Internship Report

CosmoStat - CEA Paris-Saclay

Weak gravitational lensing effects on the UNIONS survey data-set : systematic tests on galaxy shape measurements

---

Clara Bonini - M1 General Physics
clara.bonini@universite-paris-saclay.fr

Supervisor : **Martin Kilbinger**

May-June 2023

Abstract

Gravitational lensing is a powerful method to provide information about the distribution of matter in galaxies, clusters of galaxies and on large scales. It consists of the deflection of light by massive objects and can occur in different scales from cosmic webs to smaller objects. There are different types of gravitational lensing such as strong lensing, weak lensing and micro lensing which depend on the mass and the geometry of the lens bending the light. The main observable effects of weak lensing are the shear and the convergence which act respectively on the distortion of the image lensed shape and its magnification.

Several sources of bias can affect weak lensing shear measurements. There are for instance, the calibration biases induced by the conversion between observed image and shear distortion. The Point Spread Function (PSF) distortion is the blurring effect on galaxy images caused by various factors such as diffraction, imperfect optics, and atmospheric turbulence and can alter galaxy shapes. The PSF Leakage occurs when the PSF is not accurately modeled or corrected, which can cause errors in the shape measurements. It is then primordial to quantify and correct the leakage in the weak lensing measurements.

This report presents, after a bibliographic context, the results obtained during my Internship performed in the CosmoStat laboratory of the CEA Paris-Saclay under the supervision of Martin Kilbinger, about systematic tests on galaxy shape measurements on UNIONS survey data-set. The aim of the work was to further develop an existing code that originally computes the PSF Leakage, for measuring now the influence of observational quantities on galaxy shapes.

Contents

1 Introduction

Gravitational lensing is a phenomenon in which the gravitational fields of massive objects such as galaxies and clusters of galaxies, bend the path of photons as they propagate through the Universe. Weak gravitational lensing is a type of gravitational lensing and a technique that provides a way to measure the masses of astronomical objects without knowing the composition of these objects. [1][2] This is a powerful tool to constrain the expansion history and the evolution of the universe, in particular the cosmic webs. [3] It can also be used to study the distribution of dark matter in the Universe which helps to understand the structure of dark matter halos. [3] By measuring the shape distortion of the image of galaxies caused by the gravitational effect of dark matter, we can estimate the distribution of dark matter in halos, in particular by using the Galaxy-galaxy lensing method.[4][1][5]
Weak lensing measurements are intrinsically noise-dominated, and small systematic residual shape errors can propagate into the final result. For this reason, dedicated shape measurement pipelines are created to avoid these errors and to calibrate the distortion of the images.

This report contains the results obtained during my 2,5 months Internship between April and June at the *CEA* (Commissariat à l'Énergie Atomique) Paris-Saclay in the *Cosmostat* department under the supervision of Martin Kilbinger. Cosmostat is a laboratory specialized in Cosmology and Statistics which works on the development of statistical and signal processing methods for the analysis of Cosmological data.[6] This period was dedicated to programming systematic tests for galaxy shape measurements on UNIONS survey data-set, especially on the measure of the dependency of observational variables on galaxy ellipticities.
The section 2 of the report will present an introduction about gravitational lensing, the UNIONS survey and the goal of the internship. Then we will define in 3.2 the PSF Leakage, how we can compute it according to the data and the first results obtained. The section (3.3) will be focused on the Leakage of any observational variables, for the quantification of observational effects on cosmological results. We will present the new features of the script which originally computes the PSF leakage that we generalize for all quantities. Therefore the section will present the results obtained for different quantities in galaxy catalogues and other extracted quantities.

2 Context of the Internship

In this part, we will present the context of the Internship with a global introduction on Gravitational lensing 2.1 summarized from my Research Project report. [5] Then we will discuss the galaxy shape measurement pipelines 2.1.3 and the application on the UNIONS survey 2.2, for finally talking about the goals of the internship.

2.1 Gravitational lensing

The deflection of light by the gravity of objects in the universe is called gravitational lensing. Measurements and statistics with this method allowed to study objects such as Dark matter and exoplanets because lensing probes total matter structure in the Universe, baryonic and dark matter. [5][6] This section will summarize the main information about Gravitational lensing collected during my Research Project. [5]

2.1.1 History

The gravitational lensing effect was hypothesised in the first place by Isaac Newton in 1704 and further developed by Johann Georg Von Soldner in 1804 when he pointed out the deflection of a straight ray of light by the attraction of a massive object based on Newtonian physics. Einstein, with his theory of relativity, calculated the deflection of a light ray passing the sun in 1916 and concluded of a deflection of 1.7 arc-seconds. [1] This value will be confirmed by a total Solar eclipse in 1919 by Eddington, Davidson, and Dyson. In 1936, Rudi W. Mandl asked Einstein to calculate the gravitational lensing effects created by a star more distant than the sun and he concludes that the chance to observe lensing phenomena from distant stellar-mass lenses is unlikely.
The year after Fritz Zwicky obtained the first indication of dark matter by observations of galaxy clusters as lenses. Finally, in 1979, Dennis Walsh et al. detect the first double image of a lens quasar. [1] New technologies like the combination of CCD (Charge Coupled Device) and computers allowed the observation of, for instance, strongly distorted objects behind galaxy clusters in 1987 and the measurements of weak gravitational lensing properties as the cosmic shear in 2000. The creation of dedicated surveys increased in addition, the precision of the observation. [1]

2.1.2 General introduction of Weak Lensing

This section will present the main effects of gravitational lensing from the Research Project Report. [5]
First of all, a gravitational lens system can be schematized as the following figure 1 :

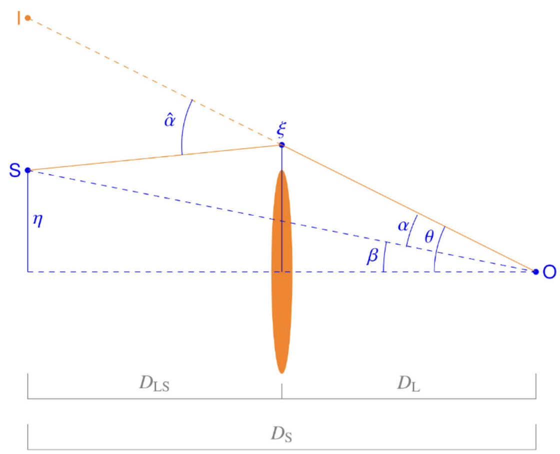

Figure 1: Scheme of gravitational lens system from [1], with O the observer, S the real position of the source, I the image of the source deflected by the lens. β is the angle between the observer and the source. $\hat{\alpha}$ is the deflection angle and α the reduced deflection angle measured by the observer

The lens equation [2][1] can be written with the convention of the figure 1 as :

$$\beta = \theta - \alpha \tag{1}$$

With β the source coordinates
θ the lens coordinates
α the deflection angle, which can also be written as the gradient with respect to the angle θ of the lensing potential ψ approximate for a point mass lens [1] :

$$\alpha = \vec{\nabla}_\theta \psi = \frac{\partial \psi}{\partial \theta} \quad with \quad \psi = \frac{4GM}{c^2} \frac{D_{LS}}{D_L D_S} \ln|\theta| \tag{2}$$

The lensing potential incorporates the properties of a gravitational lens. There are different types of gravitational lensing as strong lensing, weak lensing and micro lensing which depend on the mass and the geometry of the lens which bends the light. Strong lensing produces visible distortion, in opposite to weak lensing which produces more subtle effects and cannot be detected on individual galaxies. For studying weak lensing effects, we have to average a large number of galaxies for producing a statistical signal. [2] [6]

In the context of the Internship, we focused on weak lensing effect. The main observable effects in weak lensing are the γ shear and the κ convergence. They can be defined by using the linear derivation of the lens equation (1) :

$$\frac{\partial \beta_i}{\partial \theta_j} = \frac{\partial \theta_i}{\partial \theta_j} - \frac{\partial \alpha_i}{\partial \theta_j} = \delta_{ij} - \partial_i \partial_j \psi = A_{ij} \tag{3}$$

We can rename the matrix A_{ij} as the Jacobi matrix. This matrix can be also written as following [2] :

$$A_{ij} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} \tag{4}$$

Hence, we can observe in the matrix, the convergence κ and the complex shear $\gamma = \gamma_1 + i\gamma_2 = |\gamma| e^{i2\gamma}$. The factor 2 in the exponential term is due to the fact that shear or ellipticity is a spin-2 quantity, not a vector. [2]

2.1.3 Galaxy Shape measurement pipelines and biases

Each galaxy has an intrinsic shape or ellipticity ϵ. The ellipticity is described by two components e1 and e2, which correspond to the orthogonal direction of an ellipse. The ellipticity is related to the shear which induces the distortion of the galaxy shape due to weak lensing.

Weak lensing measurements are intrinsically noise-dominated, and small systematic residual shape errors can propagate into the final result. The observed ellipticity is in particular an estimator of the shear but it is noisy, and this noise can be reduced by doing statistics over a large sample of galaxies. For this reason, dedicated shape measurement pipelines are created to avoid these errors and parameterize the distortion of the images.

There are several methods to obtain the shape measurement of galaxies through weak lensing data. The methods can be divided into two groups: the parametric model by using the fitting method and the non-parametric model by using direct estimations and applying perturbative or non-perturbative methods. [2] [5] We will focus on the fitting method because it is the method used for the data analysed during the Internship.
There exist several model fitting methods, for instance, *lensfit* (Miller et al. 2007) is a method to estimate the galaxy size and measure the weak gravitational lensing shear. This method is interesting because it has been proven to perform ellipticity and galaxy size measurements by including in particular the PSF effects in the estimation. [2][1]
The lensfit method, like most methods in weak lensing, consists of the convolution of a model of the data with the PSF measured by the instrument for creating dithered images. Dithering is a method that consists of combining multiple exposures of the same region of the sky with small offsets to reduce the effects of pixelation and systematic errors in the measurement. [7] The lensfit method is effective when the model used for the fit is an accurate representation of the data. [8]
Another pipeline for weak lensing galaxy shape measurement is ShapePipe. It is developed in the CosmoStat laboratory and uses the multi-epoch Gaussian model fitting shear measurements Ngmix. Multi-epoch refers to the fact that the observations are taken at different times which can help to reduce the impact of systematic errors. Thus, this method can accurately measure the shear signal from the noisy, pixelised, and blurred distorted shapes of the galaxies. [9] ShapePipe is a powerful tool for processing and analyzing weak-lensing data.
During the Internship, we analysed catalogues with galaxy shapes measured with ShapePipe and Lensfit for comparing them and making systematic tests.

Moreover, several sources of bias can affect weak lensing shear measurements as the calibration biases induced by the conversion between observed image and shear distortion, the pixelation, the PSF distortion, noise bias and other biases which can be related to instruments, unknown galaxies with unknown morphologies, blending of the sources... [2][7][1] Therefore, the observed image is not the lensed image, but the result of various effects that smear the image and change the ellipticity.
Then it is primordial for each step of the pipeline to calibrate and validate the data to reach an accurate precision of extraction for cosmological analysis. [10]. For instance, the Metacalibration [11] measure the response of a shape measurement algorithm to a shear artificially apply to an image. [10] It is used for the calibration of ShapePipe, and involves distorting the image, using only the imaging data available and without needing the information about galaxies properties. It has been used for the calibration of the multiplicative bias. A basic representation of the additive (c) and multiplicative (m) bias [10] is shown in the following expression :

$$< e_{obs} >= (1+m) < \gamma_{true} > +c \tag{5}$$

This equation induces that the intrinsic ellipticity of galaxies vanishes on average. [10] The Metacalibration measure the response of a shape measurement algorithm to a shear artificially apply to an image.

2.2 UNIONS survey

UNIONS (Ultraviolet Near-Infrared Optical Northern Survey) is a collaboration between two observatories located in Hawaii : the CFHT (Canada France Hawaii Telescope) and the Pan-STARRS (Panoramic Survey Telescope and Rapid Response System), and WHISHES (Wide Imaging with Subaru HSC of the Euclid Sky) which collected the data over 4800 square degrees which is the largest multi-band optical photometric survey of the Northern Hemisphere. [10] [12] The CFHT with the MegaCAM built by the CEA, started in 2017 and takes images in the r-band (640 nm) and u-band (355 nm), this part of the survey is called CFIS (Canada-France Imaging Survey) and provides a state-of-the-art weak lensing survey. Moreover, CFIS area overlaps with other wide surveys such as SDSS-BOSS, eBOSS and DESI [10] which together provide a strong data set for weak lensing analysis.

2.3 Goal of the Internship

The goal of the Internship is to check for systematic errors in the weak lensing data against various observational variables to detect undesired dependencies in the catalogue. For this, we generalize the PSF Leakage principle (cf. 3.2) from an existing code developed by the laboratory to any quantities in a catalogue, by implementing new features in the code. (cf. 3.3) The language used for programming was Python and for large scale technical computing, we used the data remote cluster from IAP (Institut d'Astrophysique de Paris), CANDIDE. The aim was also to extract new quantities for the leakage that are not in the catalogue, from an ID associated with each galaxy representing a region of the sky survey and extract interesting values from associated files.

3 Methods and Results

In this part, we will present the methods and the results obtained during the Internship. We will first focus on the presentation of the catalogues used for the data analysis and the PSF Leakage principle. Then, we will discuss the generalization of the Leakage for any quantities, the technique of extraction for finding more quantities for the leakage in galaxy image files and the results obtained from these quantities.

3.1 Presentation of the catalogues

During this Internship, the type of data set used for the analysis was a catalogue that contains the measured shapes of 100 million galaxies in a `.FITS` (Flexible Image Transport System) file provided by ShapePipe and Lensfit pipelines on the UNIONS/CFIS data. (cf 2.1.3).
The UNIONS shapepipe catalogue has been studied globally but also by focusing on the P3 Patch which represents a part of the sky collected by the UNIONS survey. Each catalogue contains information on the deformation caused by the gravitational lensing on galaxies :

QUANTITY	DESCRIPTION
RA (deg)	Right Ascension (RA) : astronomical coordinate that specifies the angular distance of a particular point measures eastward from the equator. (celestial equivalent of the terrestrial longitude)
Dec (deg)	Declination (Dec) : astronomical coordinate that specifies the angular distance of a particular point measures northward or southward from the equator. (celestial equivalent of the terrestrial latitude)
e1	Shear/ellipticity component 1
e2	Shear/ellipticity component 2
SNR	Signal to Noise Ratio : Measure of the strength of the lensing signal relative to the background noise [13], an higher SNR indicates a better quality signal
w	Weight : parameter assigned to each galaxy which indicates their contribution in the weak lensing analysis. It is computed based on the SNR of the galaxy and it will down-weight galaxies with SNR for avoiding noise biases in the weak lensing analysis.
mag	Magnitude : parameter used to describe the brightness of galaxies
$e1_{PSF}$	PSF ellipticity component 1
$e2_{PSF}$	PSF ellipticity component 2
$fwhm_{PSF}$ (arcsec)	Full Width at Half Maximum of the PSF is the measure of the width of the PSF of an image. Also called PSF size, it can be used to measure the resolution of the image.

Table 1: Usual components of a weak lensing extended catalogue

3.2 PSF Leakage

PSF Leakage is a type of systematic error that can affect weak lensing measurements. It is caused by the Point Spread Function (PSF), which is the response of the instrument to a point source that corresponds to the blurring effect on galaxy images, and its ellipticity can affect the accuracy of measurements of the shapes and orientations of astronomical objects, such as galaxies. [2] It is caused by various factors such as optical distortions and atmospheric turbulence. [10] PSF ellipticity (e_{PSF}) refers to the degree to which the point spread function (PSF) of a telescope or an instrument deviates from a circular shape. The PSF ellipticity was measured at the position of galaxies averaged over the contributing single exposures. This means that the ellipticity induced by the point spread function (PSF) was measured for each galaxy in the image, and the measurements were averaged over all the individual exposures that contributed to the final image. [10]

The modeling of the PSF needs a selection of pure and homogeneously distributed stars. Stars are bright, numerous and with different colors which makes them ideal for modeling the PSF model of an image calibrated for different observing conditions. [10] In the case of ShapePipe, the star selection was performed at the single-exposure level because of the spatial stability of the distribution of the stars. [10] After selecting the stars by removing the saturated, faint or noisy objects with `SExtractor` tool (Bertin & Arnouts 1996), the PSF model can be estimated by two methods. The first method is by using `PSFEx` (Bertin 2011) which produces independent models with polynomial variations of the PSF for each CCD (Charge-Coupled Device which is an highly sensitive photon detectors) [9]. The second method used the MCCD (Multi-CCD) technique which computes the PSF model for all the CCD simultaneously. [9]

The corrected galaxy ellipticity can be written as a function of the PSF Leakage :

$$e^{gal}_{obs} = (1+m)\gamma_{true} + \alpha e^{PSF} + c \tag{6}$$

With e^{gal} the observed ellipticity of the galaxy, γ the true shear, α the Leakage and c the additive bias. A correlation i.e. the measured slopes of the linear fit between galaxy ellipticity (in y axis) and PSF ellipticity (in x axis) is called α and corresponds to the PSF Leakage. For the validation of the null test, the slope has to be close to 0. We estimated PSF Leakage of the UNIONS catalogue by using the `leakage_object.py` script of the `sp_validation` package.

3.2.1 Presentation of `leakage_object.py` script

The `leakage_object.py` script is a Python script from the `sp_validation` package developed by the CosmoStat laboratory. The main use of this package is the validation of the ShapePipe catalogue by running diagnostic tests such as the PSF Leakage. Figure 2 is a flowchart of the operations performed during the execution of the leakage_object.py scripts.

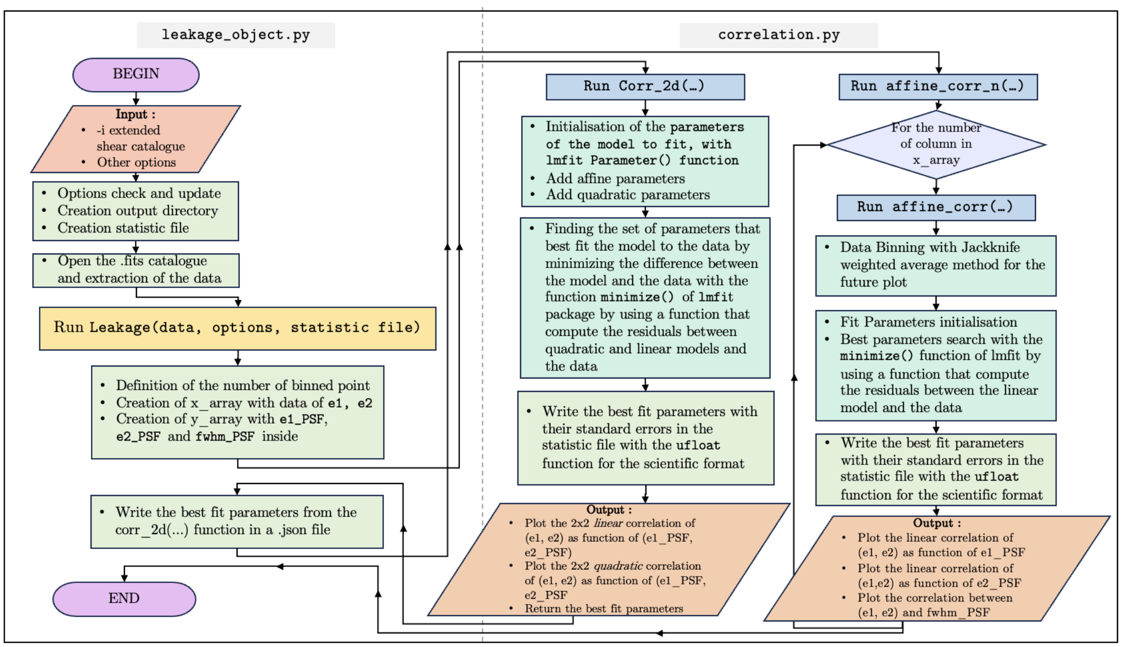

Figure 2: Algorithm flowchart of the leakage_object.py script

For running the script, the user needs to type a command line in the terminal with different options. The several options implemented to the script with the function `OptionParser`, are summarized in the table 2 below. The (string) argument means that following the option in the command line should appear a sequence of characters.

OPTION	DESCRIPTION
-i (string)	input path of the extended shear catalogue (mandatory)
-o (string)	label of the output directory
--e1_col (string)	for changing the e1 column name in the galaxy catalogue
--e2_col (string)	for changing the e2 column name in the galaxy catalogue
--e1_PSF (string)	for changing the e1_PSF column name in the galaxy catalogue
--e2_PSF (string)	for changing the e2_PSF column name in the galaxy catalogue
--size_PSF_col (stirng)	for changing the fwhm of the PSF name in the galaxy catalogue
-s (string)	changing the name of the shape measurement method
-v (store_true)	verbose output
-t (string)	test of 2D fit

Table 2: Options description of the leakage_object.py script

The mandatory option for well running the code is `-i <path/catalog.fits>` which corresponds to input the extended shear catalogue that contains at least the two components of galaxy ellipticity e1 and e2, the two PSF ellipticities components e1_PSF and e2_PSF and the FWHM of the PSF. After typing the command line with the different options, the script will check and update the input options which creates a variable "`param`" which contains all the options input.
The script will after create an output directory where the plots and the files will be stored and a statistic file that will handle the best fit parameters.
We can now extract the data from the input catalogue with the function `fits.open()` from `astropy` package as we can see in the example below, which returns an `HDUlist` (Header Data Units list) which is in general composed with a header and a data array table.

```
hdu_list = fits.open(param.input_path_shear)
data = hdu_list[1].data
```

After extracting the data, the script will run the `Leakage()` function which takes in arguments the data extracted, the options parameters set in the command line and the statistic file handler.

`Leakage()` **function :** Compute and plot object-by-object PSF Leakage relation. This function computes the PSF Leakage for the input path catalogue with the help of other sub-functions contained in the `correlation.py` library. We first set the number of bins. Binning is a data visualisation technique to reduce the number of data points in a plot. A full catalogue contains 100 million galaxies, thus binning is necessary to be able to interpret a scatter plot. For this, we have to divide the range of the data into sets of intervals called bins. For each bin, we then compute a jackknife weighted average by using the jackknife method (M. Quenouille, 1949). The jackknife method consists in omitting one data point at a time and recalculating the average for estimating the uncertainties in the average. The jackknife weighted average is then computed by taking the average of the weighted averages. The number of bins will determine the number of points in the plot.
After setting the number of bins, we define an array containing the quantities which will be in abscissa and ordinate. For the PSF Leakage, the array in abscissa will contain the two components of ellipticity of the PSF (e1_PSF and e2_PSF) and the FWHM of the PSF (fwhm_PSF) and the array in ordinate will contain the two components of ellipticity of galaxies (e1 and e2). Thereafter, the function will run the sub-function `corr_2d()` of the `correlation.py` library in the `sp_validation` package which contains several functions to deal with correlation methods such as the PSF Leakage.

1. `corr_2d()` **function :** Compute and plot 2D linear and quadratic correlations of (e1,e2) as a function of (e1_PSF, e2_PSF) and return the best fit parameters computed with the lmfit package functions. (cf. fig 2) We first initialise the parameters of the model to fit with the function Parameters() of lmfit which is used to create a collection of parameters used for the fit. Therefore, we add linear and quadratic parameters to the parameter collection variable, with a value of 0 for the initialization. The linear and quadratic fit models are represented in the equation 7 below :

$$y_{model} = mx_{data} + c \qquad y_{model} = q^2 x_{data} + mx_{data} + c \tag{7}$$

The linear and quadratic parameters correspond respectively to m and q with also c for the additive bias in the equation 7.

Thereafter, the next step is to find the best fit parameters to the models with the `minimize()` function of the lmfit package. We use as argument of `minimize()`, a function that computes the residuals, which means the difference between the models and the data. The `minimize()` function will optimize the residuals with the least square method. This method works by minimizing the sum of the squares of the differences between the observed data and the predicted values from the model. The objective is to reduce as much as possible the sum of the squared errors. The minimization takes into account the weight of each galaxy.
After obtaining the best fit parameters, the script will write them in the `.txt` statistic file handler and in a .json file. The function plots at the end, the 2x2 linear and quadratic correlation between the galaxy ellipticity and the PSF ellipticity in two different plots with 4 subplots distributed in a 2x2 mosaic such as figure 3 or figure 5.

2. `affine_corr_n()` and `affine_corr()` functions : After computing the 2D correlation, the Leakage() function will fit a separate linear 1D model and will plot the linear correlation between galaxy ellipticities and PSF ellipticities. It will also computes the correlation between galaxy ellipticities and FWHM of the PSF with the functions `affine_corr_n()` and `affine_corr()`.
The function `affine_corr_n()` will run `affine_corr()` for the n elements in x_array which means e1_PSF, e2_PSF and fwhm_PSF in our case. The `affine_corr()` will first bin the data with the same method as `corr_2d()` function, with the jackknife weighted average method. The function will find the best fit parameters of a linear model as the equation 7 and with using the `Minimize()` function (cf. corr_2d section). For the element of the x_array given in argument, two linear fits are performed with the data of the two components of the galaxy ellipticity.

After writing the best fit parameters in the statistic handler file, the function will plot the two galaxy ellipticity components e1 and e2 as a function of the PSF ellipticity e1_PSF and another plot with (e1,e2) as a function of e2_PSF with the two linear fits for each plot. It also plots e1 and e2 as a function of the fwhm_PSF and the two fits associates for measuring the dependencies between the quantities. (cf. fig 4 and fig 6)

Therefore for an input extended shear catalogue, we have in output :

- Two plots that show the 2D correlation of the 2-spin galaxy ellipticity components as a function of PSF ellipticity with the best quadratic and linear fit.
- Two 1D plots of the 2-spin galaxy ellipticity components as a function of each PSF ellipticity component.
- The correlation between the 2-spin galaxy ellipticity components and the FWHM of the PSF in a 1D plot.

3.2.2 Results

This part will present the result plots obtained for different catalogue with different shape measurement pipelines (ShapePipe and Lensfit) with the `leakage_object.py` script.
We obtain the following plots for the ShapePipe pipeline :

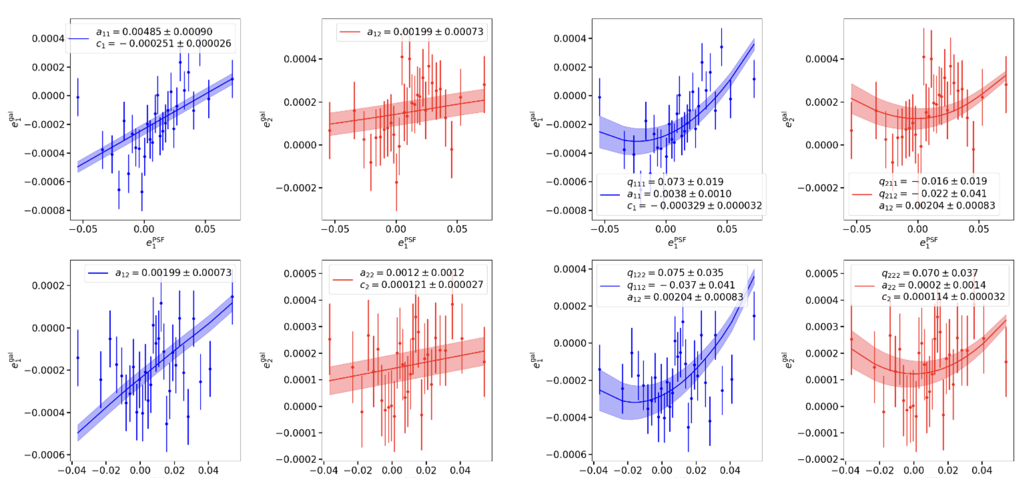

Figure 3: 2D correlation plots for the ShapePipe, catalogue of the mean galaxy ellipticity for 2-spin ellipticity components as a function of PSF ellipticity (e1 left and e2 right), the solid lines show the best linear fits (linear left and quadratic right) without binning with q being the second order parameter (for the quadratic fit), m the slope and c the additive correction

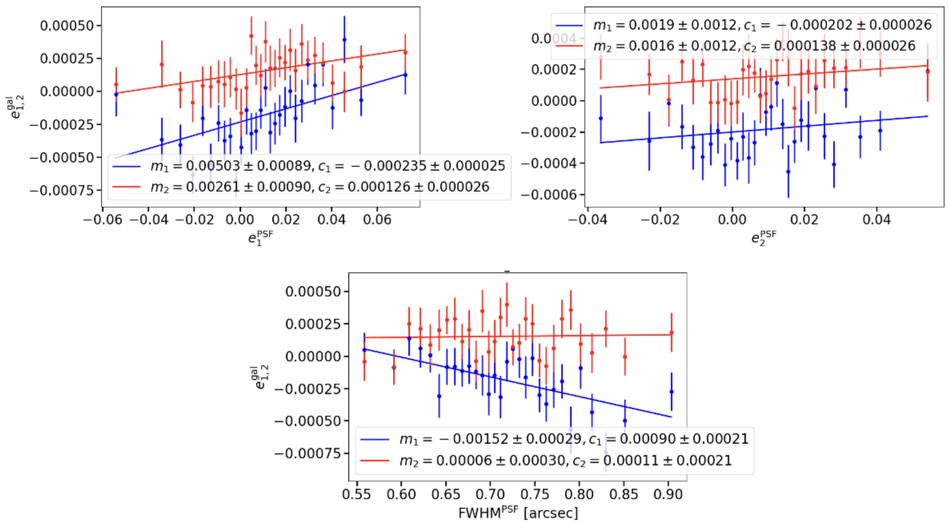

Figure 4: 1D correlation plots for the ShapePipe, catalogue of the mean galaxy ellipticity for 2-spin galaxy ellipticity components as a function of PSF ellipticity (e1 up left and e2 up right) and the FWHM of the PSF (bottom), the solid lines show the best linear fits without binning with m the slope and c the additive correction

Figure 3 presents the 2D plots for a linear and quadratic fit model for the PSF Leakage with in y-axis e_i^{gal} and in x-axis e_i^{PSF} (i=1,2). The linear fit seems to work better with the data because of the error associated to the parameter. As the PSF leakage is a null test, the expected results are no correlation and a slope close to 0. However, we can observe a correlation between the e_1^{gal} and the two components of the PSF ellipticity (2 plots on the first column in the left figure of fig (3). The leakage tends to increase with high values of the PSF ellipticity. The second component shows a weaker correlation.
Figure 4 presents 1D plots of the PSF Leakage with in the x-axis both components of the galaxy ellipticity and in the y-axis, one component of the PSF ellipticity or the FWHM of the PSF ellipticity. We can observe a slope $|m| <$ 1% for the two cases of the PSF ellipticity. Despite a slight correlation, the dependency can be considered negligible for this order of magnitude. However, for the correlation between e2 and the FWHM, $|m_1|$ is significantly greater than 0 and shows a strong correlation. The leakage tends to decrease for large values of the FWHM of the PSF. But the absolute slopes are less than 1% this effect could be neglected.
Therefore, the PSF Leakage test for the ShapePipe catalogue shows some correlation for some cases such as the first component of the galaxy ellipticity with the first component of the PSF ellipticity and the FWHM. But with the order of magnitude we can consider that the correction brought to the galaxy ellipticity is efficient enough.

We also run the PSF Leakage for the Lensfit catalogue and got the output plots in figure 5 and 6.

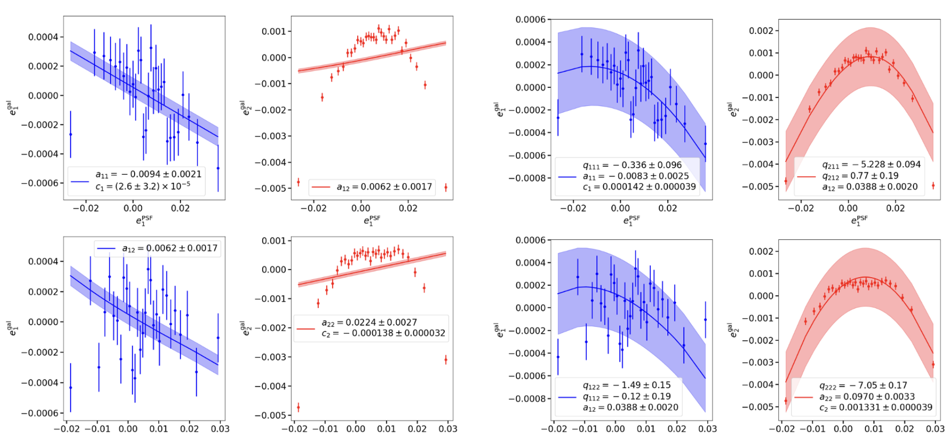

Figure 5: 2D correlation plots for the Lensfit catalogue, for the mean galaxy ellipticity for 2-spin ellipticity components as a function of PSF ellipticity (e1 left and e2 right), the solid lines show the best linear fits (linear left and quadratic right) without binning with q being the second order parameter (for the quadratic fit), m the slope and c the additive correction

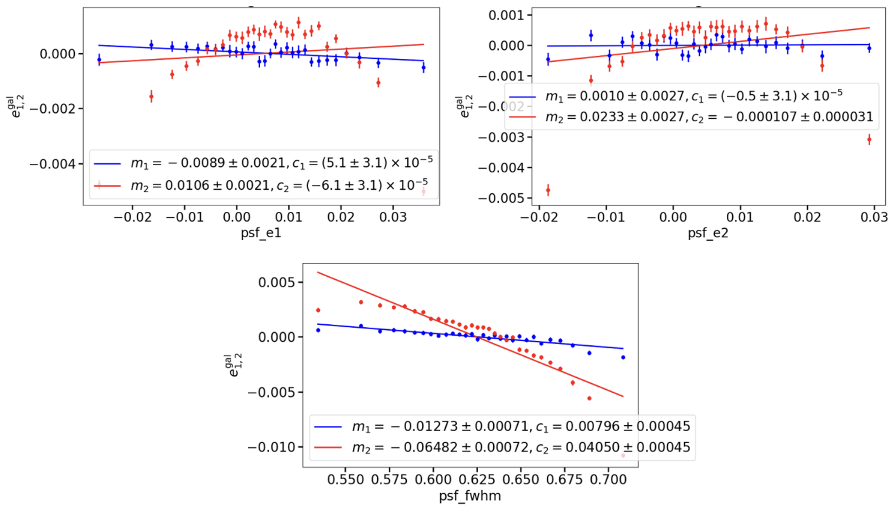

Figure 6: 1D correlation plots for the Lensfit catalogue, of the mean galaxy ellipticity, for 2-spin ellipticity components as a function of PSF ellipticity (e1 up left and e2 up right) and the FWHM of the PSF (bottom), the solid lines show the best linear fits without binning with m the slope and c the additive correction

In Figure 5 which shows the 2D PSF Leakage for each component, we can observe by comparing the two fit methods that the quadratic fit is more accurate for the PSF Leakage. We can indeed observe for the second component of the galaxy ellipticity e2 a strong quadratic correlation with no data deviation. For the first component of the galaxy ellipticity e1, we can in addition see that the leakage decrease for large values of PSF ellipticity in both linear and quadratic fits. In addition, in figure 6 we can also observe that the linear fit between e2 and the FWHM of the PSF shows a strong dependency around 7%. The correlations can be due to errors in the PSF model which have to adjusted. Therefore, the values of galaxy ellipticities for the Lensfit catalogue have to be corrected in consideration to avoid errors in the cosmological analysis.
Thus, the shape measurements of the UNIONS ShapePipe catalogue have a better correction for the PSF ellipticity then the Lensfit catalogue.

3.3 Observational variables Leakage

After computing the PSF Leakage for the PSF ellipticity, the aim is to measure the dependencies between the galaxy ellipticity and various Observational variables. For this, we edited the `Leakage_object.py` script to generalize the Leakage for any quantity. This program still can compute the PSF Leakage with a dedicated command line option or by choosing manually the PSF ellipticity as the quantity for the Leakage. As we can see in figure 8, we obtain for ShapePipe the same result as the PSF Leakage shown before in the section 3.2.

This section will present the operation of the new features added in the script and the results obtain for various quantity present in ShapePipe and LensFit catalogues. Then, for the computation of the leakage for more quantities, we create a new script that can extract values from a text file associated with the CFIS tile image or the header of the exposure image file. We will discuss the operation of this script and the results obtained combined to `leakage_object.py`.

3.3.1 New features in `Leakage_object.py` script

The generalization of the Leakage for any quantity corresponds to measuring the correlation between the components of the galaxy ellipticity and any quantity for checking any undesirable dependency. Then, the script `leakage_object.py` presented in the section 3.2.1 has been edited. (cf. section 3.2.1) First, several new options have been added to the script. Then, functions which perform 1D fit in the original script have been adapted for any quantities with linear and quadratic models. A feature for computing a summary plot has been added which presents the n slopes as a function of the label of the quantity for the linear model. A quadratic summary plot has also been added by comparing the second-order parameter q and the first-order parameter m (eq 7) as a function of the label of the quantities. This section will focus on the operation of the script with the new features presented in a flowchart in figure 7. The `correlation.py` library has been also updated and all the functions regarding the leakage have been transferred to a new library file named `leakage.py`.

First of all the following tables (tab 3,4, 5, 6) will present the new command line options implemented in the script. The (string) argument means that following the option in the command line should appear a string such as the section 3.2.1 and the (store_true) means that if the option is input in the command, it will return true for the parameter associate with this option.

OPTION	DESCRIPTION
`--RA` (string)	for changing the RA column name in the galaxy catalogue
`--Dec` (string)	for changing the Dec column name in the galaxy catalogue
`--mag` (string)	for changing the mag column name in the galaxy catalogue

Table 3: New Options for changing the column names of the input galaxy catalogue in the leakage_object.py script

OPTION	DESCRIPTION
`--PSF_Leakage` (string)	for performing a PSF Leakage of the input catalogue
`--Obs_Leakage` (string)	for performing Leakage for any quantity in the catalogue

Table 4: New Options for choosing the type of Leakage in the leakage_object.py script (at least one of the two options is mandatory)

OPTION	DESCRIPTION
`--header` (store_true)	Run interactive session : print the header of the data catalogue and allow to choose the variables for the Leakage by typing them in the terminal. Without this option, will perform the leakage for default variables put in an array.
`--linear` (store_true)	fit option for performing a linear fit between galaxy ellipticity and each chosen quantity. (default fit)
`--quadratic` (store_true)	fit option for performing a quadratic fit between galaxy ellipticity and each chosen quantity

Table 5: New Options for `--Obs_Leakage` in the leakage_object.py script

OPTION	DESCRIPTION
`--ratio` (store_true)	will compute the Leakage of a ratio between two column will continue the script for other quantities in the catalogue (if --header option will ask the user to input new quantities for the leakage or add the default quantities for the leakage)
`--ratio_alone` (store_true)	will compute the Leakage of a ratio between two column and end the program
`--ratio_label` (store_true)	change the label of the ratio for the plot ticks (can be long in the plot compared with the other names)

Table 6: New Options regarding the Leakage for a ratio between two columns in the catalogue in the leakage_object.py script

These new options allow the user to choose the type of Leakage test he wants to perform on the catalogue for validation among the two options --PSF_Leakage and --Obs_Leakage (cf. tab 5). Then, the user will input the extended galaxy catalogue such as the original script and the choice of leakage and the different other options for running the script properly. The script will first check and updates the options and extract the data as the section 3.2.1. Depending on the option --PSF_Leakage and --Obs_Leakage input by the user, the script will run differently. (cf. fig 7).

- **PSF Leakage** : Run the `PSF_leakage()` function which correspond to the `leakage()` function in the original script describe in section 3.2.1

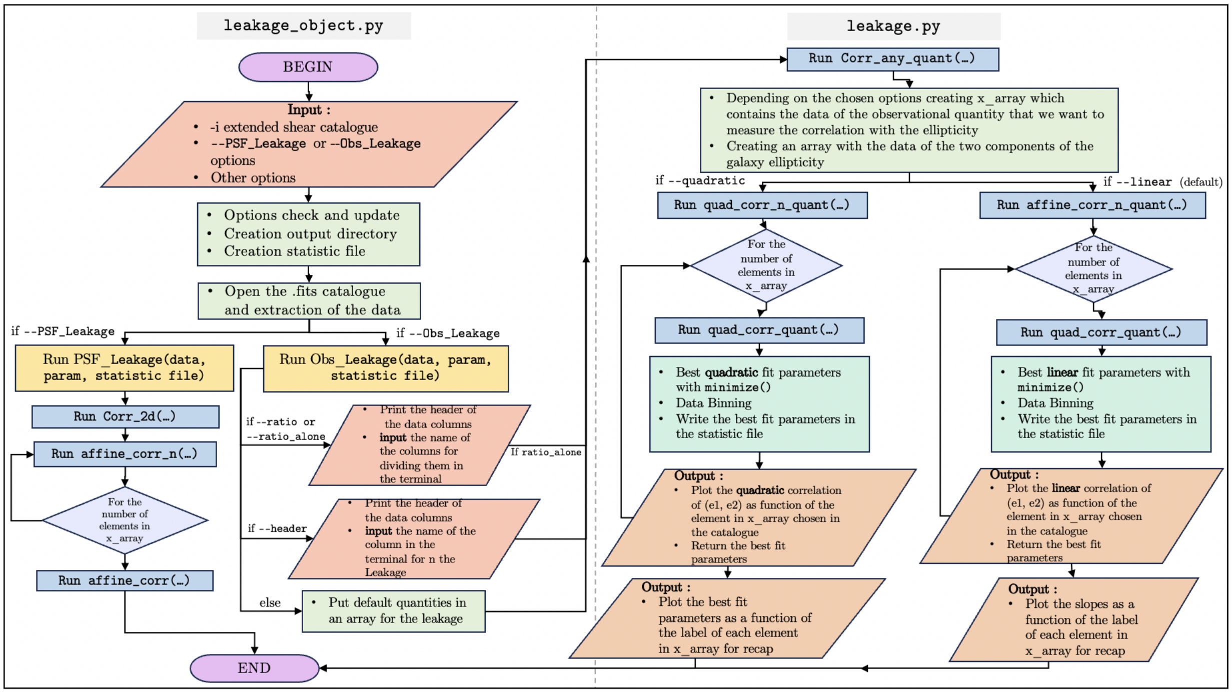

Figure 7: Algorithm flowchart of the leakage_object.py script with the new features added for the generalisation of the Leakage for any quantity

- **Observational Leakage** : Runs the `Obs_leakage()` function. The behavior of this function depends on the options input by the user. If the user needs to compute the leakage for a ratio of two columns, the option `−−ratio` or `−−ratio_alone` is input. The difference between the two options is for the `−−ratio_alone`, the script will measure the correlation only for the ratio of the two columns and the galaxy ellipticity. It will end the script after this task. On the opposite the `−−ratio` will ask the user if there is more quantity to add for computing the Leakage if he inputs the `−−header` option or will add the default quantities for the leakage.
  If there are no ratio options but the `−−header` option, the label of the data columns will be printed and the user has to enter the label for the leakage in the terminal. If none of these three options have been input, the leakage will be performed for default quantities.
  Therefore, depending on these different options the sub-function `corr_any_quant()` (cf. fig (7)) will put different quantities in an array (x_array). In another array will be stored the data of the two components of the galaxy ellipticity.
  Thereafter, depending on the fit options (tab 5) chosen by the user (`−−quadratic` or `−−linear`), (if no fit option given, the default fit will be linear) the `corr_any_quant()` function, will compute the leakage by correlate with a linear or quadratic best fit for the n quantities of the x_array, with the fits functions `affine_corr_n_quant()` and `quad_corr_n_quant()` from the `leakage.py` library.
  - `affine_corr_n_quant()` : edited version of the function `affine_corr_n()` (cf. section 3.2.1 and fig 2) which now is adapted for any quantity (not just for PSF ellipticity) and can extract the slopes with the errors and the labels associate from the function `affine_corr_quant()` which compute it for each quantity in x_array. It will also plot a summary plot of the slopes as a function of quantities labels for easier comparison between the slopes. (cf. for instance, fig 10)
  - `affine_corr_quant()` : edited version of the function `affine_corr()` which computes the best linear fit for each quantity for all the values with the `minimize()` function (cf. section 3.2.1 and fig 2) and plot the 1D binned ellipticity of galaxies (e1,e2) with the jackknife weighted average method (section 3.2.1 as a function of the quantity selected (cf. fig 8, 9 with their associated linear fits and return the slope, the error and the label of the quantity for the summary plot.
  - `quad_corr_n_quant()` and `quad_corr_quant()` : edited version of the linear fit functions presented before : `affine_corr_n_quant()` and `affine_corr_quant()` but for a quadratic fit model. It will plot the 1D binned ellipticity of galaxies (e1,e2) with the jackknife weighted average method (section 3.2.1 as a function of the quantity selected (cf. fig 8 with the associated quadratic fit, (cf. fig 11) and return the second order (q) and first order (m) fit parameters the error and the label of the quantity for the summary

plot. The summary plot will contain the two parameters as a function of the quantity label for an easy comparison of the values of the parameters between the quantity.

Therefore, the new version of the script, for an input extended galaxy catalogue and the different options selected by the user, will plot the PSF Leakage as the original script or the Leakage for any quantity (or ratio between two quantities) in the catalogue. It will plot the binned two components of the galaxy ellipticity (e1,e2) as a function of the quantities selected in the catalogue with the associated linear or quadratic plot for measuring the dependency between the two quantities. It will also, after extracting the best fit parameter, plot these parameters as a function of the quantities labels for comparing the parameters for each quantity.

3.3.2 Results

This section will present the plots obtained by using the new features of the `leakage_object.py` script for several catalogues and for different type of fit.

3.3.2.1.a Linear Leakage with ShapePipe First, we applied the Leakage script for the PSF ellipticity (with the `--PSF_leakage()` options of the section 3.3.1) for checking if the original and the new version of the script give the same results. It is the case as shown in figure 8 and 9 comparing to figure 4.
Thereafter, we used the new features of the script for computing the **linear** Leakage for several quantities (with the `--Obs_leakage()` options of the section 3.3.1) in the ShapePipe catalogue which have already been corrected, in the aim to check if correlations are null and if we need to still correct undesired dependencies for the accuracy of the galaxy ellipticity.

We first measured the leakage for the astronomical coordinates the Right Ascension (RA) and the Declination (Dec) (tab 1, fig 8), which a correlation can highlight an environmental dependency. Indeed, ground-based telescopes such as the telescope of the UNIONS survey have disadvantages because of the environment as atmospheric effects, can affect the image quality and then induce biases in the data. The linear correlation between the RA and the ellipticity components in figure 8 with the x axis the RA and the y axis the two galaxy ellipticity components, gives slopes of an order of magnitude of 10^{-6}. We can therefore consider the correlation between the two quantities negligible. Same conclusion for the declination which also has a leakage of an order of magnitude $10^{-5} - 10^{-6}$. However, we can observe that the leakage for the first component increases for low values of the declination and large values for the RA. The atmospheric effect can explain the correlation in the plots between the e1 component and the astronomical coordinates in fig 8. We can observe that m_1 is significantly greater than 0, which can be related to the airmass that we will study in the section 3.3.3.
Another explanation independent of the airmass, could be linked to the telescope pointed more toward the horizon which may affect the image quality.

After measuring the leakage of the astronomical coordinates, we will focus on the leakage of the magnitude, the SNR and the weight (cf. tab 1). The SNR and the weight are measures of the accuracy of the data, these quantities should give a null correlation with the ellipticity. We can observe in the output plots in figure 9, that the leakage for both components is at the order of magnitude of 10^{-6} for the SNR. However, the leakage increases towards low SNR which is expected since for those galaxies, shape measurement is more difficult because of the noise.
For the weight, the slope is at the order of 10^{-5} for e1 which we can consider as negligible. The second slopes related to e2 and the weight show a stronger correlation of 10^{-3}. We can see in the plot that the linear fit model does not correspond to the distribution of the binned points which can explain a greater leakage. The magnitude is used to describe the brightness of galaxies and has no apparent relation with the galaxy ellipticity. A potential correlation can be due to the large size of bright galaxies which could introduce a bias since there are some edge effects from the finite size of postage stamps (section of an astronomical image that isolates a specific object for detailed study as a galaxy). The leakage of the magnitude in figure 9 is negligible for the second component (10^{-5}) but the first component has a stronger leakage of 10^{-3}, but as the first component of the weight leakage, we can observe that the weight is not fit accurately with the data for a linear model which can explain a greater leakage value.

The summary plot in figure 10 gives an easy way to compare the leakage for all the quantities selected with their errors. It was inspired by M. Gatti et al. article for the analysis of the DES survey [14]. We can observe that all the absolute measured slopes are less than 1% which makes them negligible and validate the systematic tests for this catalogue. We can observe anyway that the PSF leakage is the greater leakage compared to the other observational quantities close to 0. This is due to the strong relation between PSF ellipticity and galaxy ellipticity because PSF ellipticity is a critical error source in galaxy ellipticity measurements. But the correction for the PSF ellipticity in this catalogue is efficient according to these results.

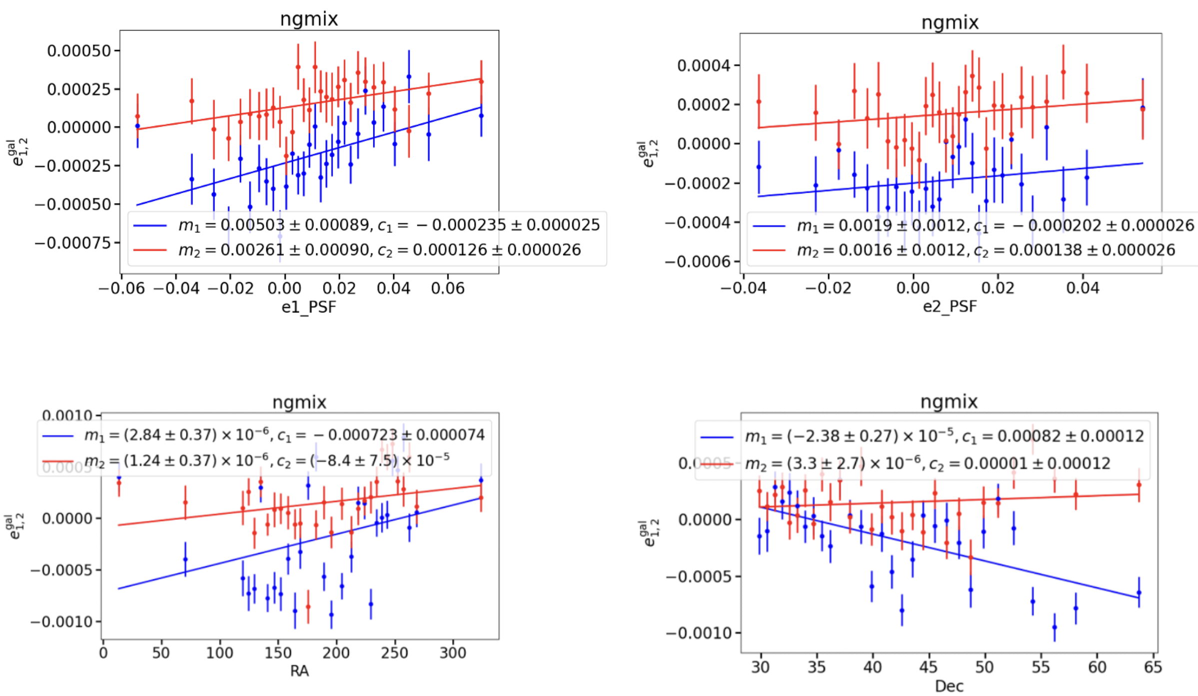

Figure 8: 1D plots of the mean galaxy ellipticity for 2-spin ellipticity components as a function of various quantities : e1, e2, RA, Dec, (cf. table 1 of the section 3.1) the solid lines shows the best linear fits without binning with m the slope and c the additive correction

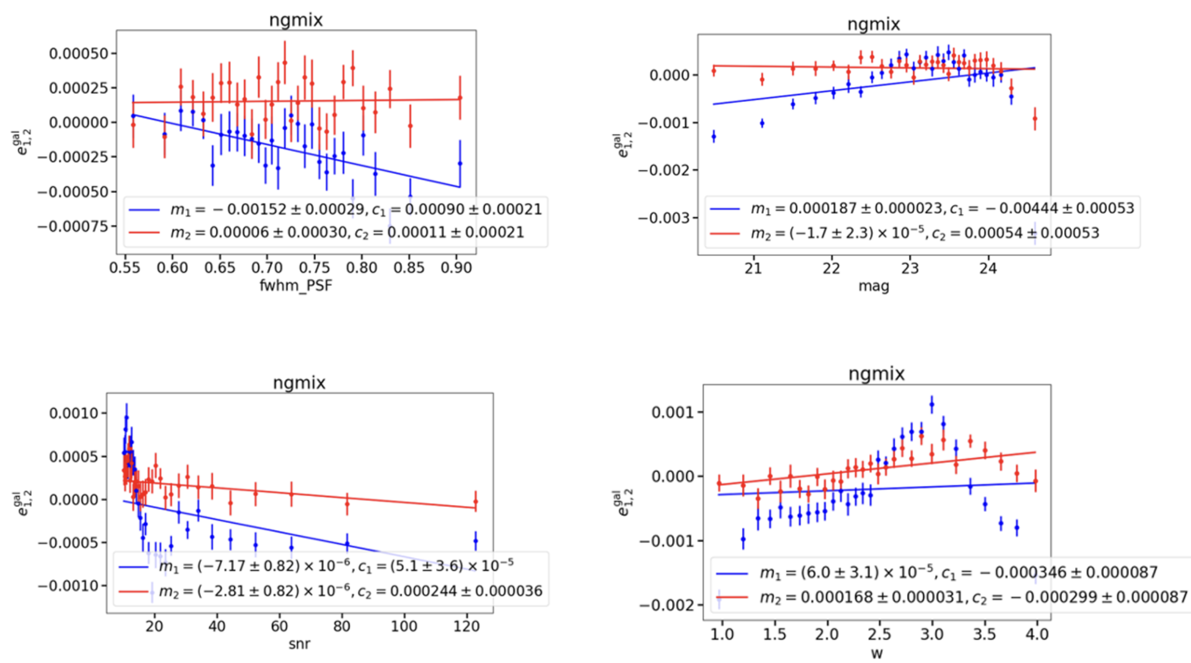

Figure 9: 1D plots of the mean galaxy ellipticity for 2-spin ellipticity components as a function of various quantities : $fwhm_{PSF}$, mag, snr, w, (cf. table 1 of the section 3.1) the solid lines shows the best linear fits without binning with m the slope and c the additive correction

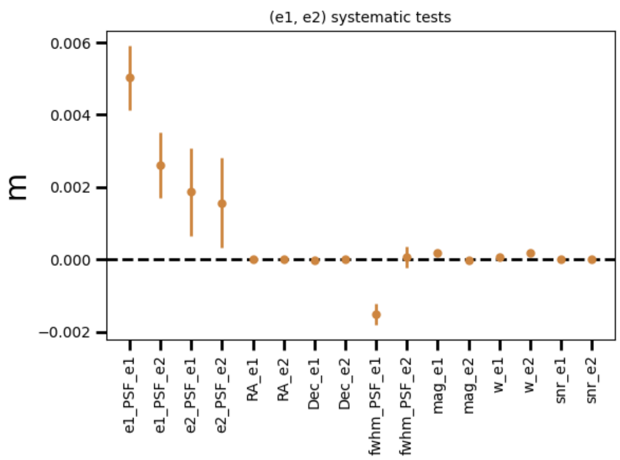

Figure 10: Summary of the slopes extracted m for all the quantities computed in the Figure 8 and 9

3.3.2.1.b Quadratic Leakage with ShapePipe For the case of the magnitude where the linear fit seems not accurate, we tried to perform the Observational leakage with a quadratic model, with the −−quadratic option of the script.

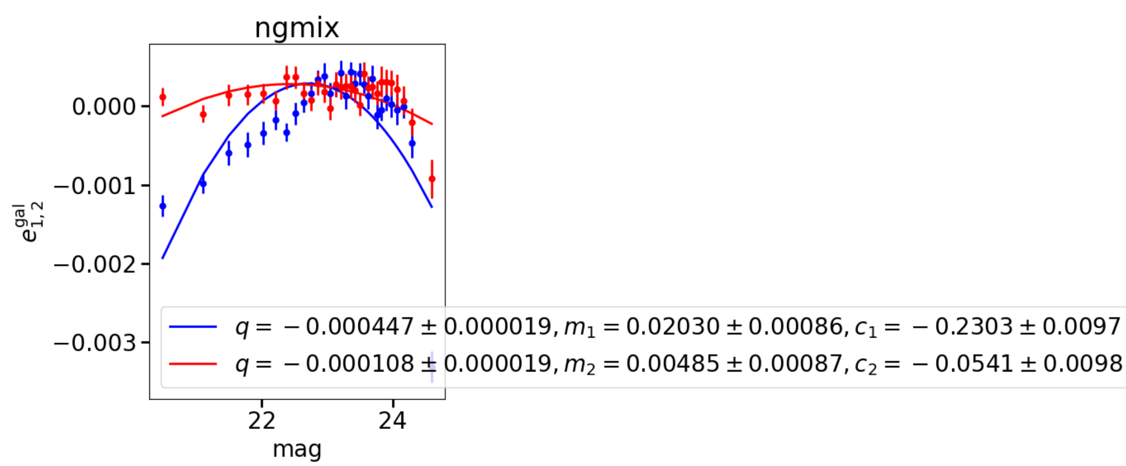

Figure 11: 1D plot of the mean galaxy ellipticity for 2-spin ellipticity components as a function of the magnitude, the solid lines shows the best quadratic fit without binning with q being the second order parameter, m the first order parameter and c the additive correction

The quadratic fit without binning in the figure 11 shows a correlation between the first component of the ellipticity e1 and the magnitude (blue solid line). But the second order parameter has an order of magnitude of 10^{-3} which is close to 0 but still significant when looking at the error bar. However we can observe that the blue points deviate from the quadratic model, it is therefore not a good fit and then we can't conclude a correlation between the two quantities.

3.3.2.2 Linear and quadratic Leakage with Lensfit We now perform the same steps but for the lensfit catalogue. We obtain the following plots in figure 12 for the linear fit model for many quantities but less than ShapePipe because the Lensfit catalogue doesn't have all the quantities. We can observe that for the astronomical coordinates, the leakages have an order of magnitude of $10^{-5} - 10^{-6}$ (fig 12) which can consider negligible. We can see, as for ShapePipe, a stronger correlation for the magnitude but the linear fit model does not correspond to the data which could be better with the quadratic model. The leakage for the weight is also negligible as we can see the slope is at the order of 10^{-5} for the first component and 10^{-3} for the second component. However the leakage for the second component of the weight seems increasing toward large value of the weight.
We also compute the leakage for a quadratic model of the PSF ellipticity in figure 14, which is a summary of the information given by the 2D plot of the PSF ellipticity in the section 3.2.2. It shows again the strong correlation

of the PSF Leakage for the quadratic model for the second components. We thereafter compute the leakage for the magnitude in figure 14. The figure shows a quadratic correlation but with a second order parameter close to 0. The figure 15 summarizes the first and second order best fit parameters computed in the figure 14. Therefore we can see that the PSF ellipticity has to be corrected in the lensfit catalogue galaxy shapes before further analysis given its strong leakage.

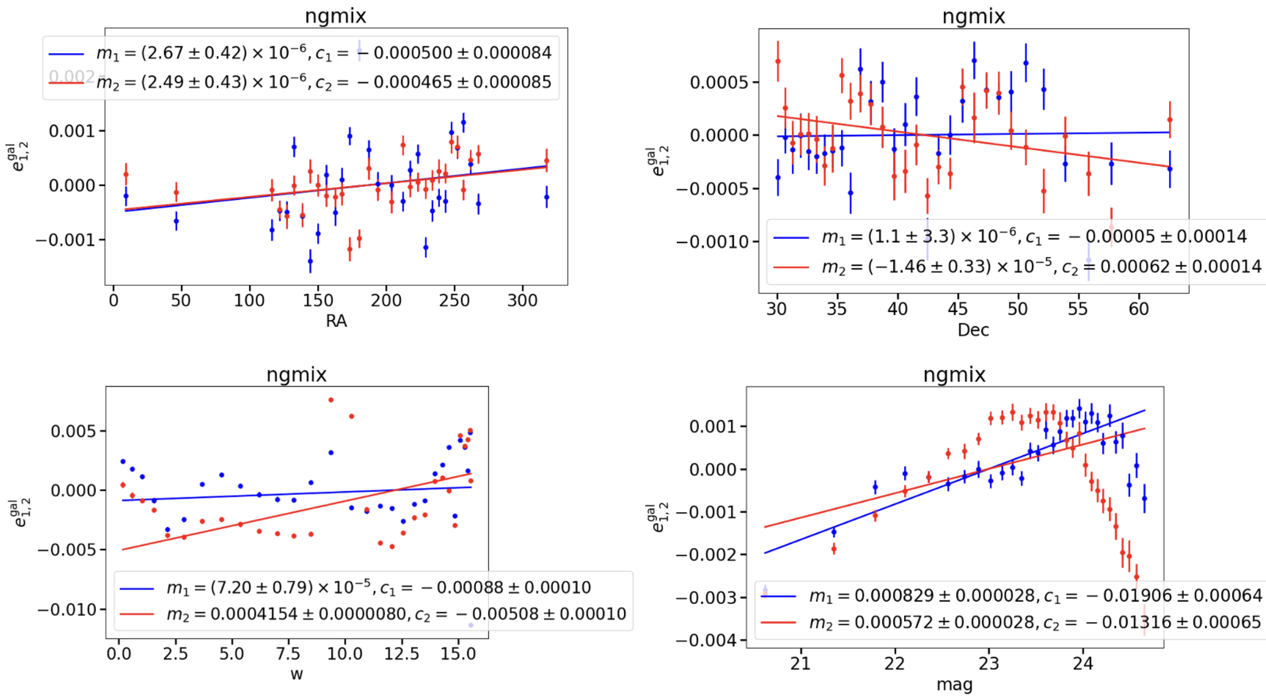

Figure 12: 1D plots of the mean galaxy ellipticity for 2-spin ellipticity components as a function of the RA,Dec,mag,w the solid lines shows the best linear fit without binning with q being the second order parameter, m the first order parameter and c the additive correction

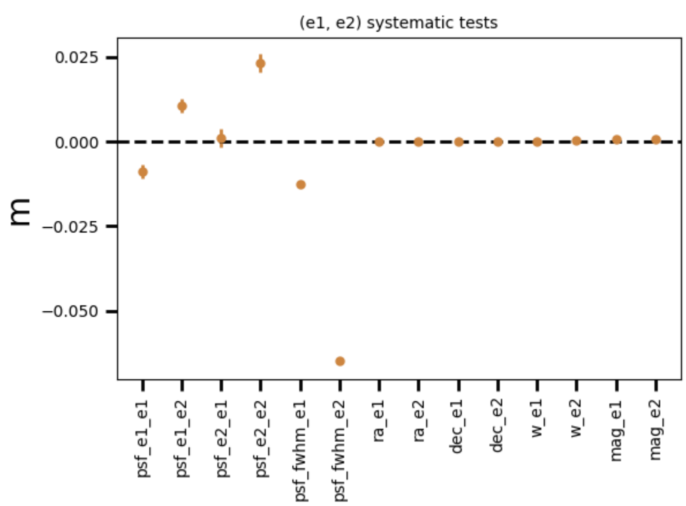

Figure 13: Summary of the slopes extracted m for all the quantities computed in Figure 12

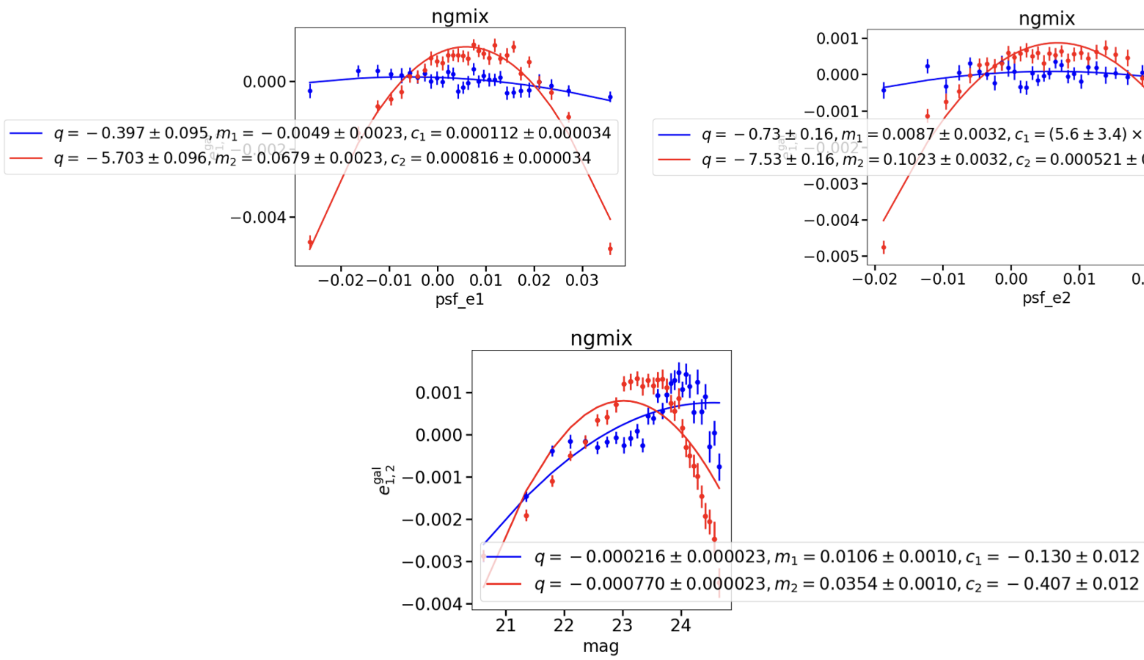

Figure 14: 1D plots of the mean galaxy ellipticity for 2-spin ellipticity components as a function of the PSF ellipticies (e1_PSF and e2_PSF) and the magnitude, the solid lines shows the best quadratic fit without binning with q being the second order parameter, m the first order parameter and c the additive correction

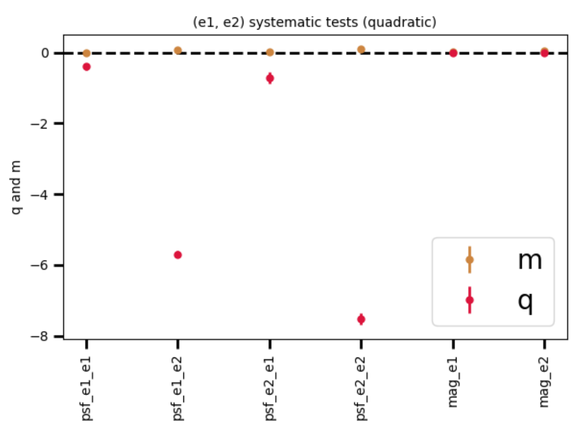

Figure 15: Summary of the slopes extracted m for all the quantities computed in Figure 14

The observational leakage for the two catalogues is in most cases under the threshold of 1% which the correlations can be considered as negligible. But for some quantities as the weight and the magnitude, the fits are not matching with the data and the correlation remains unclear.

3.3.3 Extraction of new data outside the catalogues

For measuring the atmospheric biases which can occur in the data, we want to perform the leakage on two other quantities which are not present in the catalogue. These quantities are the airmass and the seeing. The airmass corresponds to the measure of the amount of air or the atmosphere that the light passes through from a celestial object to an observer. [15] It can be used to measure the effect of the atmosphere on Earth's ground telescope and thus measure the quality of the observations and the limiting magnitude detectable by telescopes. The airmass is in general related to the angle between the zenith and the object being observed, with a larger angle corresponding to a larger airmass. [15]
The seeing refers to the blurring of the images of galaxies due to the Earth's atmosphere and creates biases in the

shape measurements. In general, this effect is correct in the shape measurement pipelines because it can be measured using the FWHM of the PSF and models to simulate PSF and atmosphere effects. [16]

We will extract the seeing and the airmass for a part of the ShapePipe catalogue called P3 which contains more information than the full catalogue. It contains in particular tiles ID of the galaxies which refers to a unique identifier assigned to a specific tile in a survey area divided into a set of tiles which corresponds to regions of the sky. This tiles ID associated with the galaxy in the catalogue allows us to find related information in other files. Two types of files are interesting for finding the seeing and the airmass, a text file associated with each tile image with information as the mean seeing of the tile inside and the name of exposure images files which create the tile image. The exposure image files contain the airmass data we need. As there are several exposure images for one tile image, we will take the average of the airmass in exposure images for each tile image. Therefore, these files allow us to obtain one value of the seeing and one value of the airmass for each galaxy. But as the values are mean values and the tile is a large region in the sky with many galaxies, the extracted values are not precise measures of these quantities for each galaxy but can be used as an order of magnitude.

3.3.3.1 Extraction method For extracting this values, we create a new script which, for an input galaxy catalogue with tiles ID associated for each galaxies, will find the seeing in the tile image file and the airmass in the exposure image file.

- **Seeing** : For finding the seeing in the tile image text files, we use the `re` package in python that provides support for regular expressions which is a language for matching text pattern. The `search()` function of the package, searches the first occurrence of the pattern in the file. We use this method to first finding the tile images files associated to the tile ID of each galaxy, we used the following pattern :

```python
pattern = r'(\d{3})[\.-](\d{3})'.format("quantity_to_extract")
match = re.search(pattern, line)
```

  This pattern will find a number in the following format XXX.YYY, the `\d{3}` means that we are searching a group of 3 digits followed by a point and 3 other digits and separate the two numbers in two separate values. We create a function that returns the tile ID associated with each tile image file for testing all the tile image files if they correspond to the tile ID for each galaxy in the catalogue. If they correspond then we open the file and extract the data to perform another research with the `search()` function of the `re` package. For finding the seeing in the text we used the following pattern :

```python
pattern = r'(\d+\.\d+\D+{}'.format("seeing")
match = re.search(pattern, line)
value
value = match.group(1)
```

  the `\d+` correspond to the one or more digits which associate to `\.`and `\d+` makes a decimal. Following by `\D+` occurrence which means one or more characters between the number created before and the following word (in this case "seeing" but it can be any quantity to extract). After finding the values for each galaxy in the catalogue, the script will write it in an array gradually. When it finishes, a new fits file catalogue is created from the input catalogue and with adding a column with the seeing for each galaxy.

- **Airmass :** For the airmass, we have to find in the tile image files for each galaxy, the exposure image file labels which have been used for creating the tile image. This information is stored after the word `"HISTORY"` in the file. With the same searching method as the seeing, we search the labels of the files in the tile images and store them in an array for each galaxy. Thereafter, for each exposure file which are .FITS file, it is easier to find the value from the header. We first open the file with the same method as the leakage script but this time we want to extract the header not the data. The header provides information about the data in the FITS file such as the size and shape of the image, the observing conditions, and the instrument used to acquire the data. There is an `AIRMASS` header in this file thus for extracting the airmass value, we take the value associated with this header in an array.

```python
hdul = fits.open(exposure_file.fits)
header = hdul[0].header
value  = header["AIRMASS"]
```

After putting the airmass value of all the exposure files associated with one tile image file, we compute the average of this airmass value and store it in an array. We perform the same step for each tile image file associated with each galaxy in the column of the input catalogue. Finally, as for the seeing, we create a new catalogue with the airmass column added in the .fits file.

3.3.3.2 Results Therefore, after extracting the seeing and the airmass, we compute the leakage for these quantities using the new features of the leakage script. We can observe in the output plots that the data points for the seeing are stacked between 0.6 and 0.8 arcseconds and those for the airmass between 1.1 and 1.5. We can conclude in the figure 16 that the absolute leakages are less than 1% at the order of 10^{-4} and show no correlation. Therefore, we can conclude that the seeing and the airmass have been corrected during the shape measurement pipeline of ShapePipe and don't influence the image quality and then the galaxy shape data. Therefore, the hypothesis that the leakage for the astronomical coordinate can be due to airmass and seeing are unlikely but other atmospheric effects or the telescope orientation can be the reasons for these leakages.

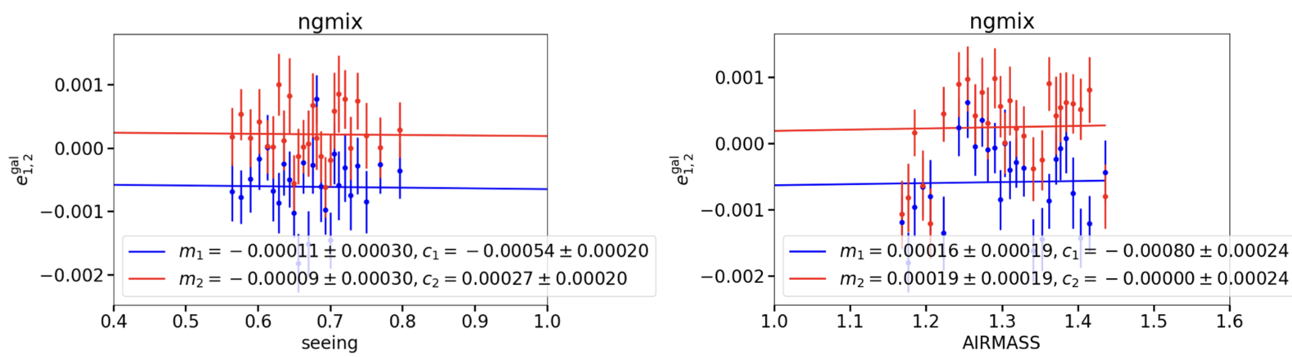

Figure 16: 1D plots of the mean galaxy ellipticity for 2-spin ellipticity components as a function of seeing and airmass , (cf. table 1 of the section 3.1) the solid lines shows the best linear fits without binning with m the slope and c the additive correction

4 Conclusion

Gravitational lensing occurs when a massive object such as a galaxy or a galaxy cluster deflects the light by gravity predicted by Einstein's General Relativity. [1] There are different types of gravitational lensing such as strong lensing, weak lensing and micro lensing which depend on the mass and the geometry of the lens bending the light. Strong lensing produces visible distortions, in opposite to weak lensing which produces more subtle effects which can be studied by averaging a large amount of data. [2] The weak lensing main observable effects are convergence and shear which contain information about the magnification of the lensed image, the enlargement and the stretching. [2] Many methods exist to measure galaxy shapes, for example model fitting methods. Those methods are implemented in different pipelines.
During the Internship, the data analysis was done with data sets called catalogues with galaxy shape data from two different pipelines : ShapePipe and Lensfit. The measurements of weak lensing effects are intrinsically noise-dominated and small residual shape errors can affect the analysis. PSF Leakage is a type of systematic error that can affect weak lensing measurements. It is caused by the Point Spread Function which is the response of the instrument to a point source that blurs the images. The PSF ellipticity can affect the accuracy of the shape measurement and refers to the degree to which the PSF ellipticity deviates from a circular image.

After presenting the PSF Leakage and the script for computing it, the output plots show correlations for both shape measurements pipelines. ShapePipe has a better correction and doesn't need more adjustments for the leakage of the PSF. In comparison, the Lensfit catalogue still has a leakage of 7% for the FWHM of the PSF and has a strong quadratic correlation for the PSF ellipticity. This catalogue has to be corrected by making adjustments in the PSF model or by a direct correction in the galaxy shape measurements.

Thereafter, we presented the new features for the script that generalize the leakage to any quantities for measuring the dependencies between galaxy ellipticity and various quantities. These measures are important for the validation of catalogues and the correction of systematic errors. The leakage test is expected to be a null test and would show a correlation close to 0 for the validation of catalogue.
Then we applied the observational leakage for a number of other quantities on the two catalogues. We observed correlations for certain quantities such as the astronomical coordinates that could be due to atmospheric effects such as the airmass or the seeing. We extracted the data of these quantities from other files for checking this hypothesis but we didn't find any correlations. Another hypothesis could be due to the telescope pointing more toward the horizon which creates biases and affects image quality. We also observed that the leakage increased towards low SNR which is expected since for those galaxies, shape measurement is more difficult because of the noise. However, we consider the validation of the catalogue for a leakage of less than 1% which is the case for the ShapePipe quantities and most of the Lensfit quantities except the PSF and the FWHM.
In addition for some quantities the linear or quadratic model is not a good fit to the data, thus we cannot make clear conclusions about these quantities.

For the leakage of lensfit quantities, we found dependencies for the weight and the magnitude, but we can observe that the linear fit doesn't fit well with the data. We tried to apply a quadratic fit to the magnitude leakage of the lensfit catalogue, it also shows a correlation but the points also deviate from the fit. A dependency with the magnitude could be due to the large size of bright galaxies which could introduce a bias since there are some edge effects from the finite size of the postage stamps. Nevertheless, the order of magnitude of the correlation could make them negligible for the validation of the catalogue.

Limitations of leakage could appear if the relation between the ellipticity and the quantities is not linear or quadratic. In this case, the fit could not match well with the data and creates an unclear correlation which biases the analysis. Moreover, the assumptions we made for all the correlations have to be confirmed by future surveys.
Nevertheless, PSF leakage is a strong method for checking systematic errors related to the PSF model and galaxy shape correction. The leakage for any quantity is important to check for atmospheric effects which can create biases for ground-based telescopes and to search for undesired dependencies between quantities to correct, for a deeper validation of the catalogues. These systematic tests method could be an asset for the analysis of future weak lensing surveys such as Euclid, launched at the beginning of July.

References

[1] M. Bartelmann M. Maturi. Weak gravitational lensing. *Scholarpedia*, 2017.

[2] M. Kilbinger. Weak gravitational lensing course, euclid summer school, banyuls, 2019.

[3] Dragan Huterer. Weak lensing, dark matter and dark energy. *General Relativity and Gravitation*, 42(9):2177–2195, jul 2010.

[4] P. Simon, H. Saghiha, S. Hilbert, P. Schneider, C. Boever, and A. H. Wright. Comparison of the excess mass around CFHTLenS galaxy-pairs to predictions from a semi-analytic model using galaxy-galaxy-galaxy lensing. *Astronomy & Astrophysics*, 622:A104, feb 2019.

[5] C. Bonini. Weak gravitational lensing. *M1 Research Project report*, 2023.

[6] CosmoStat. Weak lensing page of the website, https://www.cosmostat.org/research-topics/weak-lensing.

[7] F William High, Jason Rhodes, Richard Massey, and Richard Ellis. Pixelation effects in weak lensing. *Publications of the Astronomical Society of the Pacific*, 119(861):1295, 2007.

[8] B. Casaponsa A. F. Heavens T. D. Kitching L. Miller R. B. Barreiro E. Martinèz-Gonzalez. Size magnification as a complement to cosmic shear, 2013.

[9] S. Farrens, A. Guinot, M. Kilbinger, T. Liaudat, L. Baumont, X. Jimenez, A. Peel, A. Pujol, M. Schmitz, J.-L. Starck, and A. Z. Vitorelli. ShapePipe: A modular weak-lensing processing and analysis pipeline. *Astronomy & Astrophysics*, 664:A141, aug 2022.

[10] Axel Guinot, Martin Kilbinger, Samuel Farrens, Austin Peel, Arnau Pujol, Morgan Schmitz, Jean-Luc Starck, Thomas Erben, Raphael Gavazzi, Stephen Gwyn, Michael J. Hudson, Hendrik Hildebrandt, Liaudat Tobias, Lance Miller, Isaac Spitzer, Ludovic Van Waerbeke, Jean-Charles Cuillandre, Sébastien Fabbro, Alan McConnachie, and Yannick Mellier. ShapePipe: A new shape measurement pipeline and weak-lensing application to UNIONS/CFIS data. *Astronomy& Astrophysics*, 666:A162, oct 2022.

[11] Erin S. Sheldon and Eric M. Huff. Practical weak-lensing shear measurement with metacalibration. *The Astrophysical Journal*, 841(1):24, may 2017.

[12] CosmoStat. Unions page of the website, https://www.cosmostat.org/projects/unions-cfis.

[13] Alexandre Refregier, Tomasz Kacprzak, Adam Amara, Sarah Bridle, and Barnaby Rowe. Noise bias in weak lensing shape measurements. *Monthly Notices of the Royal Astronomical Society*, 425(3):1951–1957, aug 2012.

[14] M Gatti, E Sheldon, A Amon, M Becker, M Troxel, A Choi, C Doux, N MacCrann, A Navarro-Alsina, I Harrison, D Gruen, G Bernstein, M Jarvis, L F Secco, A Ferté, T Shin, J McCullough, R P Rollins, R Chen, C Chang, S Pandey, I Tutusaus, J Prat, J Elvin-Poole, C Sanchez, A A Plazas, A Roodman, J Zuntz, T M C Abbott, M Aguena, S Allam, J Annis, S Avila, D Bacon, E Bertin, S Bhargava, D Brooks, D L Burke, A Carnero Rosell, M Carrasco Kind, J Carretero, F J Castander, C Conselice, M Costanzi, M Crocce, L N da Costa, T M Davis, J De Vicente, S Desai, and H T Diehl et al. Dark energy survey year 3 results: weak lensing shape catalogue. *Monthly Notices of the Royal Astronomical Society*, 504(3):4312–4336, apr 2021.

[15] Lei Lei, Qing-Feng Zhu, Xu Kong, Ting-Gui Wang, Xian-Zhong Zheng, Dong-Dong Shi, Lu-Lu Fan, and Wei Liu. Limiting magnitudes of the wide field survey telescope (WFST). *Research in Astronomy and Astrophysics*, 23(3):035013, mar 2023.

[16] Saavidra Perera, Richard W Wilson, Tim Butterley, James Osborn, Ollie J D Farley, and Douglas J Laidlaw. SHIMM: a versatile seeing monitor for astronomy. *Monthly Notices of the Royal Astronomical Society*, 520(4):5475–5486, feb 2023.