

Weak lensing PSF correction of wide-field CCD mosaic images

Marissa Cevallos

California Institute of Technology

Office of Science, SULI Program
Stanford Linear Accelerator Center
Menlo Park, California

August 20, 2005

Prepared in partial fulfillment of the requirements of the Office of Science, U.S. Department of Energy Science Undergraduate Laboratory Internship (SULI) Program under the direction of Dr. Phil Marshall, at the Kavli Institute for Particle Astrophysics and Cosmology (KIPAC) at the Stanford Linear Accelerator Center (SLAC).

Participant:

Signature

Research Advisor:

Signature

Contents

1	Introduction	4
1.1	Gravitational lensing	5
1.2	Characterizing the point spread function	5
2	Methods	6
2.1	Extracting galactic and stellar objects	6
2.1.1	Separating stars from galaxies	6
2.2	Object shape estimation and PSF correction	7
2.3	Efficient computing	7
2.4	Correlation functions	8
2.5	Mass maps	9
3	Results	9
4	Discussion and Conclusions	10
5	Acknowledgments	12
6	Figures	13

Weak lensing PSF correction of wide-field CCD mosaic images

Marissa Cevallos

California Institute of Technology, Pasadena, CA 91125

Abstract

Gravitational lensing provides some of the most compelling evidence for the existence of dark matter. Dark matter on galaxy cluster scales can be mapped due to its weak gravitational lensing effect: a cluster mass distribution can be inferred from the net distortion of many thousands of faint background galaxies that it induces. Because atmospheric aberration and defects in the detector distort the apparent shape of celestial objects, it is of great importance to characterize accurately the point spread function (PSF) across an image. In this research, the PSF is studied in images from the Canada-France-Hawaii Telescope (CFHT), whose camera is divided into 12 CCD chips. Traditional weak lensing methodology involves averaging the PSF across the entire image: in this work we investigate the effects of measuring the PSF in each chip independently. This chip-by-chip analysis was found to reduce the strength of the correlation between star and galaxy shapes, and predicted more strongly the presence of known galaxy clusters in mass maps. These results suggest correcting the CFHT PSF on an individual chip basis significantly improves the accuracy of detecting weak lensing.

1 Introduction

Dark matter is predicted to compose as much as 90 percent of the Universe's baryonic matter, yet experimentalists have yet to detect the actual dark matter particle. Dark matter's presence is suggested in the analysis of the large-scale structure of the Universe, but was first postulated by Fritz Zwicky in the 1930s as non-luminous material to account for the extra gravitational pull in the Coma cluster of galaxies. When scientists in the 1960s began making galactic observations similar to Zwicky's, the concept of dark matter became a real cosmological area of interest, and alternate methods were sought to detect its elusive presence.

Weak gravitational lensing is one of these methods. In weak gravitational lensing, dark matter halos act as lenses to distort light from background sources. The lensing effect is similar to that of classical optics, but caused by the gravitational field of the dark matter—Figure 1 depicts an object that has been lensed by a massive source. Weak gravitational lensing attempts to use a net statistical lensing distortion over a field with large numbers of galaxies to predict the regions and amounts of dark matter. We assume that in a large sample of elliptical galaxies *without* gravitational lensing that the average shape of a galaxy would be a circle¹. Therefore, any net ellipticity is a lensing effect of primarily dark matter (1).

Weak gravitational lensing has only recently stemmed from its precursor, strong gravitational lensing. Strong gravitational lensing occurs when a single massive object, such as a star or galaxy, bends light from a source to form another image. These curved images are known as *Einstein rings*, as they were predicted by Einstein's theory of general relativity. The first such lens was discovered in 1979 at Kitt Peak Observatory, though Fritz Zwicky confidently and correctly wrote forty years earlier that such a lens would not be difficult to detect (2). The important difference between weak and strong gravitational lensing is the absence of Einstein rings in weak methods. One cannot distinguish a single weakly lensed galaxy from a single unlensed galaxy—it is only by understanding the correlation between thousands of galaxy shapes that one can observe a weak lensing signal.

Recent observational progress in weak lensing has been rapid, with good use being made of the large format CCD mosaic cameras now installed at many telescopes: these permit the observation of a good fraction of a square degree of sky in a single exposure, and so are well-suited to weak lensing observations requiring the measurement of many thousands of background galaxy shapes.

¹ In three dimensions, a sphere, but our images are two-dimensional.

In particular, a wide field of view is required to match the outer radii of nearby clusters, for which good data exist in other wavebands but have been lacking a lensing mass determination (1).

1.1 Gravitational lensing

Ellipticity is a critical property of the weakly lensed galaxies that we seek to measure. The ellipticity of an object due to gravitational influence can be derived from a few equations. Assuming an ellipse with a semi-major axis a , semi-minor axis b , and orientation angle ϕ , the ellipticity ϵ of an object is given by

$$\epsilon = \frac{a - b}{a + b} e^{2i\phi}. \quad (1)$$

This ellipticity is what we measure for each galaxy. For a large number of galaxies, we can estimate the ensemble average ellipticity, g , or reduced shear:

$$\langle \epsilon \rangle = g. \quad (2)$$

The reduced shear is proportional to the shear γ and convergence κ by

$$g = \frac{\gamma}{1 - \kappa}, \quad (3)$$

where γ and κ are linear combinations of second derivatives of the projected, scaled gravitational potential: κ is in fact proportional to the projected total mass density. Equations (1), (2), and (3) provide the link between measurable ellipticities and projected mass distribution.

1.2 Characterizing the point spread function

As weak lensing is an intrinsically weak effect, slight error might distort what we interpret as gravitational lensing—these problems hail mostly from the detector and atmospheric distortion. A space-based telescope, such as the Hubble Space Telescope, can ignore the problems associated with atmospheric distortion, but the images we consider in this research come from the ground-based Canada-France-Hawaii Telescope (CFHT). Therefore, it is necessary to characterize the point-spread-function (PSF) in order to describe the unwanted distortion and eliminate it from the fields of interest. The PSF is convolved with the true, undistorted image to produce the image that is actually seen by the detector. Because we know nothing of the intrinsic ellipticities of the galaxies we study, we must calibrate the galaxy field using a PSF derived from images of stars (which are intrinsically point-like).

The images we consider in this research come from the Canada-France-Hawaii Telescope (CFHT) in Hawaii. The CFHT uses the CFH12K camera (3), a CCD mosaic camera with 12 chips arranged in two rows of six as seen in Figure 1. Though there is no reason to suspect that atmospheric distortion will be more prevalent in one chip over another, the flaws in individual pixels may present errors unique to each chip. Moreover, slight deviations from planarity across the mosaic could introduce systematic ellipticity in the PSF—the CFH12K camera has a tilt height of about 20 microns. However, when performing weak lensing analysis, one typically corrects the PSF averaged over the entire image. This research aims to characterise the PSF in each chip, as we suspect the concatenation of the 12 PSF corrections will be different from the PSF of the entire image.

2 Methods

2.1 Extracting galactic and stellar objects

In order to create a comprehensive catalog of objects in a given field, it is necessary to extract objects through an automated method. `Source Extractor`, or `SExtractor`, is a program that estimates shapes rather poorly, but accurately and rapidly locates the centers of all objects in an image – it is the (x, y) coordinates we are interested in determining from `SExtractor`. `SExtractor` creates a catalog with not only the position of each object, but the Full Width at Half Maximum (FWHM) of the object’s brightness profile. This is useful information—we will use it to differentiate between stars and galaxies.

2.1.1 Separating stars from galaxies

The question of how to separate stars from galaxies in an object catalog is a pertinent one. Because the PSF of the stars will be used to calibrate the PSF of the galaxies, it is important that one maintains two catalogs for stars and galaxies. Unfortunately, with limited resolution from the detector, it is often unclear whether a high magnitude object is a star or a galaxy, since there is no definitive algorithm for such a classification. However, it is known that stars must appear point-like, such that stars at varying magnitudes will have about the same full width at half maximum (FWHM), while the FWHM of a galaxy will vary along with its magnitude. By plotting magnitude against FWHM, regions of objects with similar FWHM values can be selected from the graph and placed in a star catalog.

In Figure 3, magnitude and FWHM are compared for a given catalog containing both stars and galaxies. Though stars may be present throughout the plot, it is a safer bet to assume that all of the objects in the vertical band—the stellar locus—between about $x = 4$ and $x = 5$ are stars which reached the CCD’s saturation point. Under close inspection, one notices that Figure 3 has a double stellar locus, or two vertical bands. After closer inspection², it appeared that both bands contained stellar objects, so the chosen star catalog was composed of objects from both bands.

It is trickier to extract galaxies. Not all of the “leftover” objects in Figure 3 are galaxies—the lower left corner between $x = 0$ and $x = 4$ is most likely full of very faint stars, galaxies, and noise peaks, which are undesirable in weak lensing. Galaxies were chosen by constructing a rectangular region to the right of the stellar locus, but above very high magnitudes (note that the vertical axis in Figure 3 is reversed, running up from high magnitude to low magnitude). It is almost inevitable that one will include stars in the galaxy catalog and likewise mask potential galaxies by labelling them as stars. This small error is expected and does not significantly change the result of the weak lensing data.

2.2 Object shape estimation and PSF correction

Objects detected with SExtractor were re-measured using the tools in the imcat package (4), yielding accurate ellipticities for both stars and galaxies. The PSF patterns were characterized by applying a polynomial fit of degree n to give a smoothly interpolated stellar ellipticity pattern – in this research, we take $n = 5$. The Perl script `psfcorrect` performs this polynomial fit, and then uses the model stellar ellipticity field to correct the catalog of galactic objects, using the scheme derived by (5). The corrected galaxy ellipticities are then taken to be estimates of the local weak lensing shear.

2.3 Efficient computing

Analyzing an image for weak lensing involves both manual and automated methods—the researcher inevitably has to mask images by hand to eliminate spurious objects, but after selecting stars and galaxies, the process of creating a mass map involves no manual decisions. Thus, creating an algorithm to automate the process in between is a time-saving task that minimizes the chance that the researcher introduces more error. In addition, because we are not only interested

² We select objects in one band and replot them over the original image to see whether they correspond to stellar objects; we repeat for the other locus.

in the PSF correction across the total image, but the correction in all 12 chips, analyzing each chip individually is a tedious process. For the purposes of this research, several Perl scripts were written to run pre-existing code. Most involved reading in a directory of catalogs, and then iteratively running the appropriate process for each catalog.

Code now exists to divide a catalog into separate catalogs according to chip, run `psfcorrect` on these individual catalogs, and concatenate the catalogs into one catalog (of the same size as the original). After running the scripts, the only difference between the resulting catalog and one run through `psfcorrect` normally is when the objects underwent PSF correction chip-by-chip or with the entire catalog.

2.4 Correlation functions

A correlation function relates how closely two quantities are correlated. If the PSF alters the apparent ellipticity of a star, then every nearby object should be more or less affected in the same way. This is a correlation we can measure and attempt to reduce by correcting the galaxy images for the PSF.

The ellipticity vectors of two objects can be compared by taking the dot product between them. For two ellipticity vectors ϵ^a and ϵ^b with components ϵ_1 and ϵ_2 , the dot product is given by

$$\epsilon^* = \epsilon^a \cdot \epsilon^b = |\epsilon_1^a \epsilon_1^b + \epsilon_2^a \epsilon_2^b|.$$

We take the magnitude because we do not care about the direction of the vectors: if ϵ^* is 1 or -1, then the vectors are parallel. However, the ellipticity values found by `psfcorrect.pl` are not strictly vectors, so we have to use a more sophisticated product that is dependent on the position angle ϕ between the two objects.

Incorporating the angle between ellipticity “vectors”, we have

$$\gamma\gamma_+ = (\epsilon_1^A \cos 2\phi + \epsilon_2^A \sin 2\phi) \times (\epsilon_1^B \cos 2\phi + \epsilon_2^B \sin 2\phi) \quad (4)$$

$$\gamma\gamma_x = (\epsilon_1^A \sin 2\phi - \epsilon_2^A \cos 2\phi) \times (\epsilon_1^B \sin 2\phi - \epsilon_2^B \cos 2\phi) \quad (5)$$

in which $\gamma\gamma_+$ represents the “E-mode” and $\gamma\gamma_x$ is the “B-mode” correlation. In weak lensing with exactly calculated PSF corrections, the correlation function given by the B-mode should be zero, while the E-mode should be some non-zero (positive) number. If the two quantities compared are stars against stars, then the E-mode and B-mode correlation strengths should be ap-

proximately equal, assuming there is no reason for stars to be aligned in one formation more than the other.

The star-galaxy correlation function will show a similar, although much noisier shape, since the galaxies (before PSF correction) have ellipticities partially due to the PSF. After a perfect PSF correction, the star-galaxy correlation function should fall to zero, as there is no other physical connection between the apparent shapes of the stars and galaxies. If there *is* a correlation between the two, then this indicates that the applied PSF correction was not accurate enough.

A Perl script was written to read in two catalogs and computes both their E-mode and B-mode correlation functions, comparing the ellipticity between every pair of objects using Equations (4) and (5) and binning the pairs as a function of the distance by which the pair is separated.

For a pair of galaxy catalogs with thousands of objects, the required run-time (which scales as $O(n^2)$) presented a difficulty; an option was introduced to eliminate duplication of comparisons, lowering the run time when using two identical catalogs by a factor of two. However, in investigating the correlation between stars and galaxies, the star catalog generally contains an order of magnitude less objects, so the runtime is reasonable.

2.5 Mass maps

Though it is straightforward to predict galaxy ellipticities from a mass distribution, performing the inverse operation is non-trivial. We must first crudely estimate mass from ellipticities, and then iteratively refine the mass map to improve the match between predicted and observed ellipticities. One such code to do this is `LensEnt2` (6). More specifically, `LensEnt2` infers κ from solving the system of equations (Equations (1), (2), and (3)), but because of the presence of noise, the solution must be found iteratively.

In order to compare the chip-by-chip analysis to the standard method of weak lensing, we may use mass maps as a test. This requires running both catalogs through `LensEnt2` and compare the resulting mass maps. If one method clearly predicts known mass clusters more than the other, then it may be the more useful method.

3 Results

In order to compare the PSF pattern of the entire detector to that of the individual chips, three approaches were used: plotting ellipticity components, analyzing the correlation function between

catalogs of stars and galaxies, and testing the level of observed signal (believable and spurious) in the resulting mass maps. A wide-field image of Abell 399 was used for the tests.

Figure 4 presents four plots of ellipticity against position, with the magnitude and angle of the ellipticity represented by sticks centered on the object positions. The left-most plots show stellar ellipticities before PSF correction, and the right-most plots shows ellipticities after PSF correction. The top-right plot has chip-by-chip PSF correction, while the bottom-right plot has the traditional whole-image PSF correction.

In Figure 5, a statistical argument is unnecessary to convince the reader that the correlation functions resulting from the chip-by-chip method are less noisy than the original method. All four graphs are correlation functions between the same stars and galaxies. The top two plots used catalogs that had corrected for the PSF on a chip-by-chip basis; the bottom two plots used catalogs with traditional PSF corrections. The plots to the left are of E-mode correlation, and the ones to the right are of B-mode correlation: both show a significant improvement when moving to the chip-by-chip correction scheme.

LensEnt2 generates a reconstructed projected mass map at a user-defined angular resolution: typically the data quality favours a low resolution map, but the higher resolution maps provide useful diagnostics when trouble-shooting. Figure 6 shows mass maps derived from chip-by-chip catalogs and traditional catalogs. The top two maps are at a higher angular resolution (3 arcmin FWHM) than the bottom two (9 arcmin FWHM). The left-most plots are from chip-by-chip corrected catalog, while the right-most maps are from the traditionally-corrected catalog. The red outline is tilted by about 5 degrees to compensate for the tilt in the original image. The center of the mass maps is the position of the brightest cluster galaxy.

4 Discussion and Conclusions

In two of the three tests—correlation functions and mass maps—the PSF correction appears to have been significantly improved by splitting the image by chip. It is especially evident in Figure 5 that the chip-by-chip plots are less noisy than the other two. In fact, it even appears that the bottom two plots still contain some smooth, large-scale PSF variation, which is exactly what the PSF correction tries to eliminate. Because such a pattern is not present in the chip-by-chip plots, one can conclude that the PSF correction in these plots is better.

It is not obvious whether the ellipticity plots show improvement because of the presence of

high-ellipticity objects that were not filtered out during an earlier stage of the analysis. The results are reasonable, however—the left-most plots are identical (they both represent the uncorrected ellipticities), while the plots on the right are difficult to distinguish between and would require further study in order to reach solid conclusions about the effectiveness of splitting catalogs by CCD chips.

While a quick glance at the correlation functions can show that one is clearly better, it is not so intuitive to understand the mass maps. Because the mass maps are supposed to show where regions of dark matter are located, and we cannot see dark matter, it seems to be difficult to gauge whether a mass mapping was successful. However, we can look at where we know there are high-density regions of mass (for example, the central galaxy cluster) and compare how strongly the mass map predicts its presence. The left-most maps, from chip-by-chip catalogs, more strongly predict a single mass in the center than the other maps—the right-most maps predict a single mass in the center, but with fewer contour lines (bottom right of Figure 6) or as two masses in the center (top right of Figure 6). It would be too great a coincidence to have the PSF correction of 12 different chips conspire to give a residual PSF pattern that mimicked and increased weak lensing signal at the center of the map!

We must then ask why these differences arise: the 20 microns in tilt height is perhaps a factor, as is the make-up of individual chips. If a better PSF correction is obtained simply by regionalizing the image, but not necessarily by chips, then one could entertain the possibility of splitting the chips into even smaller regions. However, this leaves each region with fewer objects with which to do a PSF correction. The problem then becomes one of optimization—at what point (ie, how many galaxies, or at what fraction of the sky) does it become disadvantageous to split the image into smaller regions? This could be explored in future work.

It may be the case that a higher order polynomial fit ought to be used. Future work should test the outcomes of PSF corrections with n degree fits other than $n = 5$. It is not always advantageous to use very high order polynomials: we are seeking a smooth characterization of the PSF, not one that fluctuates with the stellar shape estimation noise. However, it may be that the tests devised here would show further improvement in the PSF correction if the polynomial order were to be varied. The analysis pipeline put in place in the course of this work will allow this, and other subtle effects, to be rigorously investigated.

5 Acknowledgments

I would like to thank the Department of Energy Office of Science for the funding to participate in the SULI program – it is a wonderful program and great learning experience. I would like to thank my mentors Phil Marshall and Maruša Bradač and for their time, patience, and good advice. In addition, I would like to thank Helen Quinn and James Lindesay for their tireless efforts in running the SULI program, and the Stanford Linear Accelerator Center for hosting me. This work was supported in part by the U.S. Department of Energy under contract number DE-AC02-76SF00515.



References

- [1] Philip Marshall. *Bayesian Analysis of Clusters of Galaxies*. PhD thesis, Cavendish Astrophysics and King's College, Cambridge, September 2003.
- [2] Peter Schneider. Introduction to gravitational lensing and cosmology. In *Gravitational Lensing: Strong, Weak and Micro*. Springer-Verlag: Berlin, 2003.
- [3] Jean-Charles Cuillandre. Cfh12k: a 12,288 by 8,192 pixel ccd mosaic camera for high resolution wide field imaging.
- [4] Nick Kaiser. Available as <http://www.ifa.hawaii.edu/kaiser/imcat/>.
- [5] N. Kaiser, G. Squires, and T. Broadhurst. A method for weak lensing observations. *The Astrophysical Journal*, pages 460–475, August 1995.
- [6] Phil Marshall. Lensed version 2. Available as <http://www.mrao.cam.ac.uk/projects/lensed/version2/>.

6 Figures

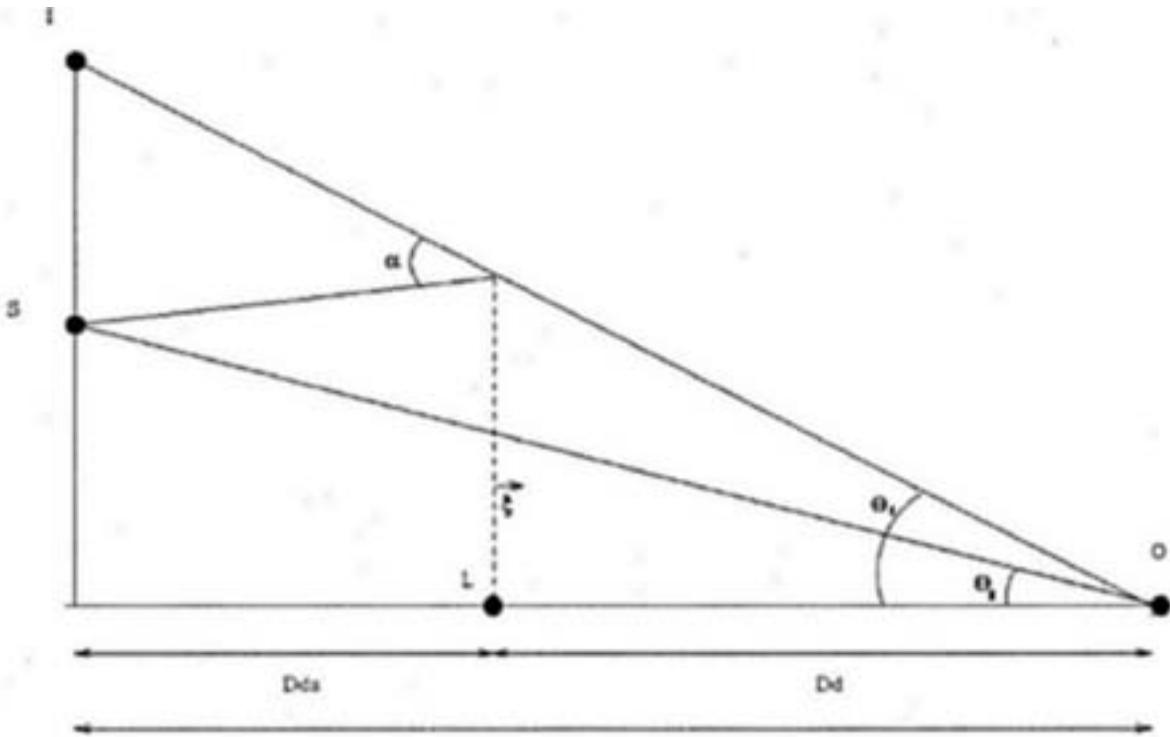


Figure 1: Gravitational lens geometry. D_s is distance between observer and source; D_d is the distance between observer and lens, and D_{ds} is the distance between the source and lens. In the absence of the lens, the angle between the observer and source is θ_S ; with the lens, the image appears at an angle of θ_I .

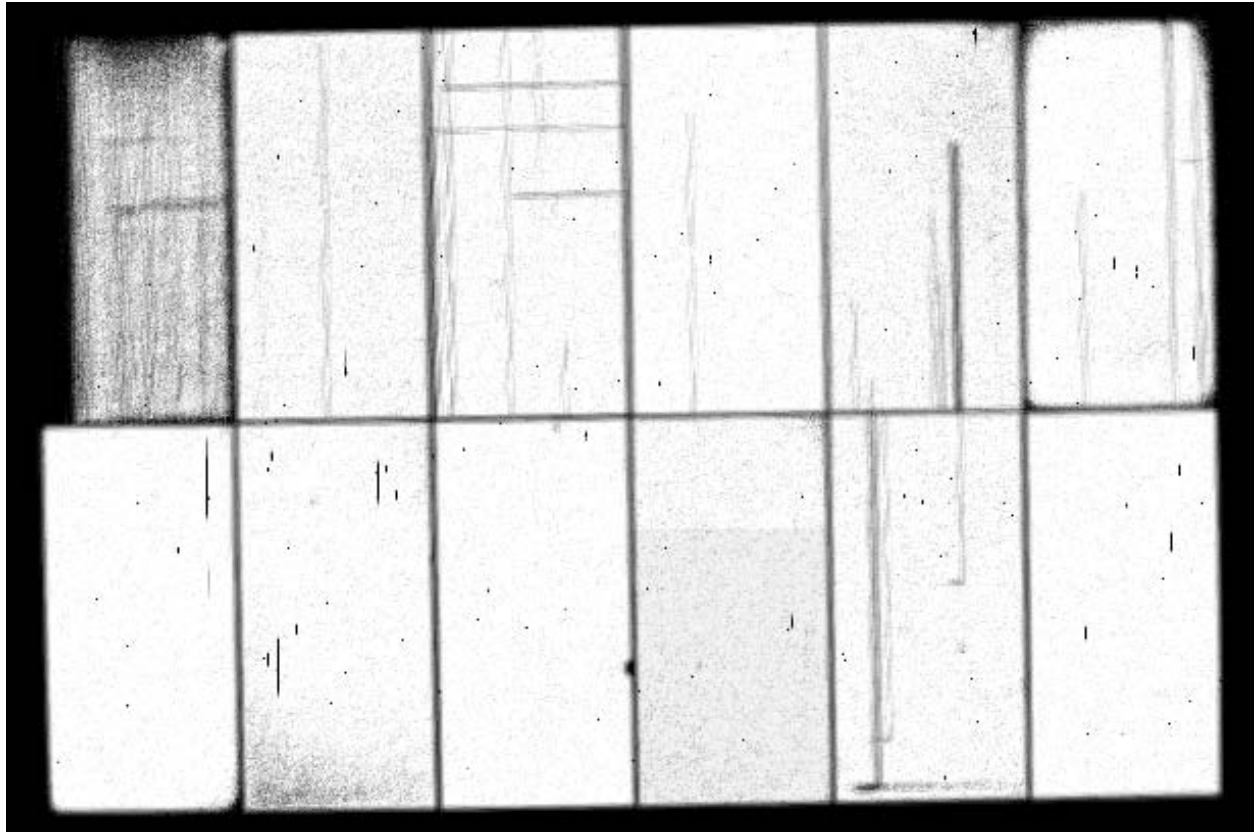


Figure 2: Weight image of A399. The gaps between chips are clearly visible as under-exposed regions.

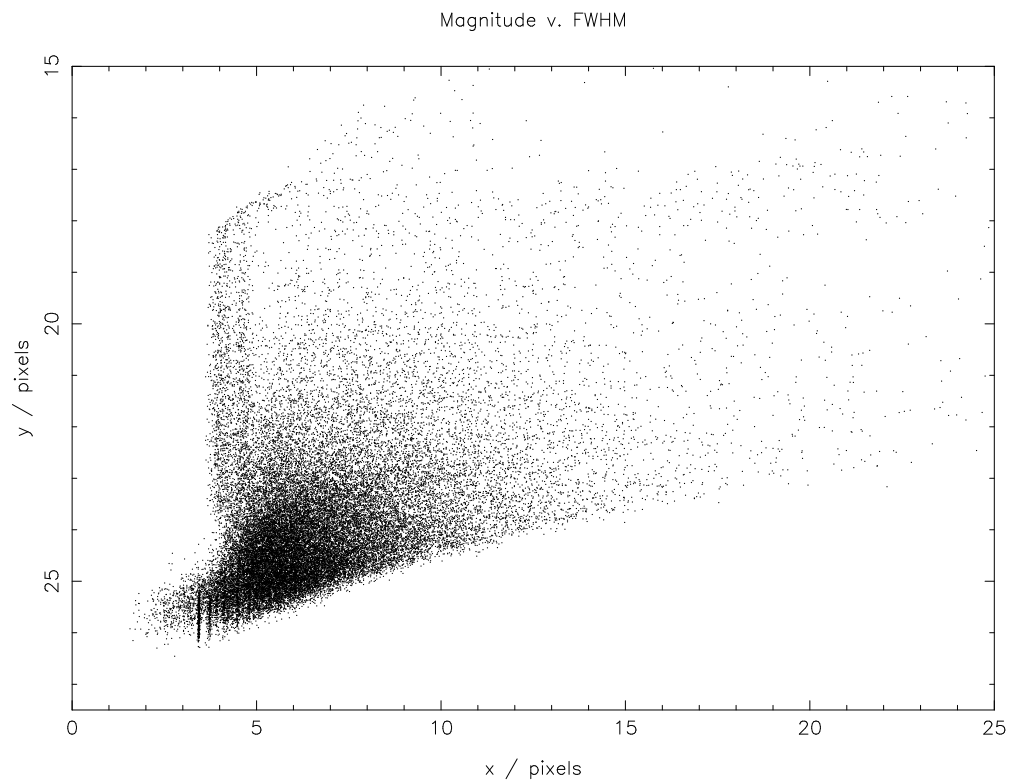


Figure 3: Object magnitude vs. FWHM. Stars are selected from the vertical band between about $x = 4$ and $x = 5$.

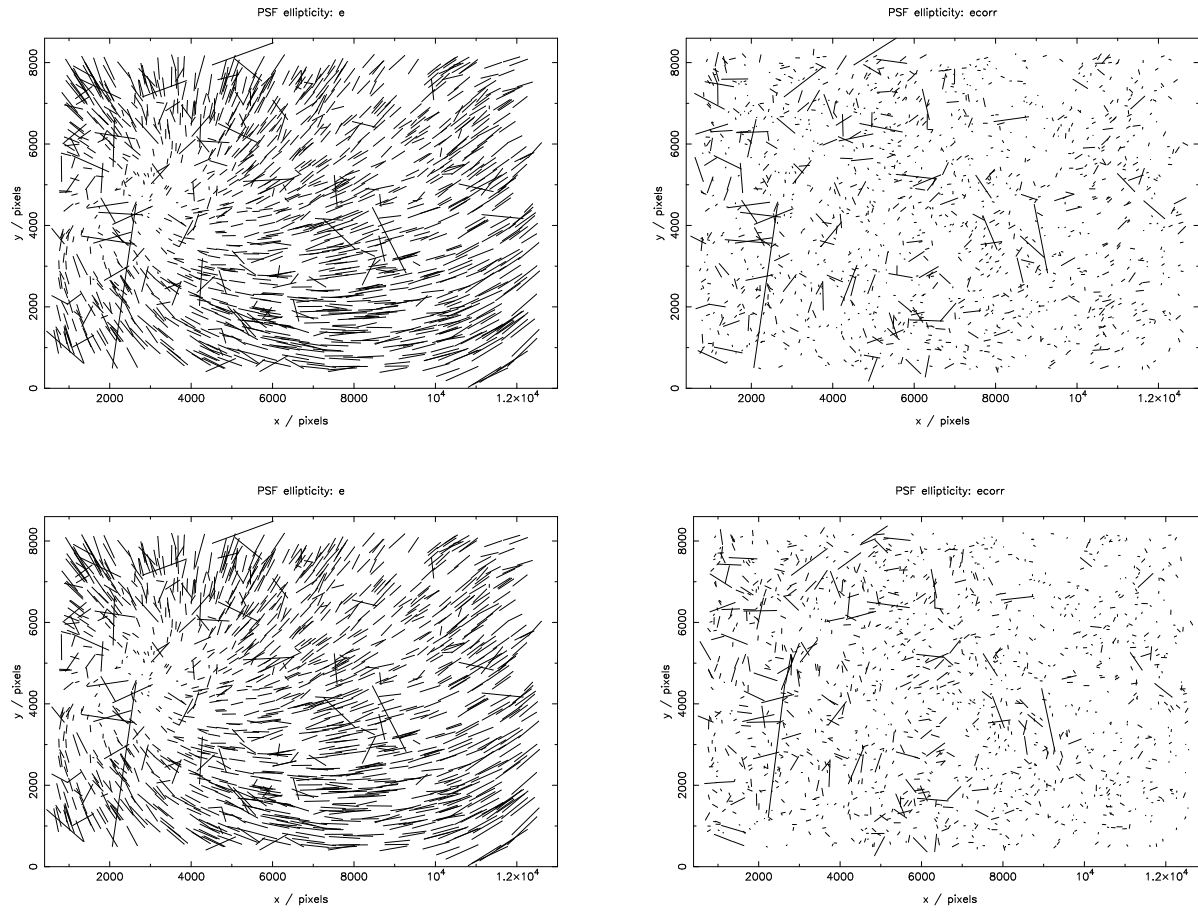


Figure 4: Stellar ellipticity vs. position. Left-hand plots show raw ellipticity plotted as vectors; the plots on the right show the same stars after ellipticity correction. Top row: traditional whole-image correction; bottom row: chip-by-chip analysis.

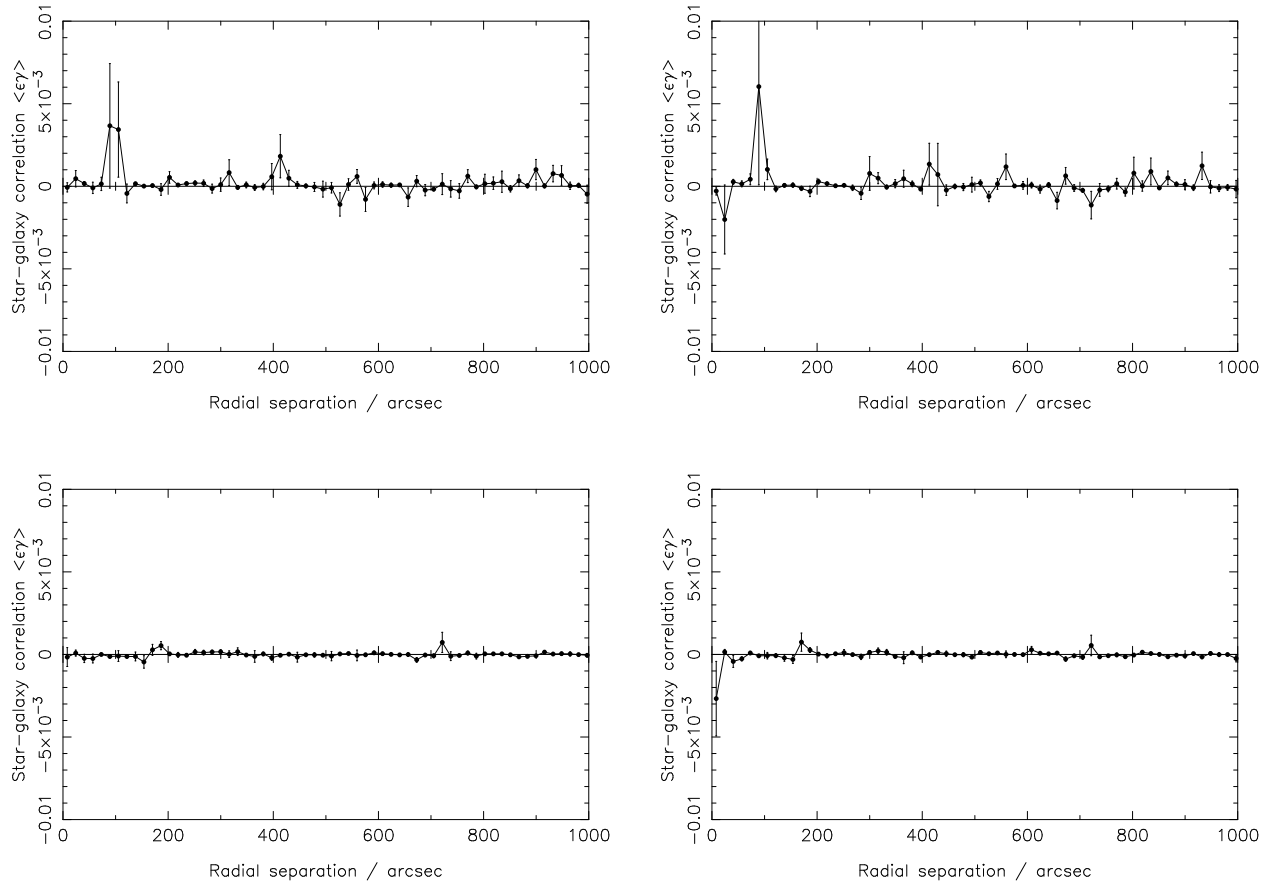


Figure 5: Correlation functions between star and galaxy catalogs. Left-hand plots show E-mode correlation, right-hand ones show B-mode. Top row: traditional whole-image correction; bottom row: chip-by-chip analysis.

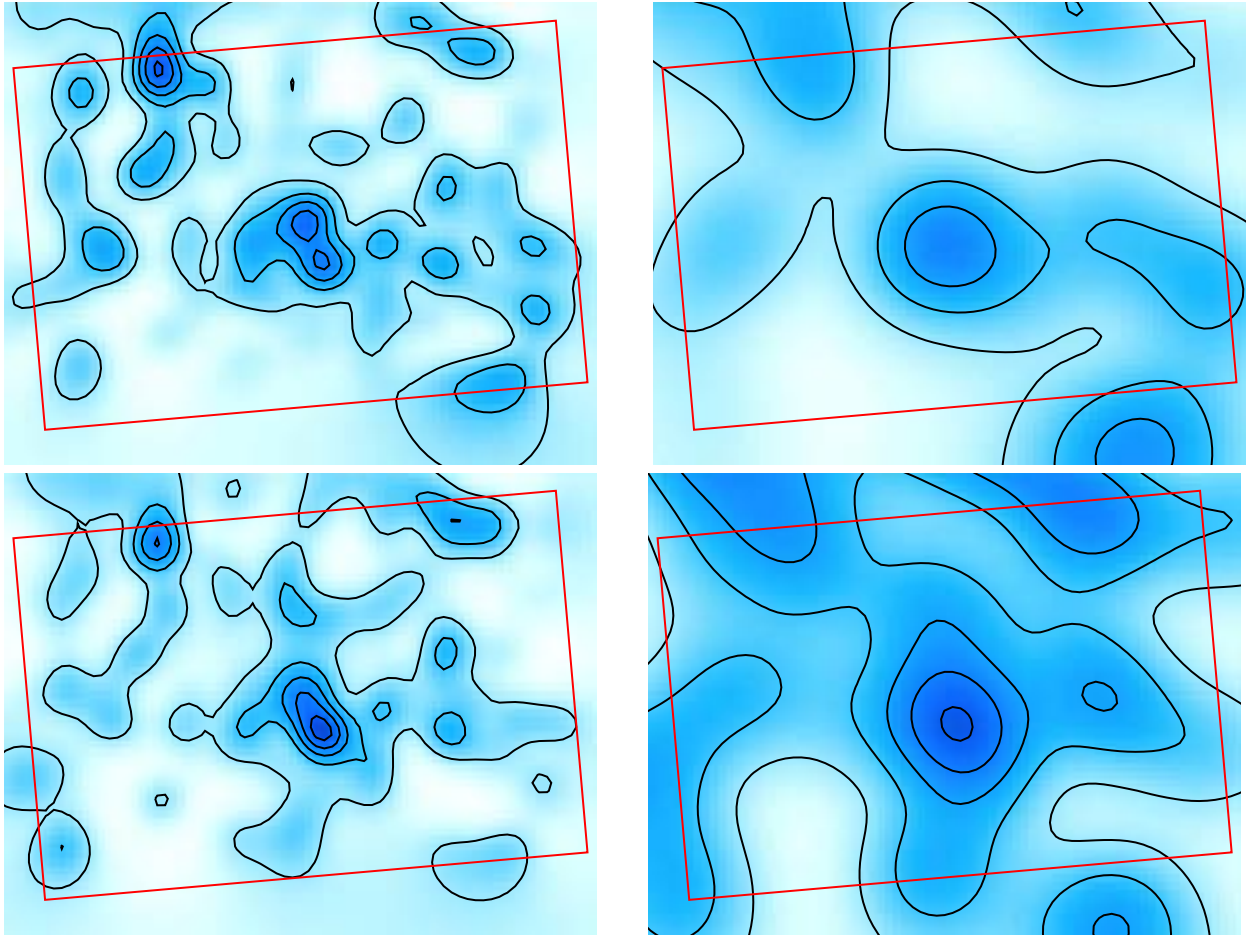


Figure 6: LensEnt2 mass maps inferred from the PSF-corrected A399 galaxy catalogues. Left-hand maps are at 3 arcmin resolution, right-hand ones 9 arcmin. Top row: traditional whole-image correction; bottom row: chip-by-chip analysis.