigny, D., Courbin, F., Dawson, W. A., Donnarumma, A., et al. (2015). Great3 results–i. systematic errors in shear estimation and the impact of real galaxy morphology. *MNRAS*, *450*(3), 2963–3007.

Marr, J., Rogers, A., Fish, V., Wilkin, F., Arndt, M., Holodak, G., & Durkota, K. (2019). Demonstrating the principles of aperture synthesis with table-top laboratory exercises. *EPJ Web of Conferences*, *200*, 02004.

Massey, R., & Refregier, A. (2005). Polar shapelets. *MNRAS*, *363*(1), 197–210.

Massey, R., Rowe, B., Refregier, A., Bacon, D. J., & Bergé, J. (2007). Weak gravitational shear and flexion with polar shapelets. *MNRAS*, *380*(1), 229–245.

Massey, R., Stoughton, C., Leauthaud, A., Rhodes, J., Koekemoer, A., Ellis, R., & Shaghoulian, E. (2010). Pixel-based correction for charge transfer inefficiency in the hubble space telescope advanced camera for surveys. *Monthly Notices of the Royal Astronomical Society*, *401*(1), 371–384.

Meier, L., Van De Geer, S., & Bühlmann, P. (2008). The group lasso for logistic regression. *J. R. Stat. Soc. B Stat. Meth.)*, *70*(1), 53–71.

Miralda-Escude, J. (1991). The correlation function of galaxy ellipticities produced by gravitational lensing. *ApJ*, *380*, 1–8.

Moreau, J.-J. (1965). Proximité et dualité dans un espace hilbertien. *Bulletin de la Société mathématique de France*, *93*, 273–299.

Nammour, F., Schmitz, M. A., Ngolè, F., Starck, J.-L., & Girard, J. N. (2019). Radio astronomical image restoration with shape constraint. *Wavelets and Sparsity XVIII*, *11138*, 111380A.

Nocedal, J., & Wright, S. (2006). *Numerical optimization*. Springer Science & Business Media.

Onose, A., Carrillo, R. E., Repetti, A., McEwen, J. D., Thiran, J.-P., Pesquet, J.-C., & Wiaux, Y. (2016). Scalable splitting algorithms for big-data interferometric imaging in the ska era. *MNRAS*, *462*(4), 4314–4335.

Patel, P., Abdalla, F. B., Bacon, D. J., Rowe, B., Smirnov, O. M., & Beswick, R. J. (2014). Weak lensing measurements in simulations of radio images. *MNRAS*, *444*(3), 2893–2909.

Patel, P., Bacon, D., Beswick, R., Muxlow, T., & Hoyle, B. (2010). Radio weak gravitational lensing with vla and merlin. *MNRAS*, *401*(4), 2572–2586.

Patel, P., Harrison, I., Makhathini, S., Abdalla, F., Bacon, D., Brown, M. L., Heywood, I., Jarvis, M., & Smirnov, O. (2015). Weak lensing simulations for the ska. *preprint (arXiv:1501.03892)*.

Reber, G. (1940). Cosmic static. *Proceedings of the IRE, 28*(2), 68–70.

Refregier, A. (2003a). Shapelets i. a method for image analysis. *MNRAS, 338*(1), 35–47.

Refregier, A. (2003b). Weak gravitational lensing by large-scale structure. *ARA&A, 41*(1), 645–668.

Refregier, A., & Bacon, D. (2003). Shapelets ii. a method for weak lensing measurements. *MNRAS, 338*(1), 48–56.

Refregier, A., Rhodes, J., & Groth, E. J. (2002). Cosmic shear and power spectrum normalization with the hubble space telescope. *ApJ, 572*(2), L131.

Rivi, M., Lochner, M., Balan, S., Harrison, I., & Abdalla, F. (2019). Radio galaxy shape measurement with hamiltonian monte carlo in the visibility domain. *Monthly Notices of the Royal Astronomical Society, 482*(1), 1096–1109.

Rivi, M., Miller, L., Makhathini, S., & Abdalla, F. B. (2016). Radio weak lensing shear measurement in the visibility domain-i. methodology. *preprint (arXiv:1603.04784)*.

Schneider, P. (1998). Cosmic shear and biasing. *ApJ, 498*(1), 43.

Schneider, P., Ehlers, J., & Falco, E. E. (1992). Springer.

Schwarz, U. (1978). Mathematical-statistical description of the iterative beam removing technique (method clean). *A&A, 65*, 345.

Sérsic, J. (1963). Influence of the atmospheric and instrumental dispersion on the brightness distribution in a galaxy. *Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 6*, 41–43.

Setzer, S., Steidl, G., & Teuber, T. (2010). Deblurring poissonian images by split bregman techniques. *Journal of Visual Communication and Image Representation, 21*(3), 193–199.

Southworth, G. (1945). Microwave radiation from the sun. *Classics in radio astronomy* (pp. 168–181). Springer.

Starck, J.-L., Murtagh, F., & Fadili, J. M. (2010). *Sparse image and signal processing: wavelets, curvelets, morphological diversity*. Cambridge university press.

Swarup, G. (2015). Major advances in radio astronomy: some key questions today. *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences, 85*(4), 465–481.

Thompson, A. R., Moran, J. M., & Swenson Jr, G. W. (2008). *Interferometry and synthesis in radio astronomy*. John Wiley & Sons.

Thompson, A., Moran, J., & Swenson Jr, G. (2004). Interferometry and synthesis in radio astronomy, wiley-vch.

Van Waerbeke, L., Mellier, Y., Erben, T., Cuillandre, J., Bernardeau, F., Maoli, R., Bertin, E., Mc Cracken, H., Fevre, O. L., Fort, B., et al. (2000). Detection of correlated galaxy ellipticities on cfht data: first evidence for gravitational lensing by large-scale structures. *preprint (arXiv:astro-ph/0002500)*.

Van Waerbeke, L., & Mellier, Y. (2003). Gravitational lensing by large scale structures: a review. *preprint (arXiv:astro-ph/0305089)*.

van Uitert, E., Hoekstra, H., Schrabback, T., Gilbank, D. G., Gladders, M. D., & Yee, H. (2012). Constraints on the shapes of galaxy dark matter haloes from weak gravitational lensing. *A&A*, *545*, A71.

Villumsen, J. V., Freudling, W., & Da Costa, L. N. (1997). Clustering of galaxies in the hubble deep field. *The Astrophysical Journal*, *481*(2), 578.

Vũ, B. C. (2013). A splitting algorithm for dual monotone inclusions involving cocoercive operators. *Advances in Computational Mathematics*, *38*(3), 667–681.

Walsh, D., Carswell, R. F., & Weymann, R. J. (1979). 0957+ 561 a, b: twin quasistellar objects or gravitational lens? *Nature*, *279*(5712), 381–384.

Wiaux, Y., Jacques, L., Puy, G., Scaife, A. M., & Vandergheynst, P. (2009). Compressed sensing imaging techniques for radio interferometry. *Monthly Notices of the Royal Astronomical Society*, *395*(3), 1733–1742.

Wittman, D. M., Tyson, J. A., Kirkman, D., Dell'Antonio, I., & Bernstein, G. (2000). Detection of weak gravitational lensing distortions of distant galaxies by cosmic dark matter at large scales. *Nature*, *405*(6783), 143–148.

Yang, J., & Zhang, Y. (2011). Alternating direction algorithms for ℓ_1-problems in compressive sensing. *SIAM journal on scientific computing*, *33*(1), 250–278.

Zhang, S., Huang, J., Huang, Y., Yu, Y., Li, H., & Metaxas, D. N. (2010). Automatic image annotation using group sparsity. *CVPR*, 3312–3319.

Vasileios Angelopoulos

CONTACT
INFORMATION

Chemin des Condémines 1,
1587 Constantine,
Switzerland

E-mail: vasileios.c.angelopoulos@gmail.com
Skype: vasilis_angelopoulos88
LinkedIn: vasilis-angelopoulos

STRENGTHS

- **Expert in data analysis, machine learning, and signal processing**
- **Experience in software development: MATLAB, Python, C, C++, Java**
- **Skills: Ability to learn fast, analytical thinking, problem-solving expertise, focused on details and organization**

EDUCATION

2017 - 2021
Doctor of Philosophy in Electrical Engineering
École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

2013 - 2016
Master of Science in Communication Systems
École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
 - Specialization in Signals, Images, and Interfaces

2005 - 2012
Bachelor of Science in Informatics and Telecommunications
National and Kapodistrian University of Athens, Athens, Greece
 - Specialization in Communications and Signal Processing

CORE EXPERIENCE

2016 - 2021
École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
Research Assistant
I was working on a project aiming to develop, evaluate and characterize several models, revealing the underlying dark matter distribution in the Universe. Using algorithms based on the most modern imaging and machine learning techniques, I implemented solutions analyzing data from surveys performed by recently built radio telescopes. For this task, we formed a collaboration with other groups that work in closely related domains in the university and industry.

2015-2016
École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
Student Research Assistant
This project aimed to apply image processing and analysis techniques to a core problem of cosmology, the weak gravitational lensing effect. Our model is based on advanced regularization techniques and can make the required measurements without prior knowledge of the data characteristics, unlike the already developed methods, achieving a great reduction in computational resources.

2015
Hellenic General Insurance Co S.A. "THE ETHNIKI", An Affiliate of National Bank of Greece S.A., Athens, Greece
Software Engineer - Internship
I was part of a group working on image restoration and analysis to recover useful information during car accidents (e.g., registration plaques), applying pattern classification to extract useful information from the images automatically.

2011-2013	**National and Kapodistrian University of Athens**, **National Observatory of Athens**, Athens, Greece
	Student Research Assistant
	Project on materials' identification through their spectral signatures in real data from the ESA's Mars Express Satellite. We extended an existing mathematical model, providing an increase in the identification accuracy while reducing the computation time by almost 70%.

2007-2011	**Private Insurance Advisors' Office**, Athens, Greece
	Administrative Assistant
	My responsibilities included editing and managing contracts, policies, and general legal documents, collecting and analyzing policy data for statistical purposes, communicating with the customers for open issues, and coordinating with the office's partners.

ADDITIONAL
EXPERIENCE

2019-2021	**École Polytechnique Fédérale de Lausanne (EPFL)**, Lausanne, Switzerland
	Course Development & Administration
	Motivated by the important developments in this field, I have been co-developer and co-instructor of a series of MOOCs in radio astronomy (https://www.edx.org/course/radio-sky-1). I was also responsible for managing the development procedure, the editing of the required documents, the coordination with the partners in the edX platform, and the interaction with the students.

2009-2021	**National and Kapodistrian University of Athens**, **École Polytechnique Fédérale de Lausanne (EPFL)**
	Course Teaching Assistant
	During my studies, I have participated in course teaching, exam supervision, and correction for numerous courses at bachelor and master levels.

2014-2015	**École Polytechnique Fédérale de Lausanne (EPFL)**, Lausanne, Switzerland
	Semester Projects
	1. Application & interface development for image segmentation, useful to professionals in biomedical research, in collaboration with scientists from many subfields. The software is published as a plug-in on the platform's official website: http://icy.bioimageanalysis.org/plugin/Potts_Segmentation.
	2. Utilization of a specific wavelet decomposition basis to tackle the problem of precise texture segmentation in histopathological whole slide images, producing a highly efficient algorithm.

LANGUAGES

	Understanding		Speaking	Writing
	Listening	*Reading*		
English	C2	C2	C2	C2
Greek	Native	Native	Native	Native
French	A2	B1	A1	A2
Russian	A2	A2	A1	A1

TECHNICAL SKILLS
- Programming Languages: MATLAB, Python, C, C++, Java, Bash.
- Applications/Platforms: ImageJ, Icy (Bioimage Processing), LaTeX, Git, MS Office.
- Development environments: MATLAB, Eclipse, NetBeans, Microsoft Visual Studio.
- Operating Systems: Unix (Solaris), Linux (Ubuntu), Android, Windows (XP, Vista, 7, 8, 8.1, 10).

PERSONAL
INTERESTS

Reading, puzzles, music, chess, football.

148