Cosmic Shear Variance from 50 Uncorrelated VLT FORS1 Fields Using Shapelets method*

*Article in preparation

ABSTRACT — A new pipeline based on the shapelets (Refregier, 2003) technique has been developed for the reduction of a huge amount of data which will be produced in the future by the VLT Survey Telescope - OmegaCAM system and for the study of galaxy-galaxy lensing and cosmic shear.

The pipeline extracts the objects present in the images, determines parameters to characterize the objects, and compensates the point spread function introduced by atmospheric disturbances and optical defects from the instrument that affect the image of the galaxies (Kuijken, 2006).

The shapelets pipeline has proved reliable with simulated images. It is able to extract the ellipticity of simulated galaxies and the artificial shear introduced in the simulated images of the galaxies. To validate definitely the shapelets pipeline, the latter has to prove its efficiency on real images. In this chapter we present a study performed on the data used by Maoli et al. (2001). Using the shapelet pipeline and a second pipeline based on the Kaiser, Squires and Broadhurst (KSB) method, we reproduce the calculations of Maoli et al. (2001) which were performed to extract a significant cosmic shear signal from their images. The comparison of the data from the two pipelines, the computation of the cosmic shear signal as well as the analyses of possible systematic errors have shown that the pipeline based on the shapelets algorithm fulfills all requirements for performing a cosmic shear analysis and providing more accurate results.
6.1 Introduction

Gravitational lensing provides powerful tools to quantify and locate the dark matter in galaxies, cluster of galaxies and in large scale structures. Theoretical analyses have been performed these past decades concerning the effects of massive astronomical objects on the path of light coming from sources located in their background. They have shown that mass can act as a lens and that finding has been confirmed recently by several observations. Strong lensing (Huchra et al., 1985), galaxy-galaxy lensing (Brainerd et al., 1996a), weak lensing by a cluster of galaxies (Tyson et al., 1990) and cosmic shear analyses (Van Waerbeke et al., 2000; Bacon et al., 2000; Wittman et al., 2000; Kaiser et al., 2000) have been carried out over the past decade demonstrating the veracity of the predictions. Gravitational lensing has become an important tool for determining the presence of the dark matter as well as its mass contribution on all scales throughout the universe.

The anisotropy in the mass of large scale structures causes a weak distortion of the images of lensed objects. This effect is called cosmic shear and the detection of such distortions remains a major goal (Mellier, 1999; Bartelmann and Schneider, 2001). To study the distribution of matter on such a large scale, the gravitational lensing theory has to be applied to new more accurate observations. It is now possible to study the variation of the cosmic shear signal, even though its amplitude is very small. We have cited above the first four groups that have managed to measure a cosmic shear signal. Others did it also latter with either small field instruments like the VLT/FORS1 and HST/STIS (Maoli et al., 2001; Haemmerle et al., 2001) or using data from large surveys, e.g. VIMOS-Descart, Red-sequence Cluster Survey, CTIO Lensing Survey (Van Waerbeke et al., 2005; Hoekstra et al., 2002; Jarvis et al., 2005). The reasons for these successes are the high-quality of the data and the improved estimates of the gravitational shear from the galaxies shapes.

In general, the measurements are performed on galaxy images that are faint and poorly sampled (only few pixels per galaxy). They are also strongly affected by pixel noise and, optical and atmospheric disturbances (point spread function, PSF) which can mimic a cosmic shear signal (Kaiser et al., 1995). The distortions induced by gravitational lensing have a magnitude of a percent. The magnitude of these fake signals has the same amplitude as a real cosmic shear signal, and sometimes even larger. To remove the parasite signals several techniques have been developed (Kaiser et al., 1995; Bonnet and Mellier, 1995; Luppino and Kaiser, 1997; van Waerbeke et al., 1997; Fischer and Tyson, 1997; Hoekstra et al., 1998; Kuijken, 1999; Kaiser, 2000; Bernstein and Jarvis, 2002; Hirata and Seljak, 2003). A technique which is now widely used by the scientific community is the Kaiser, Squires & Broadhurst method (KSB henceforth). This method is rather simple and permits to correct the vast majority of the defects that affects the ellipticity of the objects (isotropic and anisotropic perturbation, see Chapter 5 for more detail). This technique works well if a calibrated shear coefficient, \( \sim 0.85 \), is applied to compensate biases (Bacon et al., 2001) that are induced by the problem of "Kaiser flow" (Kaiser, 2000). Until now, the KSB method and its precision were acceptable. The technology which was at the disposal of the scientific community could not offer measurements with higher accuracies (Van Waerbeke et al., 2000). This statement is no longer true with new optical wide field cameras that have become available to astronomers during this decade. Because of positive results obtained this last 10 years from gravitational
6.2: Description of the Data

We have chosen to work on 50 images used by Maoli et al. (2001) from which the authors have extracted a significant cosmic shear signal. The images have been kindly provided by Pr. Dr. Y. Mellier (IAP, France). We briefly summarize here the main characteristics of the data that are described in detail in the Maoli et al. (2001) article.

The data set consists of 50 FORS1 fields (Appenzeller et al., 1998) which have been selected randomly in a sky area of about 1000 square degrees. This sample of uncorrelated images cover 0.64 square degrees. Each image has been taken with a TK2048 EB4-1 backside thinned CCD\(^1\) which provides an image of 2048 \(\times\) 2048 pixels. Each pixel has a sensitive area of 24 \(\times\) 24 microns which samples the images in the focal plane of the VLT every 0.2 arc seconds. The images have been taken with the I-band filter where no fringe patterns are present and have an average seeing of 0.63 \(\pm\) 0.13" (1\(\sigma\)). Each image is a composition of 6 scientific acquisitions of the same sky field with an exposure time of 6 minutes per image. This procedure yields an image which has an equivalent of 36 minutes exposure time per set of sky fields, and provides a density of 30 galaxies per square arc minute and an average red-shift of one for the source galaxies. The data have been reduced using the TERAPIX pipeline\(^2\) to obtain calibrated images.

\(^1\)http://www.eso.org/projects/odt/Fors1/Fors1.html
\(^2\)http://terapix.iap.fr
6.3 Data Reduction

The built pipelines, based either on the KSB method or the shapelets, follow the same structure. First we extract from each image a catalog of objects. Next, from each catalog we extract the stars and the galaxies separately and form a stars catalog and a galaxies catalog, respectively. Using the stars information, first we compute the parameters which are necessary to estimate the PSF at the galaxy positions and then we correct the ellipticity of each galaxy that is disturbed by the PSF. This procedure yields a catalog of galaxies per each image with their corrected ellipticity and location. These catalogs are then used to compute the cosmic shear signal as a function of the angular scale of the top hat filter in which the cosmic shear signal is computed.

6.3.1 Data Reduction Based on the IMCAT Software and the KSB Method

In chapter 5, we described in detail the object shape measurement and the correction of the point spread function (PSF). The KSB method is used here again with some modifications in the evaluation of the star parameters to estimate the PSF at galaxy positions. We also include a correction factor introduced by Bacon et al. (2001) to compensate for the systematic underestimation of the object ellipticities. The procedure based on the KSB method is used to correct the defects in the images introduced by the electronic and photon noise, atmospheric turbulence, imperfect optics and oscillation of the telescope due to the wind. We briefly summarize the different steps of this procedure.

6.3.1.i Principle

After the extraction of the objects we estimate their parameters, position, object magnitude, ellipticity \( e \), the smear and shear polarisability tensors \( P_{sm} \) and \( P_{sh} \), respectively. The ellipticity is given by,

\[
e = \frac{Q_{11} - Q_{22} + 2iQ_{12}}{Q_{11} + Q_{22}}
\]

where \( Q_{ij} \) are the components of the second brightness moment tensor (see equation 5.7).

In order to correct the PSF, higher order moments are used to estimate the \( 2 \times 2 \) tensors, \( P_{sm} \) and \( P_{sh} \), of each object. These two tensors describe a smearing and a shearing of the image of an object. The complete information about the PSF of the CCD images is contained in the images of the stars. It is crucial to identify the stars in the CCD images and to determine \( e_{obs,i}^* \), \( P_{sm,i}^* \) and \( P_{sh,i}^* \) for each of them\(^3\). Employing \( e_{obs,i}^*, P_{sm,i}^* \) and \( P_{sh,i}^* \), first we extrapolate values for the star tensors \( \tilde{e}_{obs,i}, \tilde{P}_{sm,i}^*, \tilde{P}_{sh,i}^* \) which they would have at the galaxy position and next we compensate the PSF which affects each galaxy. We have:

\[
e_{int,i} = e_{obs,i} - P_{sm,i} \left( \tilde{P}_{sm,i}^* \right)^{-1} \tilde{e}_{obs,i} - P_{sh,i} g_i + \tilde{P}_{sh,i} P_{sm,i} \left( \tilde{P}_{sm,i}^* \right)^{-1} g_i
\]

\(^3\)The symbol \(*\) in the superscript refers to the star
where $e_{\text{int},i}$ is the intrinsic ellipticity tensor of the galaxy $i$, $e_{\text{obs},i}$ is the ellipticity tensor of the observed galaxy $i$, $P_{\text{sm},i}$ and $P_{\text{sh},i}$ are the smear and shear tensors of the observed galaxy $i$, $\hat{e}_{\text{obs},i}^*$ is the estimate of the star ellipticity tensor at the galaxy position $i$, and $\hat{P}_{\text{sm},i}^*$ and $\hat{P}_{\text{sh},i}^*$ are the estimated smear and shear tensor of the star at the galaxy position $i$, $g_i$ is the reduced shear (see Chapter 5).

For convenience the isotropic ellipticity, $e_{\text{iso}}$, and the pre-seeing shear polarizability tensor, $P_g$, are introduced:

\begin{align}
    e_{\text{iso},i} &= e_{\text{obs},i} - P_{\text{sm},i} \left( \hat{P}_{\text{sm},i}^* \right)^{-1} \hat{e}_{\text{obs},i}^* \\
    P_{g,i} &= P_{\text{sh},i} - \hat{P}_{\text{sh},i}^* P_{\text{sm},i} \left( \hat{P}_{\text{sm},i}^* \right)^{-1}
\end{align}

$P_g$ appears to be inaccurate due to noise when the galaxy image has a low resolution. To improve the accuracy of this parameter, it has to be smoothed to obtain $P_{g,\text{smo}}$. This smoothing is performed in the (size, magnitude) basis (Kaiser et al., 1998; Van Waerbeke et al., 2000). Knowing also that in the weak lensing regime, the reduced shear, $g_i$, can be approximated by the shear, $\gamma$ (Bartelmann and Schneider, 2001), we can obtain an estimate of the shear, $\tilde{\gamma}$ as:

\begin{equation}
    \tilde{\gamma}_i = \left( P_{g,\text{smo}} \right)^{-1} e_{\text{iso},i}
\end{equation}

We have to keep in mind that $\tilde{\gamma}_i$ is strongly affected by the intrinsic ellipticity, $e_{\text{int},i}$, of the galaxy. To determine the cosmic shear signal $\langle \gamma^2 \rangle$ which is the variance of $\gamma$, $e_{\text{int},i}$ will interfere if a specific care has not been taken. To obtain an unbiased estimate of the cosmic shear signal, an unbiased estimator will be used (Van Waerbeke et al., 2001; Maoli et al., 2001). Before determining the variance of $\gamma$, we have to take into consideration the results of Bacon et al. (2001) which show that the KSB method underestimates systematically the results for $\gamma_i$. A compensation factor of 0.85 has to be applied to $\gamma_i$ to obtain the corrected value of $\tilde{\gamma}_i$, $\tilde{\gamma}_{i,\text{corr}} = \tilde{\gamma}_i/0.85$.

### 6.3.1.ii The Measurements

A modified version of the IMCAT software\(^4\) (Kaiser et al., 1995) has been used to extract the objects and to determine the necessary parameters which will be used in the method discussed above (see Chapter 5 for more information). The procedure yields one catalog per image. Each catalog is sorted by forming two sub-catalogs, one containing the stars, and the other the galaxies. This is done by using the size-magnitude diagram (see Figure 5.7). The peak in this plot corresponds to the stellar branch and the objects with a radius larger than the maximum radius of the stars are galaxies. The objects with the radius smaller than this limit are discarded because they are smaller than the seeing disc diameter. Before using equation (6.5) to obtain an estimate for the shear $\gamma$ we need to evaluate the star parameters $\hat{e}_{\text{st}}^*$, $\hat{P}_{\text{sm},i}^*$ and $\hat{P}_{\text{sh},i}^*$ at the position of the galaxy $i$. Two methods have been used to approximate these parameters. The first method is commonly used and it is based on a second order polynomial fit. The second method is based on

\(^4\)http://www.ifa.hawaii.edu/~kaiser
a weighted nearest neighbor average approach. The tensors, \( \tilde{\Gamma} = \tilde{e}_i \), \( \tilde{P}_{sm,i} \), \( \tilde{P}_{sh,i} \), at the corresponding grid position \( I(x_i, y_i) \) are calculated using the formula,

\[
\tilde{\Gamma}_i(x_i, y_i) = \frac{\sum_{j=1}^{n} \frac{\Gamma_j(x_j, y_j)}{\sqrt{d_j^2 + D^2}}}{\sum_{j=1}^{n} \frac{1}{\sqrt{d_j^2 + D^2}}}
\]  

(6.6)

where \( d_j \) is the distance between the point \( I(x_i, y_i) \) and \( J(x_j, y_j) \), \( D = 100 \) pixels is a smoothing parameter which corresponds to a half the distance between two consecutive stars ideally spread on the grid (we suppose here that we have 100 stars per image), and \( n \) is the number of closest data points lying around the point \( I(x_i, y_i) \).

Figure 6.1: Example of a star ellipticity reconstruction, \( e_1 \) and \( e_2 \), over one of the 50 images using a two degree polynomial fit (first column) and the weighted nearest neighbor average method (second column).

An example of both estimations can be found in Figure 6.1. The ellipticity reconstruction based on the weighted nearest neighbor average method may need more smoothing but the actual results are not sensitive to this. These methods have been employed for the star catalogs to check whether the PSF correction is rigorously applied (see Figure 6.2). Before any correction we observe an inhomogeneous cloud of points with a barycenter localized away from \((0,0)\). After the anisotropic correction we observe a homogenous circular cloud centered at \((0,0)\) for both the 2d polynomial fit and weighted nearest neighbor average method. Figure 6.2 shows that the weighted nearest neighbor average method realizes a better correction than the 2d polynomial fit. The inner circles in the diagrams A, B, and C correspond to the average norm of \( \tilde{e}^{iso} \). Before correction we have \( \tilde{e}^{iso}_b = 0.035 \).
After correction with the 2d polynomial fit technique we obtain $e_{2d} = 0.015$, which corresponds to an improvement of about 56% and with the weighted nearest neighbor average method $e_w = 0.01$, and the improvement is about 71%. The improvement achieved by the weighted nearest neighbor average method compared to the 2d polynomial fit technique is of the order of 34%. The outer circles in the diagrams A, B, and C represent the standard deviation which is defined as $\sigma_{e_{iso}} = \sqrt{\sigma_{e_{iso,x}}^2 + \sigma_{e_{iso,y}}^2}$. Before correction we determine $\sigma_{e_{iso}} = 0.015$, and after the correction with the 2d polynomial fit technique we have $\sigma_{e_{iso}} = 0.014$. Employing the weighted nearest neighbor average method we obtain $\sigma_{e_{iso}} = 0.008$. Clearly, using the 2d polynomial fit, we observe almost no change in the standard deviation before and after the correction ($\sim 7\%$), whereas using the weighted nearest neighbor average method, we see a significant improvement ($\sim 48\%$).

This information has to be considered with caution. To correct the stellar ellipticity $e_i$ we have used the parameters $\tilde{e}_i^*, \tilde{P}_{sm,i}^*$ and $\tilde{P}_{sh,i}^*$ estimated at the position of star $i$. These parameters are deduced from the stellar parameters $e_i^*, P_{sm,i}^*$ and $P_{sh,i}^*$. It can be easily seen from equation (6.6) that the estimated parameters obtained within the weighted nearest neighbor average method are better approximations to $e_i^*, P_{sm,i}^*$ and $P_{sh,i}^*$ than those derived from the 2d polynomial fit method. To prove that the weighted nearest neighbor average method is really more efficient than the 2d polynomial fit method, the star catalogs have been split in two sub-catalogs arbitrarily. $4/5$ of the stars have been used to estimate the parameters which are necessary to correct the PSF using both the 2d polynomial fit and the weighted nearest neighbor average method. The PSF correction methods were then applied to the rest of the stars. For each object we computed the norm, $e^{iso}$, the average of the norm, $\bar{e}^{iso}$, and the standard deviation of the norm. The results are plotted in Figure 6.3. This manipulation has been repeated 20 times. On average, the norm is about $0.0248 \pm 0.0015(1\sigma)$ after the correction based on the 2d polynomial fit and about $0.0214 \pm 0.0005(1\sigma)$ after using the weighted nearest neighbor average method. The scatter of the points is found to be smaller within the weighted nearest neighbor average method, $0.0164 \pm 0.020(1\sigma)$, than within the 2d polynomial fit method, $0.0199 \pm 0.056(1\sigma)$. The PSF correction of the stars obtained in this study using the weighted nearest neighbor average method is improved by $\sim 17\%$ and it is much more stable with a noise reduced by $\sim 65\%$ compared to the 2d polynomial fit.

To correct the ellipticity of the galaxies affected by the PSF, the weighted nearest neighbor average method has been used. First we compute and smooth $P_g$, and then we calculate the shear estimate of the galaxy, using equation (6.5). This complete procedure yields a catalog per image which contains the position, the magnitude, the radius, the observed ellipticity and the estimated shear.

### 6.3.2 Data Reduction Based on Source Extractor and the Shapelets

The procedure based on the shapelets algorithm decomposes the image of each optical object into shapelets (Refregier, 2003). The technique developed by Kuijken (1999) to correct the PSF has been adapted to the shapelets method in order to determine the ellipticity of the galaxies before any atmospheric and instrumental disturbances.

The first step consists of creating a catalog of objects for each image. This task is performed by the SExtractor software (Bertin and Arnouts, 1996) and provides the necessary information to localize all objects, to determine their ellipticity, their size and
Figure 6.2: Ellipticity of the stars before (A) and after anisotropic PSF correction using the KSB method (B) or the shapelets method (C). Image (B) shows the ellipticity corrected after estimating the parameters by using a 2d polynomial fit. Image (C) shows the correction after using the weighted nearest neighbor average method. The inner circle corresponds to the norm, $e^{iso}$, and the outer circle expresses the standard deviation multiplied by a factor of three.

magnitude as well as a flag parameter. The latter is used to quantify the quality of the measurements. Analogously to the procedure based on the KSB method, we have to locate the stars in each image before any further considerations. We will use the stars once again in order to estimate the PSF in the images. To extract the stars, the catalogs created by SExtractor are used. Next the plot magnitude versus full-width at half-maximum, FWHM, is realized for each catalog. The points which compose the characteristic peak in the plot correspond to the stars in the images we are interested in. Once the stars are located in the complete image, their images are decomposed into shapelets and the obtained coefficients are used to extrapolate, employing a polynomial fit, the PSF at the position of the galaxies (Kuijken, 2006). To identify the galaxies, the catalogs produced by SExtractor are used again. All objects with a FWHM larger than the largest star FWHM, which correspond to objects larger than the seeing, are considered to be galaxies. Each galaxy image is then decomposed also in the shapelets basis set. We have now the necessary information to follow the approach described by Kuijken (1999) and adapted to the shapelets formalism. This technique consists of shearing and smearing a circular source in order to obtain the ellipticity of the PSF-convolved galaxy. Knowing the PSF we can deduce the ellipticity before being affected by the PSF. The circular model of a galaxy is given by:
Figure 6.3: Ellipticity of the stars before (A) and after anisotropic PSF correction using the KSB method (B) and the shapelets method (C) for 1/5 of the stars. The other 4/5 have been used to estimate the star parameters. The legend of these plots is the same as that in Figure 6.2.

\[ G_0 = c_0 C_0 + c_2 C_2 + c_4 C_4 + \ldots \]  
\[ \text{where } G_0 \text{ is the circular source, } C_0, C_2, C_4, \ldots \text{ are circular shapelets of even order (see Appendix A of Kuijken, 2006) and } c_0, c_2, c_4, \ldots \text{ are coefficients.} \]

The observed galaxy is then expressed as:

\[ G_{\text{obs}} = P(1 + e_1 S_1 + e_2 S_2 + d_1 T_1 + d_2 T_2)(c_0 C_0 + c_2 C_2 + c_4 C_4 + \ldots) \]  
\[ \text{where } G_{\text{obs}} \text{ is the observed galaxy, } P \text{ is the PSF matrix, } S_{1,2} \text{ are the first-order shear operators (Refregier, 2003), } e_{1,2} \text{ corresponds to the ellipticity elements, } T_{1,2} \text{ are translation operators, } d_{1,2} \text{ are coefficients. Furthermore the terms } d_1 T_1 \text{ and } d_2 T_2 \text{ are used here to ensure that the center of the object coincides with the center of the shapelets expansion for a better decomposition (see Kuijken, 2006). The objective of this step consists of fitting a circular object, } G_0, \text{ sheared by an ellipticity } (e_1, e_2) \text{ and convolved with the PSF, } P, \text{ to the observed galaxy, } G_{\text{obs}}. e_{1,2}, d_{1,2} \text{ and } c_{0,2,4,\ldots} \text{ are taken to be free parameters which are obtained numerically after successive iterations. The complete procedure yields a catalog per each image in which for each galaxy the position, flux, FWHM, the first and second axis of the galaxies, the position angle, the ellipticity elements, } e_{1,2}, \text{ and their errors are present. The ellipticity elements } e_{1,2} \text{ correspond to a very inaccurate estimate} \]
of the shear, $\gamma_{1,2} = c_{1,2}$, that is strongly affected by the intrinsic ellipticity of the galaxy. To determine the unbiased cosmic shear signal, we use the technique employed by Van Waerbeke et al. (2001) and Maoli et al. (2001) (see Section 6.5).

6.4 Shapelets and KSB Catalogs Comparison

The shear estimates of the galaxies obtained by the two pipelines are compared. First the shear parameters in the KSB catalogs, which are not corrected using the compensation factor introduced by Bacon, are compared to the shear parameters present in the shapelets catalogs. For all objects present in all catalogs we plot $\tilde{\gamma}_{1,2}^{ksb}$ versus $\tilde{\gamma}_{1,2}^{shapelets}$. To avoid a multiple indexing $\tilde{\gamma}_{1,2}^{ksb}$ or $\tilde{\gamma}_{1,2}^{shapelets}$ denotes the two components of a particular $\tilde{\gamma}_i$ where $i$ stands for the galaxy $i$. When the KSB and shapelets pipelines provide the same result for $\tilde{\gamma}_{1,2}$, the slope $\alpha_{1,2}$ of the plot $\tilde{\gamma}_{1,2}^{ksb} = \alpha_{1,2} \tilde{\gamma}_{1,2}^{shapelets}$ is 1.0. Our results for $\tilde{\gamma}_{1,2}^{ksb}$ and $\tilde{\gamma}_{1,2}^{shapelets}$ pointed out $\alpha_{1,2} = 0.84 \pm 0.01(1\sigma)$. The KSB algorithm is known to underestimate the shear $\gamma_i$ (Bacon et al., 2001). We have also evaluated the results from the pipeline based on the shapelets algorithm during the 'Shear Testing Program', STEP (Heymans et al., 2006). STEP provides a set of realistic simulated images of sheared patches of the sky. The shapelets algorithm recovers the correct value of the shear, $(\gamma_{1,2}^{out}/\gamma_{1,2}^{in} = 1.0 \pm 0.02)$. In addition the pipeline recovers the ellipticity of the objects with an accuracy better than 1%. Therefore we confirm the results obtained by Bacon et al. (2001): the shear yielded by the pipeline based on the KSB algorithm has to be corrected by a factor of 0.85.

Furthermore a detailed statistical analysis of the slope, $\alpha_{1,2}$, has been performed. We have calculated the slopes $\alpha_{1,2}^k$ from $\tilde{\gamma}_{1,2}^{k,ksb}$ versus $\tilde{\gamma}_{1,2}^{k,shapelets}$ for each image $k$, $k = 1, ..., 50$. We obtained $\overline{\alpha}_{1,2} = 0.83$, as the average of all $\alpha_{1,2}^k$ and a standard deviation of $\sigma_{1,2} = 0.06$. After using the compensation factor introduced by Bacon we have $\overline{\alpha}_{1,2} = 0.98$ and a standard deviation of $\sigma_{1,2} = 0.07$. In both cases the dispersion of the data was surprisingly high. It turns out that the plots $\tilde{\gamma}_{1,2}^{ksb} = \alpha_{1,2}^k \tilde{\gamma}_{1,2}^{shapelets}$ for ten images have slopes which differ from the average slope, $\overline{\alpha}_{1,2}$, by more than 10%. Further analyses have shown that the slopes, $\alpha_{1,2}^k$, are correlated to the seeing of the images $k$ (see Figure 6.4.A and 6.4.B). In Figure 6.4.A, we observe that the KSB algorithm underestimates the shear of the objects more as the seeing gets worse. Assuming the shapelets algorithm performs correctly as suggested by the STEP simulations (Kuijken, 2006), we conclude that the compensation factor introduced by Bacon et al. (2001) has to be replaced by a function of the seeing,

$$c(s) = 1 - 0.28s$$

$$\tilde{\gamma}_{1,2}^{ksb,corr} = \frac{\tilde{\gamma}_{1,2}^{ksb}}{c(s)}$$

(6.9)

(6.10)

where $c(s)$ is the new compensation factor which we use as a function of the seeing s. This empirical equation has been deduced from the plot A in Figure 6.4 and it is valid for the data in the seeing range between 0.4 and 0.95 arcsec. $\tilde{\gamma}_{1,2}^{ksb}$ and $\tilde{\gamma}_{1,2}^{ksb,corr}$ are the shear estimate and the corrected shear estimate, respectively. To avoid multiple indexing, the superscript "corr" in $\tilde{\gamma}_{1,2}^{ksb,corr}$ is omitted latter in this chapter. $\tilde{\gamma}_{1,2}^{ksb}$ have been corrected by the compensation factor $c(s)$. 
After modifying the pipeline, based on the KSB algorithm, to include the new compensation factor, we compared once again the catalogs provided by the two pipelines (see Figure 6.5.A and 6.5.B). We obtained for the slopes $\alpha_1 = 1.01 \pm 0.01$ and $\alpha_2 = 1.01 \pm 0.01$ using the values of $\tilde{\gamma}_{1,2}^{k,\text{ksb}}$ and $\tilde{\gamma}_{1,2}^{\text{shapelets}}$ of all catalogs. The slopes $\alpha_{1,2}^k$ are plotted for each image, $k$, in Figure 6.4.C. The average of the slopes gives $\bar{\alpha}_1 = 1.01 \pm 0.05$ and $\bar{\alpha}_2 = 1.01 \pm 0.06$. After introducing the compensation factor, $c(s)$, $\tilde{\gamma}_{1,2}^{k,\text{ksb}}$ and $\tilde{\gamma}_{1,2}^{\text{shapelets}}$ have similar values, because $\tilde{\gamma}_{1,2}^{k,\text{ksb}}$ are no longer underestimated and correlated to the seeing of the images. The dispersions of the slopes have also been reduced (see Figure 6.5.C and 6.5.D). At this step of our control sequences, we conclude that both pipelines provide the correct estimates of the shear $\tilde{\gamma}_{1,2}$.
6.5 Cosmic Shear Signal Calculation

After demonstrating that there are no significant differences between the catalogs provided by the pipelines, the computation of the variance of the shear has been carried out following the work of Maoli et al. (2001). As mentioned in the sections above, we need to compute an unbiased estimate of the variance of the shear. It has been shown that the following equation fulfills this requirement (Van Waerbeke et al., 2001):

\[
E \left[ \tilde{\gamma}^2 (\theta) \right] = \sum_{\beta=1}^{2} \sum_{m \neq l}^{N} w_m w_l \tilde{\gamma}_\beta (\theta_m) \tilde{\gamma}_\beta (\theta_l) / \sum_{m \neq l}^{N} w_m w_l \tag{6.11}
\]

where \( E \left[ \tilde{\gamma}^2 (\theta) \right] \) is an estimate of the shear variance inside a window defined by a top hat filter of a size \( \theta \), \( \tilde{\gamma}_\beta (\theta_m) \) (for both components \( \beta = 1, 2 \)) are the shear estimates of our objects at the position \( \theta_m \) of the galaxy \( m \), and \( w_m \) is the weight of the galaxy \( m \). To calculate the weight we use the intrinsic ellipticity variance \( \sigma_{e_{\text{int}}}^2 \) and the ellipticity
6.5: Cosmic Shear Signal Calculation

The variance of the nearest 20 galaxies of the galaxy \( m \), \( \sigma_{e,obs}^2 \), \( w_k = (\sigma_{e,obs}^2 + \sigma_{e,int}^2)^{-1} \) (Van Waerbeke et al., 2005).

We can deduce the cosmic shear signal:

\[
\langle \gamma^2 \rangle = \langle E[\tilde{\gamma}^2(\theta)] \rangle
\]

where \( \langle \gamma^2 \rangle \) is the estimate of the cosmic shear signal within a window defined by a top hat filter of size \( \theta \), and \( \langle E[\tilde{\gamma}^2(\theta)] \rangle \) is the ensemble average of the estimate of the shear variance \( E[\tilde{\gamma}^2(\theta)] \).

To determine the cosmic shear signal from the 50 VLT for1 images we proceed as follows. Each image is sub-divided into equal squares with a size of \( \theta \). In each sub-window the estimate of the shear variance is calculated using equation (6.11). Next the ensemble average is computed in order to obtain the cosmic shear signal \( \langle \gamma^2 \rangle \) for a given size box, \( \theta \). These calculations are repeated for different values of \( \theta \).

We have applied this procedure to the catalogs produced by the two pipelines and obtained the results in Figure 6.6. A significant cosmic shear signal is present in both sets of catalogs and the two curves in Figure 6.6.A and 6.6.B are in accordance. The error bars are obtained from 100 fake sets of catalogs in which the angular position of each object has been selected randomly. The cosmic shear signal from the shapelets catalogs is smoother, better resolved and with smaller error bars than the signal from the KSB catalogs. The main difference between the two catalogs is the number of objects present in each of them. During the data reduction process several constraints are applied to the data from the KSB pipeline. After each step of the KSB procedure, thresholds are used to discard unrealistic data.

- Once the objects are localized in the images and the estimation of their parameters (position, size, magnitude, ellipticity, smear and shear polarizability tensors, and a flag) is carried out, restrictions are applied to the majority of the parameters.

- After a PSF correction we removed from the catalogs those objects, which have shown unrealistic shear estimates. If \( \tilde{\gamma}_{1,2} \) is not within specific boundaries (0 \( \leq \tilde{\gamma}_{1,2} \leq 1 \)), the object is rejected.

During the complete procedure many objects are discarded. We also apply constrains to the data from the shapelets pipeline but the number of rejected objects is much smaller compared to the KSB pipeline. This difference shows that the shapelets pipeline is more robust than the KSB pipeline. In galaxy-galaxy lensing analyses, it is important to have a large number of objects to extract weak signals such as the cosmic shear signal. The lack of a sufficient number of objects in the KSB catalogs has an impact on the computation of the cosmic shear signal and the error bars.

In Figure 6.7 we compare the cosmic shear signal obtained by the pipeline based on the shapelets algorithm to the cosmic shear signal obtained by Maoli et al. (2000) and other surveys (van Waerbeke et al., 2000; Kaiser et al., 2000; Wittman et al., 2000). We observe that our data are consistent with the data from other surveys, which indicates that the technique based on the shapelets algorithm is a reliable method for obtaining the cosmic shear signal. These results are very promising, but they have to be confirmed
Figure 6.6: $\langle \gamma^2 \rangle$ with respect to the size, $\theta$, of the top hat filter. The shear variances have been computed within different square regions of each image with a size equals to the size of the top hat filter (black dots). The crosses correspond to the estimates of the ellipticity variance before any PSF correction. In A, the data from the shapelets pipeline have been used, and in B, the data from the KSB pipeline.

Figure 6.7: Comparison of the cosmic shear signal, obtained using the shapelets pipeline, with that obtained by other surveys.
by the analysis of possible systematic errors which could be present in our data and could mimic a cosmic shear signal.

6.6 Analysis of Systematic Errors

We apply the criteria used by Maoli & al. (2000) to check whether our data are affected by systematic errors and then to validate our calculations.

First we check whether there are systematic errors in the computed isotropic ellipticity, being introduced while reading the images recorded on the CCD. Unexpected CCD defects, such as low CTE and the presence of traps, can occur during the horizontal and vertical transfer of charges while reading the scientific images and these disturbances might not have been compensated. The corrected ellipticities are averaged in horizontal (Y) and vertical (X) bands and plotted for both catalogs (see Figure 6.8 and 6.9). No correlations between \( \langle e_{1,\text{galaxies}} \rangle \) and \( \langle e_{2,\text{galaxies}} \rangle \) are observed with respect to X and Y. Our data are scattered around \( \langle e_{i,\text{galaxies}} \rangle = 0 \) which proves that no systematic errors due to a possible charge transfer inefficiency are introduced in our cosmic shear calculations.

Another possible source of systematic errors is the procedure to correct the PSF itself. This procedure is designed to eliminate systematic errors, however a wrong manipulation of the data, estimate of the parameters, etc. could introduce an unexpected error, which resembles a cosmic shear signal. To test this possibility we plot the average ellipticity, before and after PSF anisotropic correction, as a function of the average star ellipticity (see Figure 6.10). Before any kind of corrections we observe that the ellipticities of the stars, and the galaxies are correlated, which correlation disappears after an anisotropic
PSF correction. The average ellipticity of the galaxies is scattered around 0. Since we do not have an access to the ellipticity of the galaxies after a PSF anisotropic correction within the shapelets method (the anisotropic and the isotropic corrections are carried out simultaneously), we plot instead the galaxy ellipticity before the correction as a function of the star ellipticity from the KSB estimate and the average shear estimate as a function of the star ellipticity from the KSB estimate (see Figure 6.11). We observe the same feature as in Figure 6.10 namely, the PSF correction based on the shapelets formalism is carried out properly and the estimated shears contain no systematic errors.

As explained by Van Waerbeke et al. (2000) and Maoli et al. (2001), and resumed here, the tests, performed above, are necessary to check if systematic errors are present in the catalogs, but they are not sufficient to confirm the absence of biases, particularly if the latter occur at a scale smaller than the smallest bin, used to determine the cosmic shear signal. A complete test consists of calculating the variance of the shear before any correction and comparing the obtained result with the cosmic shear signal (see section 6.5). The variance of the galaxy ellipticity before PSF correction is a constant value of about $3.5 \times 10^{-4}$ (see Figure 6.6) for both sets of catalogs from the KSB and shapelets pipelines. The variance of the ellipticity, obtained within the KSB and the shapelets formalisms, is independent of the variation in the size of the top hat filter. The observed constant variance is due to the PSF which is present in each image. As we can see in Figure 6.6, the variance of the ellipticity produced by the PSF is of the same order as the cosmic shear signal. This result demonstrates the importance of the PSF correction in the calculation of the cosmic shear signal.

The final test we performed consists of comparing our cosmic shear signal with 100
Figure 6.10: Average galaxy ellipticity, before (plots on the top) and after (plots on the bottom) the anisotropic PSF correction, as a function of the star ellipticity. The KSB algorithm is used here to correct the PSF. After anisotropic correction the correlation is no longer present.

Figure 6.11: Average galaxy ellipticity before (plots on the top) and after (plots on the bottom) PSF correction, as a function of the star ellipticity. The shapelets algorithm is used here to correct the PSF. After correction the correlation is no longer present.
realizations of cosmic shear calculation after rotating randomly the shear estimates in our catalogs (see Figure 6.12). None of those 100 calculations could mimic the cosmic shear signal. For each size of the top hat filter the cosmic shear signals obtained from the additional 100 calculations are distributed around zero and the maximum values are far below the real cosmic shear signal. All these tests permit to conclude that the cosmic shear signal, obtained using either the KSB or the shapelets formalisms, is not affected by systematic errors.

6.7 Conclusions

After providing good results on simulated data (Kuijken, 2006), the pipeline used to correct the PSF based on the shapelets formalism had to be tested on real data. We have selected the set of images used by Maoli et al. (2001) in order to perform such tests. Maoli et al. (2001) have extracted from these images a significant cosmic shear signal for sizes between 0.5 and 5 arcmin of the top hat filter. The present study has consisted of comparing the catalogs of data, obtained using the pipeline based on the shapelets algorithm with the catalogs of data yielded by the pipeline based on the KSB formalism. The latter pipeline is based on the well established method developed by Kaiser et al. (1995) also known under the acronym KSB. We found from our analysis that the compensation factor, introduced by Bacon et al. (2001), was indeed necessary to fit the shear estimates from the KSB pipeline to the shear estimates from the shapelets pipeline, but was not enough to compensate the correlation which exists between the shear estimate obtained using the KSB algorithm, and the seeing of the images. The compensation factor introduced by Bacon et al. (2001) has been replaced by a one-degree polynomial which is a function of the seeing (in arcsec), \( c(s) = 1 - 0.28s \) (\( s \in [0.4; 0.95] \))
6.7: Conclusions

in $\tilde{\gamma}_{ksh, corr}^{1,2} = \frac{\gamma_{ksh}^{1,2}}{c(s)}$. After having corrected the shear estimate of the objects in the catalogs produced by the KSB pipeline, we compared these shear estimates with the shear estimates obtained from the shapelets pipeline. In their vast majority the shear estimates from both pipelines are in accordance (see Figure 6.5).

For each set of catalogs we compute the cosmic shear signal $\langle \gamma^2 \rangle$. The cosmic shear signals, obtained after using the KSB and the shapelets pipelines to compute the shear estimate of the objects, are in accordance (see Figure 6.6). Using the shear estimates yielded by the shapelets pipeline, several improvements have been observed in the computation of the cosmic shear signal. In the process of applying the PSF correction, the shapelets pipeline has discarded less objects than the KSB pipeline which results in a smoother cosmic shear curve. As expected the pipeline based on the shapelets algorithm has shown a greater robustness and a significant reduction of the noise. The analyses of the possible systematic errors have confirmed our preliminary conclusions, namely there is a cosmic shear signal in the set of 50 images. In order to detect possible systematic errors in the shear estimates from the KSB and the shapelets pipelines, we carried out tests on these data analogous to the tests carried out by Maoli et al. (2001). These systematic error analyses have revealed that both pipelines provide shear estimates free of such errors. Finally a comparison of the cosmic shear signals obtained from the shapelets pipeline with those from other surveys (van Waerbeke et al., 2000; Kaiser et al., 2000; Wittman et al., 2000; Maoli et al., 2001) has shown that the results of the present study are in good agreement with the other analyses. From these results we can conclude that the shapelets algorithm can be successfully applied to correct the PSF and to extract the shear present in the objects. Consequently the shear estimates are employed to obtain a cosmic shear signal with a better resolution than that, yielded by the standard method based on the KSB formalism. The cosmic shear signal, we calculated using a completely different formalism for the correction of the PSF and for the computation of the shear estimates, provides an evidence that in the set of images considered in the present study, there are fluctuations of mass in large scale structures in the universe.

The shapelets algorithm appears to be less demanding in terms of a computation time spent for the data reduction. On equivalent computers, the pipeline based on the shapelets technique is faster than the pipeline using the IMCAT sub-routines. This point is crucial for the future surveys where a huge amount of data will be produced, e.g. the Kilo-Degree Survey, KiDS (Kuijken et al. 2007, in preparation) which is a large survey of 1500 squared degrees.