



Universiteit  
Leiden  
The Netherlands

## Light weighed: on the statistics and systematics of weak gravitational lensing

Smit, D.M.

### Citation

Smit, D. M. (2021, December 8). *Light weighed: on the statistics and systematics of weak gravitational lensing*. Retrieved from <https://hdl.handle.net/1887/3245825>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3245825>

**Note:** To cite this publication please use the final published version (if applicable).

---

## Chasing the peak: optimal statistics for weak shear analyses

---

Weak gravitational lensing analyses are fundamentally limited by the intrinsic distribution of galaxy shapes. It is well known that this distribution of galaxy ellipticity is non-Gaussian, and the traditional estimation methods, explicitly or implicitly assuming Gaussianity, are not necessarily optimal.

We aim to explore alternative statistics for samples of ellipticity measurements. An optimal estimator needs to be asymptotically unbiased, efficient, and robust in retaining these properties for various possible sample distributions. We take the non-linear mapping of gravitational shear and the effect of noise into account. We then discuss how the distribution of individual galaxy shapes in the observed field of view can be modeled by fitting Fourier modes to the shear pattern directly. This allows scientific analyses using statistical information of the whole field of view, instead of locally sparse and poorly constrained estimates.

We simulated samples of galaxy ellipticities, using both theoretical distributions and data for ellipticities and noise. We determined the possible bias  $\Delta e$ , the efficiency  $\eta$  and the robustness of the least absolute deviations, the biweight, and the convex hull peeling estimators, compared to the canonical weighted mean. Using these statistics for regression, we have shown the applicability of direct Fourier mode fitting.

We find an improved performance of all estimators, when iteratively reducing the residuals after de-shearing the ellipticity samples by the estimated shear, which removes the asymmetry in the ellipticity distributions. We show that these estimators are then unbiased in the absence of noise, and decrease noise bias by more than  $\sim 30\%$ . Our results show that the convex hull peeling estimator distribution is skewed, but still centered around the underlying shear, and its bias least affected by noise. We find the least absolute deviations estimator to be the most efficient estimator in almost all cases, except in the Gaussian case, where it's still competitive ( $0.83 < \eta < 5.1$ ) and therefore robust. These results hold when fitting Fourier modes, where amplitudes of variation in ellipticity are determined to the order of  $10^{-3}$ .

The peak of the ellipticity distribution is a direct tracer of the underlying shear and unaffected by noise, and we have shown that estimators that are sensitive to a central cusp perform more efficiently, potentially reducing uncertainties by more than 50% and significantly decreasing noise bias. These results become increasingly important, as survey sizes increase and systematic issues in shape measurements decrease.

M. Smit, and K. Kuijken  
*Astronomy & Astrophysics*, **Volume 609**, A103 (2018)

## 4.1 Introduction

Since the first gravitational shear detections (Tyson et al. 1990), the statistical analysis of weak gravitational lensing effects has become recognized as a competitive cosmological tool. With the advent of precision cosmology, meaningful interpretations of statistical agreement or tension between various models and datasets become increasingly important.

Weak gravitational lensing produces slight magnification and distortion effects by bending the paths of light rays. Although analyses of the former have produced important scientific results (e.g., Hildebrandt et al. 2009, Van Waerbeke et al. 2010) and it has in fact been demonstrated that combined analyses can give better constraints (Hildebrandt et al. 2011, Ford et al. 2012), most scientific information has come from the analysis of weak shear distortions. To access that information, one has to be able to (1) measure the shapes of lensed background sources accurately, (2) understand the intrinsic distribution of these shapes and the effects of shear and noise on statistical inference, and (3) obtain the statistical power to probe the subtle perturbations of this distribution by weak shear.

For the first part, a multitude of shape measurement methods have been explored, among which are foremost methods based on surface brightness moments (e.g., Kaiser et al. 1995, Rhodes et al. 2000) and model fitting methods (e.g., Kuijken 1999, Bernstein & Jarvis 2002, Hirata & Seljak 2003, Refregier & Bacon 2003, Kuijken 2006, Miller et al. 2007, Kitching et al. 2008), with various alternative or combined approaches (Bernstein & Armstrong 2014, Herbonnet et al. 2017, Zhang et al. 2015).

Community-driven projects for optimal and robust shape estimates (Heymans et al. 2006, Massey et al. 2007, Bridle et al. 2010, Kitching et al. 2012, Mandelbaum et al. 2015) have led to a further decrease in measurement variances and a better understanding of remaining systematic effects and biases (e.g., Voigt & Bridle 2010, Bernstein 2010, Kacprzak et al. 2012, Melchior & Viola 2012, Refregier et al. 2012).

For the last part, the last two and a half decades have also known dramatic improvements in statistical power. Surveys that are finished, ongoing, and planned such as COSMOS<sup>1</sup> (Leauthaud et al. 2007), CFHTLenS<sup>2</sup> (Heymans et al. 2012b), RCSLenS<sup>3</sup> (Hildebrandt et al. 2016), KiDS<sup>4</sup> (de Jong et al. 2013), DES<sup>5</sup> (Dark Energy Survey Collaboration et al. 2016), LSST<sup>6</sup> (Ivezić et al. 2019), Euclid<sup>7</sup> (Laureijs et al. 2011) steadily increase in size (sky coverage and depth) and imaging quality, including a significant improvement in understanding and correcting for systematic effects (e.g., Heymans et al. 2012a,b, for CFHTLenS).

This increasing statistical power is necessary to overcome the inference limit set by the intrinsic galaxy shape distribution, known as shape noise. Unlike many forms of noise, such as measurement uncertainties that are often dominated by Poisson processes, there is no reason that the ellipticities of background galaxies follow a Gaussian distribution. In fact, studies of galaxy morphologies (Lambas et al. 1992, Rodríguez & Padilla 2013) suggest that late type galaxies may exhibit a roughly uniform axis ratio distribution.

This departure from Gaussianity is clearly demonstrated in Section 4.2, when comparing

---

<sup>1</sup><http://cosmos.astro.caltech.edu/>

<sup>2</sup><http://www.cfhtlens.org/>

<sup>3</sup><http://www.rcslens.org/>

<sup>4</sup><http://kids.strw.leidenuniv.nl/>

<sup>5</sup><http://www.darkenergysurvey.org/>

<sup>6</sup><https://www.lsst.org/>

<sup>7</sup><http://www.euclid-ec.org/>

the shape distribution of the CFHTLenS shape measurements catalog (Heymans et al. 2012b, Figure 4.2) to a simulated Gaussian distribution (Figure 4.1). This implies that commonly used Gaussian estimators, such as the (weighted) mean estimate of the central peak of the distribution or the variance for its width, are not necessarily optimal for the inference of the underlying gravitational shear.

For example, if the tails of the ellipticity distribution decline more slowly than the Gaussian  $\exp(-x^2)$ , then more elliptical galaxies contribute more shape noise. There have been many weighting and clipping schemes suggested to minimize biases and uncertainties in weak shear inference (Bonnet & Mellier 1995, Van Waerbeke et al. 2000, Bernstein & Jarvis 2002). Alternative approaches include distribution symmetrization (Zhang et al. 2017), or using ensembles of galaxies in Bayesian analyses or nulling techniques (Bernstein & Armstrong 2014, Herbonnet et al. 2017), so that the step of individual shape measurement before inference of the underlying shear is bypassed.

In this article, we explore an alternative approach by reviewing statistical estimators that are more suited to a distribution with a pronounced central cusp and slowly declining tails. Estimator optimality would include a low or vanishing estimator bias and a high accuracy by a low spread in estimates. These aspects should be robust for various possible distributions, as samples of background galaxies are comprised of different populations.

We then highlight the use of these estimators in fitting the shear pattern in the field of view with Fourier modes (Fourier Mode Fitting, FMF). This approach provides an alternative to smoothed gridding and locally sparse and therefore poorly constrained estimates. It provides statistical information constrained by the whole field of view, and incorporates fluctuations in background number densities and estimated measurement uncertainties automatically. For subsequent scientific analyses, the Fourier model allows for relatively straightforward, analytic approach to fundamental quantities, such as a power spectrum or mass density reconstruction.

We note that we focus on the statistical inference from samples of measured shapes, for various possible intrinsic shape distributions, that is, the propagation of shape noise. This is a single but fundamental step in improving the accuracy and fidelity of weak lensing analyses. We do not perform a subsequent cosmological analysis, which would require addressing other well-known sources of bias and systematic effects. These include for example selection and detection biases (e.g., Hirata et al. 2004, Miller et al. 2013, Jarvis et al. 2016) among others on the instrumental and computational side. Other sources include physical effects that affects the interpretation of the measured signal, such as the effects of baryons, or the redshift distribution and intrinsic alignments of lensed background galaxies background sources. The shear signal we recover in this paper would represent a combined signal, which would then need to be interpreted.

The remainder of this paper is organized as follows. We will briefly review the necessary definitions of galaxy shapes and the weak lensing formalism in Sect. 2, referring the reader to excellent reviews such as Bartelmann & Schneider (2001), Schneider (2006), Hoekstra & Jain (2008), for more in-depth approaches. We review the necessary statistical framework in Sect. 3, where we discuss galaxy shape distributions and statistical estimators, including definitions for efficiency and bias, before expanding on FMF. In Sect. 4 we describe the various possible simulations and data, and analysis methods. In Sect. 5 we discuss the results and the scientific implications. Section 6 gives a summary of our conclusions.

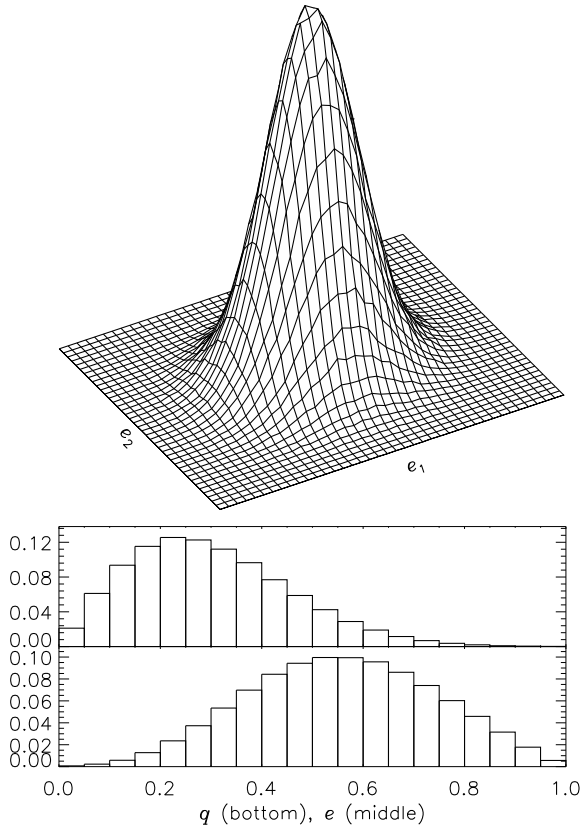


Figure 4.1: Gaussian ellipticity distribution and corresponding axis ratio distribution. Top: a 2D histogram of ellipticities. Middle: histogram of the absolute ellipticity  $|e|$ . Bottom: histogram of the ellipse axis ratio  $q$ .

## 4.2 Weak lensing

Gravitational lensing is the effect of curved space-time on the paths of light rays from distant sources to the observer as they pass through the gravitational potential of foreground structures. This geometrical effect leads to a displacement of point sources on the projected plane of the sky. The differential effect on images  $I(x, y)$  of extended sources leads to magnification and distortion effects, known as the convergence  $\kappa$  and the shear  $\gamma = \gamma_1 + i\gamma_2$ , directly related to the surface mass density. This is commonly described as a coordinate transformation

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad (4.1)$$

resulting in the lensed image  $I(x', y')$ .

Weak lensing magnification analyses (e.g., Hildebrandt et al. 2009, Van Waerbeke et al. 2010) require the intrinsic (distribution of) source sizes or magnitudes. In weak shear anal-

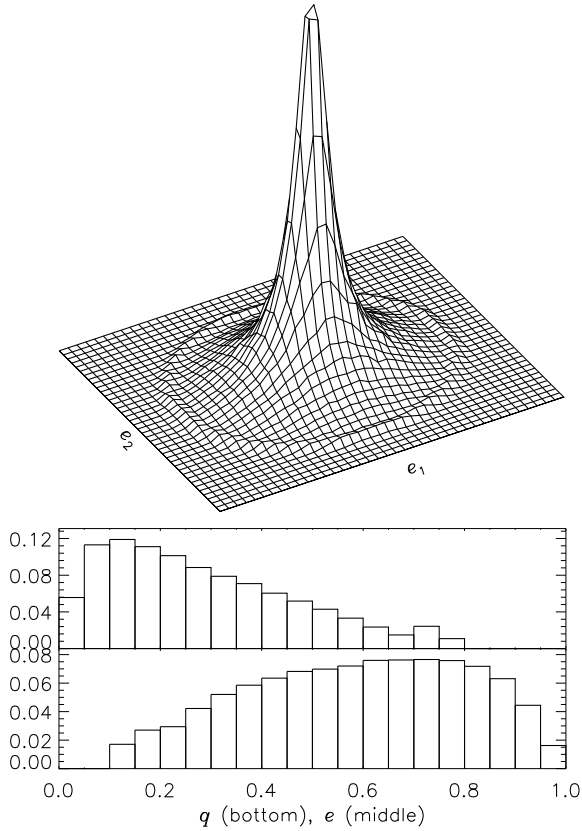


Figure 4.2: Ellipticity and axis ratio distributions of the CFHTLenS catalog. Top: a 2D histogram of ellipticities. We note that the ring-like feature at  $e \approx 0.8$  is due to noisy outliers forced to a maximum  $e$  by the shape measurement pipeline, but see also Figure 4.8. Middle: histogram of the absolute ellipticity  $|e|$ . Bottom: histogram of the ellipse axis ratio  $q$ .

yses, the focus lies on the net distortion or reduced shear  $g = g_1 + ig_2 \equiv (\gamma_1 + i\gamma_2)/(1 - \kappa)$ :

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = (1 - \kappa) \begin{pmatrix} 1 - g_1 & -g_2 \\ -g_2 & 1 + g_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad (4.2)$$

where the transformation is written as a multiplication of  $(1 - \kappa)$  (which leads to the magnification) and a traceless distortion matrix describing the alignment of lensed sources in the foreground potential.

The distortion effect of weak lensing shear on images of background galaxies depends on their intrinsic shape distribution. While galaxies often have complex morphologies, it is adequate to describe images by their quadrupole brightness moments or their ellipticities, and the respective response to weak shear distortions.

A common definition of the shape of an image with elliptical isophotes is the ellipticity  $e = e_1 + ie_2$ , defined as the reduced shear needed to create this image from an image with

circular isophotes (Bernstein & Jarvis 2002, Kuijken 2006). This gives an axis ratio  $q = \frac{b}{a}$  as

$$q = \frac{1 - |e|}{1 + |e|} \Leftrightarrow |e| = \frac{1 - q}{1 + q} = \frac{a - b}{a + b}, \quad (4.3)$$

and position angle  $\theta$  via

$$e = |e|(\cos 2\theta + i \sin 2\theta). \quad (4.4)$$

As an example, we compare a Gaussian ( $e_1, e_2$ ) distribution to the distribution observed in the CFHTLenS shape measurement catalog in Figures 4.1 and 4.2.

This complex notation gives a most straightforward formulation of the resulting ellipticity  $\tilde{e}$ , after transforming an image with ellipticity  $e$  by a distortion  $g$ , by Seitz & Schneider (1997)

$$\tilde{e} = \frac{e + g}{1 + g^*e} \quad \text{for } |g| \leq 1, \quad (4.5)$$

with  $g^*$  the complex conjugate of  $g$ .

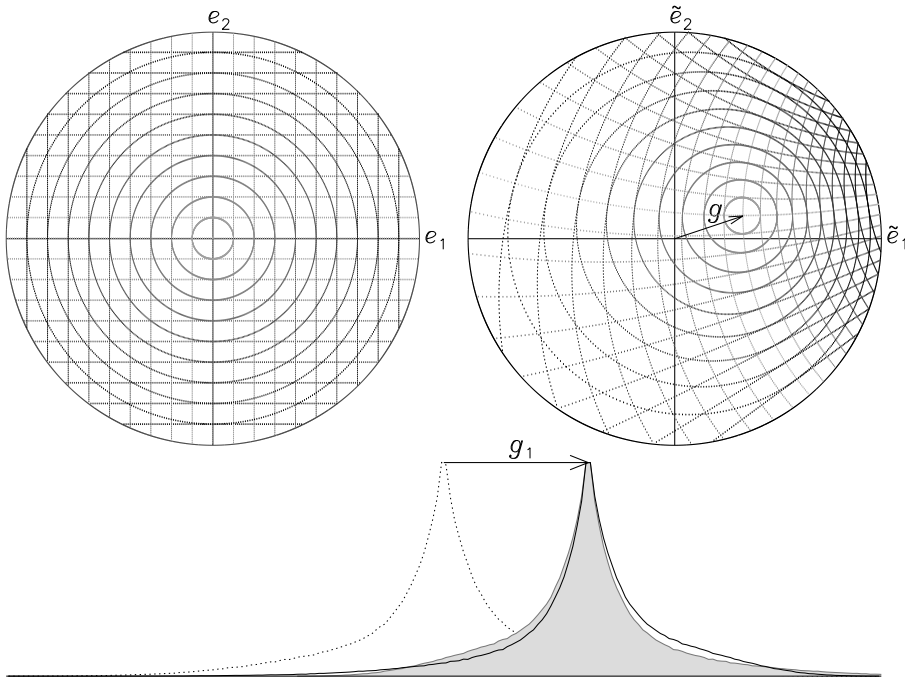


Figure 4.3: Top: the non-linear mapping of ellipticities (with  $|e| \leq 1$ ) by an exaggerated gravitational shear of  $g = 0.33 + 0.11i$ . Bottom: the asymmetry introduced in the ellipticity distribution, highlighted for the  $e_1$ -component.

The non-linear effect of gravitational shear on the ellipticity parameters is shown in the top panel of Figure 4.3. Through statistical estimation, we can attempt to infer from an ensemble of galaxy shapes the underlying shear, if we assume the intrinsic ellipticity distribution  $P(e)$  to be centered around zero ellipticity. In other words, one assumes no preferred direction on the sky.

This non-linear response to weak shear distortions gives rise to the asymmetry in the observed ellipticity distribution, as shown in the bottom panel of Figure 4.3. The shifted central peak of the distribution is unaffected by this non-linearity and therefore a direct tracer of  $g$ .

The canonical approach is a weighted mean  $\mu$ , where weighting schemes attempt to minimize systematic effects from noise, size and brightness. As observed by Seitz & Schneider (1997), the expectation value  $\langle \bar{e} \rangle$  does not depend on  $P(e)$  in the absence of noise. The mean of an ensemble of measured ellipticities is then an asymptotically unbiased estimator for the underlying shear  $g$ .

In the presence of noise, however, these estimations suffer from unavoidable biases in the estimated shear (Melchior & Viola 2012, Kacprzak et al. 2012). Furthermore, the variance of an estimator such as the mean, or more generally, the scale of the estimator distribution, does depend on the intrinsic ellipticity distribution  $P(e)$ . Informally put, the smaller the estimator variance, the more ‘trustworthy’ the estimates and the more efficient the estimator. A more efficient estimator reduces the uncertainties in and therefore the error bars or confidence intervals of parameter estimates.

The smearing of the sheared distribution by noise affects central value estimations, but the peak location itself is still an unbiased tracer of the shear.

## 4.3 Statistical framework

In this section, we discuss various estimators, after reflecting upon estimator properties, such as bias, efficiency, and robustness, and their interpretation. We then propose ways to apply this to fitting individual Fourier modes to a shear field.

### 4.3.1 Bias, efficiency and robustness

We will use the term bias, or  $\Delta e$ , when referring to the difference between the central value of an estimator, such as the expected value or mean  $\langle \hat{e} \rangle$ , and the population parameter  $e$ . We will use the term residuals, or  $r_i = e_i - \hat{e}$ , when talking about the differences between one sample estimate and the elements of that sample, that is, the individual measurements  $e_i = e_{i,1} + ie_{i,2}$ .

We note that we write  $r_i = e_i - \hat{e}$  for simplicity throughout this paper, but we employ Equation 4.5 to calculate the residuals, unless specifically noted otherwise. The absolute residual ellipticity of a single measurement with respect to the sample estimate is then the norm  $|r_i|$ .

The difference  $\Delta e = \langle \hat{e} \rangle - e$ , commonly referred to as simply the bias of the estimator, is formally called the mean-bias  $\mu_{\Delta e}$ . An estimator is then called asymptotically mean-unbiased, if for an increasing number of estimations  $\hat{e}$ , the mean estimate  $\mu_{\hat{e}}$  converges toward the parameter value of the underlying population. This is commonly simply referred to as unbiased. Here we have changed notation from  $\langle \hat{e} \rangle$  to  $\mu_{\hat{e}}$ , to emphasize the method of determining the central value of a set of estimates.

We do this, because there are other possible definitions of unbiasedness, such as median-unbiasedness, in which case the median estimate  $M(\hat{e})$  converges toward the true parameter value. By the central limit theorem, it is often appropriate to assume an asymptotically normal distribution of the estimator  $\hat{e}$  (not to be confused with the distribution  $P(e)$  of the population



parameter  $e$ ), when the number of estimations increases. This validates the general use of mean-unbiasedness. In practice, sample sizes needed for convergence toward a normal estimator distribution can be very large and one should take care when assuming asymptotic normality when making statistical inferences from a few measurements.

The efficiency of an estimator can be defined in terms of its variance. For unbiased estimators, this variance is bounded from below by the Cramér-Rao lower bound (Rao 1945, Cramér 1946), which in short means that there is an absolute maximum efficiency that can be obtained. For some distributions, such as the Gaussian distribution, this limit can be calculated analytically<sup>8</sup>. In other cases, it is useful to define a relative efficiency

$$\eta_e = \frac{\sigma_0^2}{\sigma_e^2}, \quad (4.6)$$

where  $\sigma_0^2$  is the variance of a comparison estimator, such as the mean. Then, if for example  $\eta_e > 1$ , the estimator has a lower variance than the mean and is therefore more efficient in finding the central value of the population parameter distribution. An estimator that achieves the Cramér-Rao lower bound for all possible parameter values is for this reason also known as a minimum variance estimator.

Again, if the assumption of asymptotic normality is not appropriate, another definition of the scale of distribution of the estimator can be used instead of the variance, such as the median absolute deviation (MAD). In such cases, care should be taken with the coverage of that scale, which is simply the percentage of estimates with lower residuals than the scale. In case of a Gaussian distribution, the standard deviation has a coverage of 68.3%. The MAD has, by definition, a coverage of 50%.

To avoid comparing apples with oranges, we will use chosen percentiles as scale, so the coverage is defined. For instance, we define the 68.3% scale  $s_{68.3}$  as the residual value for which 68.3% of the estimates has an equal or lower residual. In case of asymptotic normality,  $s_{68.3}^2$  will converge to the same value as the estimator variance.

We note that we can do this, since in our simulations the true population parameter value  $e$  is known<sup>9</sup>. In general, the coverage of a definition of scale is not known, confusing the interpretation of any relative efficiency.

In conclusion, we define the efficiency of an estimator  $\hat{e}$ , relative to the mean  $\mu_e$ , at a certain percentile coverage  $p$ , as

$$\eta_{\hat{e};p} = \frac{s_{\mu_e;p}^2}{s_{\hat{e};p}^2}. \quad (4.7)$$

Finally, we label an estimator  $\hat{e}$  as robust (in a qualitative manner), when  $\hat{e}$  retains low or zero bias and high efficiency in a wide range of possible distributions. A robust estimator is desirable, since it makes the choice of estimator for a parameter with unknown distribution more objective. As an example, the mean is optimally efficient in case of a Gaussian parameter distribution, but since the mean has low resistance against departures from Gaussianity (such as outliers), it is not the most robust.

---

<sup>8</sup>We will omit a more detailed discussion, since it's applicable only to certain distributions and not (directly) relevant to this discussion.

<sup>9</sup>More accurately, the underlying shear  $g$  is known.

Since we work with relative efficiencies, a conclusive statement about robustness is not straightforward. We will therefore use robustness to indicate that an estimator is equally or more efficient than the Gaussian estimator in most or all cases.

### 4.3.2 Estimators

We have explored various alternatives for well known estimators, which are optimal under Gaussian assumptions, like the mean and variance. By definition, the mean  $\hat{e}_\mu$ , or  $\mu_e$ , minimizes the variance of the residuals, which makes it a least squares estimator.

In general, optimization estimators are solutions  $\hat{e}$  that minimize a loss function

$$S_{\hat{e}} = \sum_i \rho(e_i; \hat{e}), \quad (4.8)$$

such as  $\rho = r_i^2 = (e_i - \hat{e})^2$  for the mean.

For this paper, we considered two other optimization estimators, the least absolute deviations estimator (LAD) and the biweight (BI) estimator, and an ordering estimator, namely convex hull peeling (CHP). In section 4.3.3, we describe Fourier mode fitting (FM), using a LAD regression approach.

#### Least absolute deviations

LAD is an optimization approach where the loss function to be minimized is the sum of the absolute deviations, instead of the commonly used least squares minimization:

$$S_{\text{LAD}} = \sum_i |r_i|. \quad (4.9)$$

In the one dimensional case, this is the median. In more than one dimension, we talk about the marginal median, when in each dimension the median is taken independently, or the spatial median, when minimizing the sum of the distances of measurements to a point. In many practical cases<sup>10</sup>, the spatial median is unique, contrary to the marginal median, which can have multiple solutions. This is one of the reasons we used the spatial median throughout the rest of the paper.

Another reason is that  $e_1$  and  $e_2$  should not be seen as independent parameters of the shape. An ellipticity is defined by an absolute elongation  $|e|$  and a position angle  $\theta$ . The latter is defined within the context of a chosen frame of reference and therefore so are  $e_1$  and  $e_2$ . In other words, using the marginal median would introduce an artificial anisotropy, as can be seen in Figure 4.4.

In concreto, for a set of  $(e_{i,1}, e_{i,2})$  measurements, the mean  $\hat{e}_\mu$  as an estimator for the net reduced shear  $g_1 + ig_2$  minimizes the squared residuals

$$S_\mu = \sum_i (e_{i,1} - g_1)^2 + (e_{i,2} - g_2)^2. \quad (4.10)$$

A LAD estimate minimizes the absolute residuals,

$$S_{\text{LAD}} = \sum_i \sqrt{(e_{i,1} - g_1)^2 + (e_{i,2} - g_2)^2}, \quad (4.11)$$

---

<sup>10</sup>Formally speaking: when the norm is strictly convex.

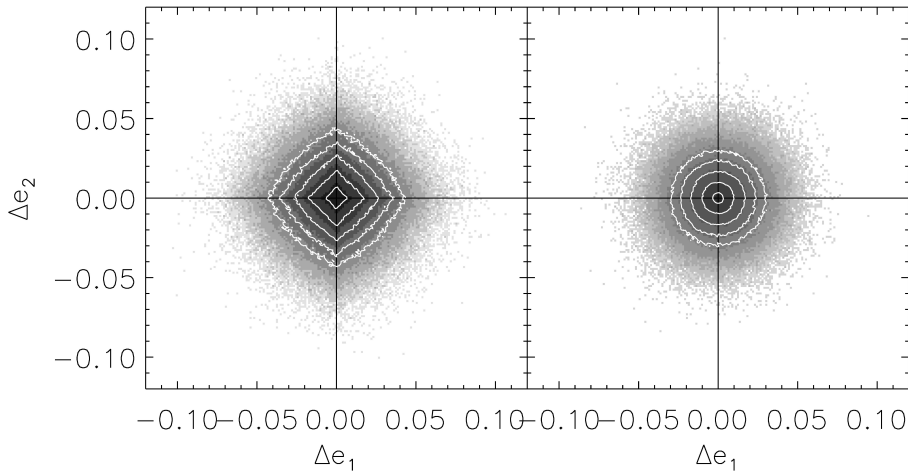


Figure 4.4: Comparison of the marginal median (left) to the spatial median, or LAD estimation (right). Plotted are the estimation biases  $\Delta e$  for  $10^6$  simulation runs, shown as a density in grayscale. Overplotted are arbitrary contours of increasing density (equal in both plots), to highlight the anisotropy in  $P(\Delta e)$ . Since  $e_1$  and  $e_2$  depend on the choice of reference frame, the marginal median introduces an artificial anisotropy. For the LAD estimations, the residual distances  $|r_i|$  do not depend on the choice of reference frame.

which reduces the effect of outliers on the estimate. In one dimension, the LAD estimate arises as the central value maximum likelihood estimator of the Laplace distribution, which has a central cusp and more slowly declining tails.

There is no general analytic solution for LAD optimization. LAD can however be formulated as a linear optimization problem for which several iterative methods exist (e.g., simplex-based methods, Barrodale & Roberts 1973). In practical weak shear analyses, convergence is generally rapid.

### The biweight

An alternative optimization approach is a bi-square weighted loss function (Beaton & Tukey 1974), called the biweight for short, given by

$$\nabla S_{\text{BI}} = \sum_i r_i \left(1 - \left(\frac{r_i}{k}\right)^2\right)^2 = 0 \quad \text{for } |r_i| < k, \quad (4.12)$$

where  $r_i = (e_i - \hat{e})$  are again the residuals and  $k$  is a tuning parameter, usually determined by (an estimate of) the scale of the measured distribution.

A robust choice for  $k$  is the median absolute deviation (MAD), setting  $k = c \cdot \text{MAD}$ , where  $c = 6.0$  is optimal for estimation of location for a broad range of distributions (Mosteller & Tukey 1977). A common approach is iteratively correcting an initial estimate  $M_0$  by the

normalized sum in Equation 4.12:

$$M_{n+1} = M_n + \frac{\sum_i r_{i,n} \left(1 - \left(\frac{r_{i,n}}{k}\right)^2\right)^2}{\sum_i \left(1 - \left(\frac{r_{i,n}}{k}\right)^2\right)^2}, \tag{4.13}$$

which can be interpreted as a normalized weighting of the residuals. In this case, the weight of a certain measurement increases toward the (current) central estimate, which makes this estimator a useful complement to the mean and LAD estimators.

In turn, a robust choice for  $M_0$  is the (spatial) median. Note that measurements with residuals  $|r_{i,n}| \geq k$  have effectively zero weight, although these points are not ‘clipped’ from the sample, since the residuals can change with each iteration. Convergence usually requires few iterations.

### Convex hull peeling

The convex hull of a set of points  $X$  in  $\mathbb{R}^n$  can be defined as the intersection of all convex sets in  $\mathbb{R}^n$  that contain  $X$ . Informally put, the convex hull is the smallest subset of points that ‘surrounds’ the rest of the set (see Figure 4.5).

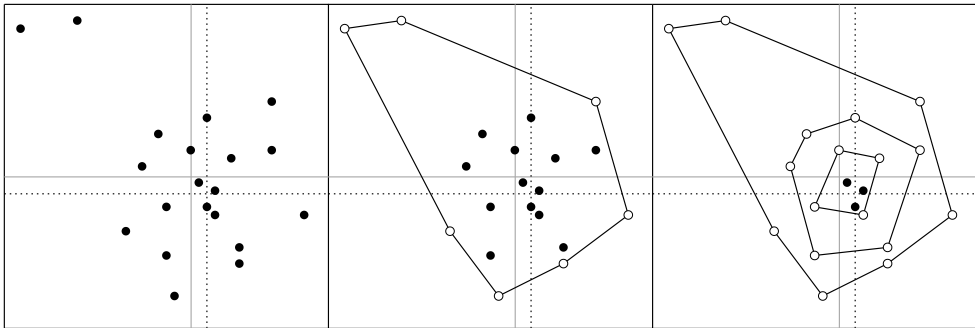


Figure 4.5: The method of CHP. The left panel shows a scatter plot with two outliers. The arithmetic mean is shown as a gray, solid line and the dotted line represents the mean without the two outliers. The middle panel shows the convex hull of the set of points, which is then removed from the set. The right panel shows the final result after repeating the process, until the final set of points is equal to its own convex hull.

There exist various algorithms for determining the convex hull (e.g., Preparata & Shamos 1985). For this paper, we used Delaunay triangulation based on the divide-and-conquer method (Lee & Schachter 1980).

In the process of CHP, the convex hull of a set of data points is determined and subsequently ‘peeled’ from the set, after which the process is repeated (Figure 4.5). When the remaining set of points is equal to its own convex hull, the final estimate is determined from these points, for example using the mean or LAD. This makes CHP an ordering approach, much like obtaining the familiar median for the one-dimensional case by sorting the data,

instead of optimization<sup>11</sup>. Among other aspects, it shares the resistance of the median against outliers.

In this paper, we are interested in the use of CHP in the two-dimensional case of  $(e_1, e_2)$  measurements in the complex plane, but CHP can be used in higher dimensions as well (see e.g., Lee 2007, for applications to SDSS quasar data).

### Weighting and collinearity

When using real data, a weighting of ellipticity measurements is necessary to avoid or mitigate effects, such as noise or intrinsic size and ellipticity, that would confuse or bias the estimation of the underlying shear. For LAD and biweight optimization, weighting schemes are readily introduced, analogous to the weighted mean. For CHP, we suggest a possible weighting scheme, analogous to the one-dimensional weighted median, as follows.

The convex hull comprises a set of points in the  $(e_{i,1}, e_{i,2})$ -plane, with  $w_i$  the associated weights, given by the measurement pipeline. The minimum weight on the convex hull is then subtracted from these weights, after which all points with updated weight  $w_i = 0$  are peeled from the sample. Note that this removes at least one point per iteration, but can lead to point-by-point peeling and large computation times. A solution with lesser precision but increased speed would be given by binning the weights in discrete steps.

We also note a possible collinearity problem of multiple ellipticity measurements with finite precision coinciding. In that case, triangulation has no solution. By combining these points into one measurement by combining the weights, this problem is resolved.

### 4.3.3 Fourier mode fitting

One can model a signal, in our case a varying ellipticity, over a one-dimensional range  $-L < x < L$ , writing that signal as a linear superposition of waves, or (Fourier) modes,  $A_n \cdot \cos(k_n x \pm \phi_n)$ , where  $A_n$  and  $\phi_n$  are the amplitude and phase of the signal mode respectively, and  $k_n \equiv \frac{n\pi}{L}$  are the wave numbers of the modes, showing the periodicity over the range  $2L$ .

It is useful to rewrite this model linearly in its coefficients  $a_n \cdot \cos(k_n x) + b_n \cdot \sin(k_n x)$ , where amplitude and the phase are now given by  $A_n^2 = a_n^2 + b_n^2$  and via  $\frac{b_n}{a_n} = \tan(\phi_n)$ .

This one-dimensional model is readily extended to two dimensions, by considering that each coefficient depends similarly on  $y$ . This gives us  $\alpha_{mn;\pm} = \cos(k_m x \pm l_n y)$  and  $\beta_{mn;\pm} = \cos(k_m x \pm l_n y)$ , or

$$\begin{aligned}
 e(x, y) = \sum_{m,n} & a_{mn} \cos(k_m x) \cos(l_n y) \\
 & + b_{mn} \cos(k_m x) \sin(l_n y) \\
 & + c_{mn} \sin(k_m x) \cos(l_n y) \\
 & + d_{mn} \sin(k_m x) \sin(l_n y) ,
 \end{aligned} \tag{4.14}$$

where the wave numbers  $k_m$  and  $l_n$  represent the spatial frequencies in the  $x$  and  $y$  directions, respectively. In two dimensions, we make a terminological distinction between a full Fourier mode, as given by Equation 4.14, and the individual waves comprising it. The amplitude of the fluctuations in ellipticity are now given for each mode in  $m, n$  by  $a_{mn}^2 + b_{mn}^2 + c_{mn}^2 + d_{mn}^2$ .

<sup>11</sup>Indeed, in one dimension, both approaches to the median are the same.

This linear model is fitted in a relatively straightforward manner to a sample of measured or simulated ellipticities. In the absence of noise and for a well-behaved field of view, each wave component of a Fourier mode is independent and can be fitted separately. We will discuss the effect of noise in Section 4.5.2.

### Applying statistics

To apply these statistics to a shear field consisting of discrete Fourier modes, which by construction is centered around  $e = 0$ , the ellipticity measurements should be properly weighted by the model of the Fourier mode under consideration. We considered the information carried by an ellipticity measurement, which is proportional to the value of the fitted model  $M$ , where  $M(x, y)$  can for instance be a single wave like  $\cos(k_m x) \cos(l_n y)$ , or a full mode.

Measurements close to the nodes of a wave carry the least information, whereas measurement close to extrema, or antinodes, carry the most amplitude information. We considered that each ellipticity measurement  $e_i$  theoretically infers an estimate of the amplitude  $A$ , where  $A \in \{a_{mn}, b_{mn}, c_{mn}, d_{mn}\}$  by  $\hat{A}_i = e_i \cdot M^{-1}$ . In the case of Gaussian variations around the model, that is, measurement error distribution, the information scales as the inverse variance of that distribution, and therefore as the square of the model:

$$\hat{A} = \frac{\sum M^2 \cdot \frac{e}{M}}{\sum M^2} = \frac{\sum M \cdot e}{\sum M^2}, \quad (4.15)$$

where we recover the well known analytic LSQ form. This can be seen as an inverse variance weighting based on the model-to-noise ratio. For different error distributions, one can allow a general scaling of the information with the model by  $M^n$ , and therefore

$$\hat{A} = \frac{\sum M^{n-1} \cdot e}{\sum M^n}. \quad (4.16)$$

For application with our proposed weighting scheme for CHP, it is instructive to view the multiplication by weights as shifting the data points, so the central data point(s) or CHP value matches the amplitude to be estimated. For this purpose, it is practical to write Equation 4.16 as

$$\hat{A} = \frac{\sum |M|^{n-1} \cdot \text{sgn}(M) \cdot e}{\sum |M|^{n-1}} \cdot \frac{\sum |M|^{n-1}}{\sum M^n}, \quad (4.17)$$

where the  $(e_1, e_2)$  data points are first shifted by  $\text{sgn}(M)$ , and then weighed by  $|M|$ , before the weighed estimate is normalized as usual. We show this in Figure 4.6, where we plot the  $(e_1, e_2)$  values, the same  $(e_1, e_2)$  points shifted by  $\text{sgn}(M)$ , with in this case  $M = \cos(k_m x) \cos(l_n y)$ , and then the associated distribution of the weights  $|M|$  over the complex ellipticity plane as normalized 2D histograms.

## 4.4 Simulations and data

For this paper, we tested various forms of  $P(e)$ , using samples of random ellipticities, assumed to be centered around zero, which we sheared by Eq. 4.5. We have used several approaches to obtaining these samples.

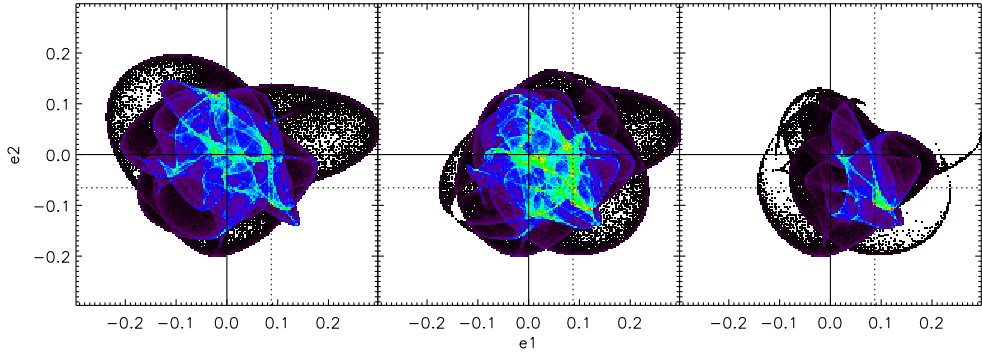


Figure 4.6: Ellipticity distribution of a superposition of 9 Fourier modes over the complex  $(e_1, e_2)$  plane, where we show how we recover a single  $(e_1, e_2)$  amplitude, indicated by the dotted lines. Left: the Fourier modes, centered around  $(0, 0)$  (normalized number counts). Middle: the  $(e_1, e_2)$  points shifted by  $\text{sgn}(M)$ , with  $M$  the model of the amplitude (normalized number counts). Right: the resulting distribution of weights over  $(e_1, e_2)$ , showing a shift toward the amplitude under consideration.

Firstly, we simulated a uniform  $q$  distribution, which seems to fit real data adequately (e.g., Lambas et al. 1992, Rodríguez & Padilla 2013), without assuming any physical mechanism that would explain this distribution.

Secondly, we modeled background galaxies as randomly orientated triaxial ellipsoids, and derived the projected ellipticities following Stark (1977), using axis ratio distributions fitted to observed ellipticity distributions (Lambas et al. 1992).

In both cases, we compared our results to samples with added Gaussian noise, using real data shape measurement error distributions to simulate the effect of noise.

Thirdly, we sampled real data, using shape measurement catalogs from weak lensing observations.

Finally, we compared these various ellipticity distributions and the results from each estimator to results in case when  $P(e)$  follows a Gaussian distribution. We examined the behavior of bias and efficiency of each estimator under the effect of noise, the input shear and the sample size.

#### 4.4.1 Simulated ellipticity distributions

##### Uniform samples

We produced random samples with a uniform  $q$ -distribution, as an ideal version of the observed distribution of spiral galaxies in for example Lambas et al. (1992), Rodríguez & Padilla (2013), henceforth referred to as a uniform sample. We used an axis ratio cut-off of  $q \approx 0.2$  to account for a finite galaxy thickness, following Lambas et al. (1992), which gives rise to standard deviations in each ellipticity component of  $\sigma_e \approx 0.25$ , comparable to the samples drawn from data.

The resulting axis ratio and ellipticity distributions are shown in Figure 4.7

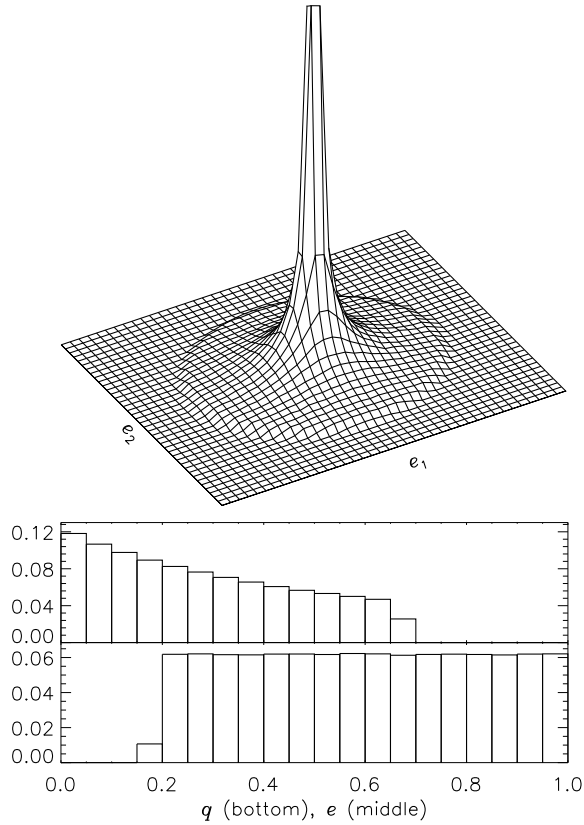


Figure 4.7: Ellipticity distributions for a uniform axis ratio distribution. Top: a 2D histogram of ellipticities. Middle: histogram of the absolute ellipticity  $|e|$ . Bottom: histogram of the ellipse axis ratio  $q$ . A cut-off near  $q \approx 0.2$  is suggested by observations and produces standard deviations in each ellipticity component of  $\sigma_e \approx 0.25$ , comparable to most survey shape measurement catalogs.

### Projected ellipsoids

A triaxial ellipsoid with axes  $\tilde{a} \geq \tilde{b} \geq \tilde{c} \geq 0$  can be described by

$$(cx)^2 + \left(\frac{cy}{b}\right)^2 + (z)^2 = \text{constant}, \tag{4.18}$$

with  $b = \tilde{b}/\tilde{a}$  and  $c = \tilde{c}/\tilde{a}$ . As given by Stark (1977), such an ellipsoid is seen as an ellipse in projection, given by

$$(j/f)x'^2 + 2(k/f)x'y' + (l/f)y'^2 = \text{constant}, \tag{4.19}$$



where  $(x', y')$  are the coordinates in the projection plane and

$$f \equiv c^2 \sin^2 \theta \sin^2 \varphi + (c/b)^2 \sin^2 \theta \cos^2 \varphi + \cos^2 \theta, \quad (4.20a)$$

$$j \equiv c^2 (c/b)^2 \sin^2 \theta + c^2 \cos^2 \varphi \cos^2 \theta + (c/b)^2 \sin^2 \varphi \cos^2 \theta, \quad (4.20b)$$

$$k \equiv ((c/b)^2 - c^2) \sin \varphi \cos \varphi \cos \theta, \quad (4.20c)$$

$$l \equiv c^2 \sin^2 \varphi + (c/b)^2 \cos^2 \varphi, \quad (4.20d)$$

with  $\varphi$  and  $\theta$  the first two orientation angles of the ellipsoid.

For simulations of projected ellipsoids, we assumed Gaussian distributions for  $b$  and  $c$ , following Lambas et al. (1992). For elliptical galaxies, we used  $b = 0.95$  and  $c = 0.55$  with standard deviations  $\sigma_b = 0.35$  and  $\sigma_c = 0.2$ . For disk galaxies, we used  $b = 1.00$  and  $c = 0.25$  with standard deviations  $\sigma_b = 0.13$  and  $\sigma_c = 0.12$ .

The axis ratio was then recovered via

$$q = \sqrt{\frac{j + l - \sqrt{(j - l)^2 + 4k^2}}{j + l + \sqrt{(j - l)^2 + 4k^2}}}, \quad (4.21)$$

and the ellipticity through Equation 4.3. The orientation angles of the ellipsoids were randomly distributed. The resulting axis ratio and ellipticity distributions are shown in Figure 4.8.

We will refer to these simulated samples as disk and elliptical samples. We also used combined samples with a disk to elliptical ratio derived from the CFHTLenS catalog. (See Table 4.1.)

#### 4.4.2 Data: CFHTLenS

We used data from Canada-France-Hawaii Telescope Lensing Survey (CFHTLenS, Heymans et al. 2012b). The CFHTLenS survey analysis combined weak lensing data processing with THELI (Erben et al. 2005, 2009, 2013), shear measurement with *lensfit* (Miller et al. 2007, 2013, Kitching et al. 2008), and Bayesian photometric redshift measurement (BPZ, Benítez 2000, Coe et al. 2006) with PSF-matched photometry (Hildebrandt et al. 2012). A full systematic error analysis of the shear measurements in combination with the photometric redshifts is presented in Heymans et al. (2012b), with additional error analyses of the photometric redshift measurements presented in Benjamin et al. (2013).

For our analyses, we selected 4.2 million objects that are well determined and resolved (*lensfit* fitclass = 0, non-zero *lensfit* weight, star\_flag = 0, CLASS\_STAR ≤ 0.5). We excluded objects that lie within a mask, with the exception of large, conservative masks around relatively faint stars and stellar haloes (MASK ≤ 1, see Erben et al. 2013).

The CFHTLenS shape catalog is not an exact representation of the ellipticity distribution of the observed galaxy population, as it includes measurement noise present in any real data set. Selecting sources on *lensfit* weight  $w$  or signal-to-noise ratio  $\nu_{\text{SN}}$  could on the other hand introduce selection biases in the galaxy population we wanted to study. We decided to use two sets of sources: the complete set, described above, to optimally sample the complete source population, and a conservative subset with  $w \geq 15$  and  $\nu_{\text{SN}} \geq 20$ , to reduce the uncertainty in observed ellipticity, at the possible cost of a bias in the selection.

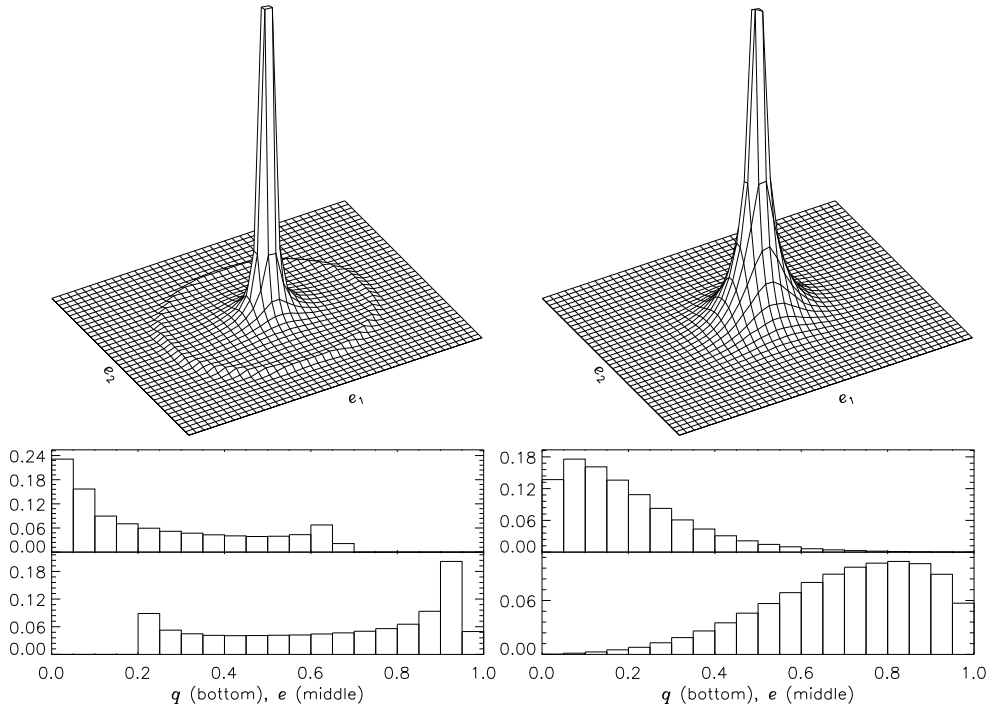


Figure 4.8: Ellipticity and axis ratio distributions for distribution of projected ellipsoids. Left: disk galaxies. Right: elliptical galaxies. Top: a 2D histogram of ellipticities. Note that the ring-like feature in the left panel is the result of a finite disk thickness. Middle: histogram of the absolute ellipticity  $|e|$ . Bottom: histogram of the ellipse axis ratio  $q$ .

For both sets, we split these sources by BPZ spectral type into red ( $T_{\text{BPZ}} < 1.5$ ) and blue ( $1.5 < T_{\text{BPZ}} < 3.95$ ) galaxies, with a further division between Sbc ( $1.5 < T_{\text{BPZ}} < 2.5$ ) and Scd ( $2.5 < T_{\text{BPZ}} < 3.95$ ). We found that our conservative selection reduced the number of galaxies to roughly 25%, almost independent of spectral bin for  $T_{\text{BPZ}} < 3.1$ . For higher spectral types, the subset decreased linearly to roughly 10% for the highest spectral bin, which was an indication that our selection did indeed introduce a modest sample bias.

Table 4.1 gives an overview of the selected CFHTLenS data, while Figure 4.9 shows the respective distributions.

We drew random subsets from the selected CFHTLenS ellipticities, which we then sheared by Eq. 4.5. This introduced the implicit assumption that, after the bias corrections described in Heymans et al. (2012b) and Miller et al. (2013), the central ellipticity was zero, and these random subsets were approximately drawn from an unsheared, noise-free background galaxy population.

#### 4.4.3 Simulated noise

In any realistic shape measurement catalog, ellipticities not only have shape noise due to a finite intrinsic distribution, but suffer from measurement uncertainties as well. For this

Table 4.1: Overview of the CFHTLenS data used. Column 1 gives the division between BPZ spectral type (red:  $T_{\text{BPZ}} < 1.5$ , blue:  $1.5 < T_{\text{BPZ}} < 3.95$ , Sbc:  $1.5 < T_{\text{BPZ}} < 2.5$ , Scd:  $2.5 < T_{\text{BPZ}} < 3.95$ ). Column 2 gives the number  $N$  of objects selected. Column 3 gives the 1D Gaussian ellipticity standard deviation  $\sigma_e$ , using both ellipticity components after bias correction. In parentheses, we give  $N$  and  $\sigma_e$  for sources with  $w \geq 15$  and  $v_{\text{SN}} \geq 20$ .

Color	$N$		$\sigma_e$	
All	4216334	(912828)	0.286	(0.242)
Red	553633	(151939)	0.267	(0.242)
Blue	3662701	(760889)	0.289	(0.242)
Sbc	870295	(219929)	0.294	(0.262)
Scd	2792406	(540960)	0.288	(0.232)

reason, we wanted to study the effect of noise or our simulated, noiseless ellipticity samples.

Measurement uncertainties depend primarily on pixel noise and therefore vary with image size and brightness. This means that errors on the ellipticities are not drawn from a single distribution. To mimic the effect of a skewed composite error distribution for our simulated samples, we randomly sampled the CFHTLenS weight  $w$ .

Miller et al. (2013) calculated an approximately inverse-variance weight using the width of the ellipticity likelihood surface by

$$w = \left[ \frac{\sigma_e^2 e_{\text{max}}^2}{e_{\text{max}}^2 + 2\sigma_e^2} + \sigma_{\text{pop}}^2 \right]^{-1}, \quad (4.22)$$

where  $\sigma_e^2$  is the variance in ellipticity of the likelihood surface,  $\sigma_{\text{pop}}^2$  is the ellipticity variance of the galaxy population, and  $e_{\text{max}}$  is a maximum ellipticity, to reflect a finite edge-on disk thickness.

Using  $e_{\text{max}} = 0.804$  from Miller et al. (2013) and refining  $\sigma_{\text{pop}}^2 \approx 0.242$  using the CFHTLenS catalog itself<sup>12</sup>, we obtained a distribution in ellipticity variance  $\sigma_e^2$  for each  $w$ . From this, we produced noise by assuming a Gaussian distribution with the ellipticity as mean and  $\sigma_e^2$  as variance.

## Estimation of errors

To assess errors on bias and efficiency from our simulations, we simply divided our simulations randomly in smaller subsets and determine the statistical variations, assuming  $t$ -distributions. While this approach may seem to lack finesse compared to a full bootstrap, the significance of our results is high enough for a proof of concept.

<sup>12</sup>Miller et al. (2013) cite  $\sigma_{\text{pop}}^2 = 0.255$  as prior, but this would lead to a negative  $\sigma_e^2$  for the maximum weight in the CFHTLenS catalog.

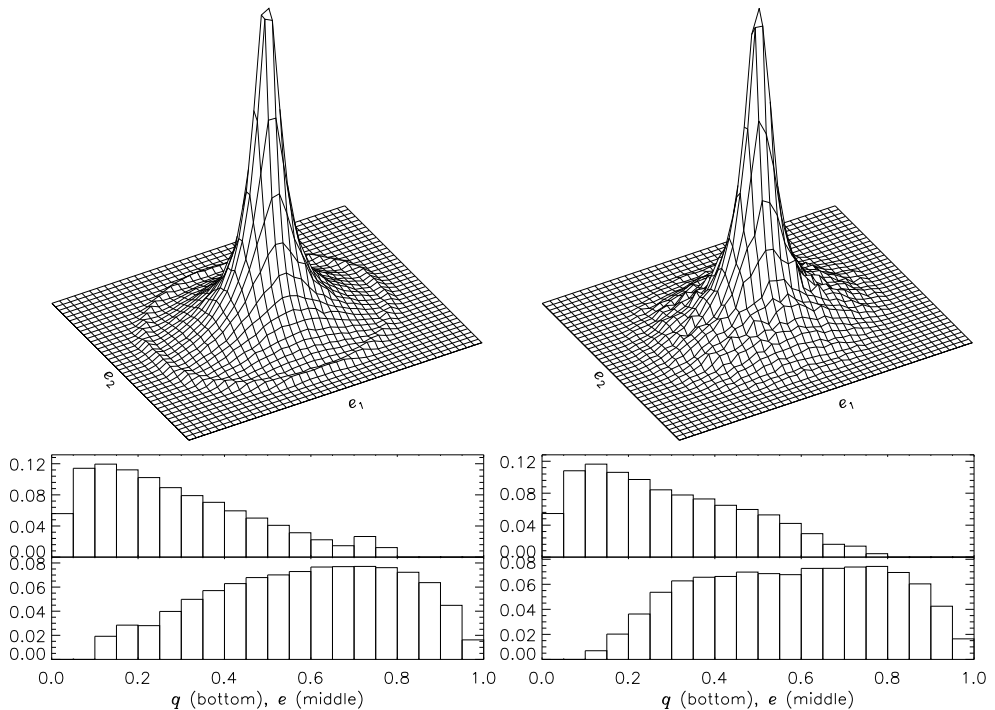


Figure 4.9: Ellipticity and axis ratio distributions for CFHTLenS blue (left), red (right). Top: a 2D histogram of ellipticities. Note that the ring-like feature at  $e \approx 0.8$  is due to noisy outliers forced to a maximum  $e$  by the shape measurement pipeline, but see also Figure 4.8. Middle: histogram of the absolute ellipticity  $|e|$ . Bottom: histogram of the ellipse axis ratio  $q$ .

## 4.5 Results

### 4.5.1 Central value estimation

For each sample type, we produced  $10^4$  random samples of 100 ellipticities, which we distorted by an absolute reduced shear of  $g = 0.2$ , and determined relative efficiencies and possible biases. We then assessed the effect of varying the shear and the sample size.

#### Asymmetry and bias

Ideally, an estimator should be unbiased in the absence of noise. For the mean, this is the case (Seitz & Schneider 1997), but since the effect of shear on intrinsic ellipticities is non-linear, the resulting, observed ellipticity distribution  $P(\tilde{e})$  is asymmetric, or skewed, which can lead to mean-biases for various estimators.

In Figure 4.10, we show this effect on the CHP estimator for  $g = 0.3$  in two directions. The distribution of the CHP estimator is clearly skewed, as shown by the convex hulls plotted, when the coverage within the current hull is equal to approximately<sup>13</sup> 38.3%, 68.3%, 86.6%,

<sup>13</sup>CHP is a discrete and not a continuous process, but this effect is negligible for  $10^4$  estimates.

and 95.4%. We note that this leads to a mean-biasedness, according to definition, but the center of the estimator distribution  $P(\hat{e}_{\text{CHP}})$  seems significantly less biased. In other words, the CHP estimator seems ‘CHP-unbiased’.

A solution to this skewness in the estimator distribution, in the absence of noise, is iteratively improving estimates by correcting the observed ellipticities  $P(\bar{e})$  by the estimated shear, using Equation 4.5, and then determining the updated residuals. We call this process of iteratively correcting the sample by the current estimate ‘de-shearing’ (or ‘de- $g$ ’). Figure 4.10 shows how this symmetrized the estimator distribution  $P(\hat{e})$ , and slightly improved the efficiency as well (see section 4.5.1). The latter seemed to be the case even for the mean  $\hat{e}_\mu$  as estimator, but the difference was not statistically significant.

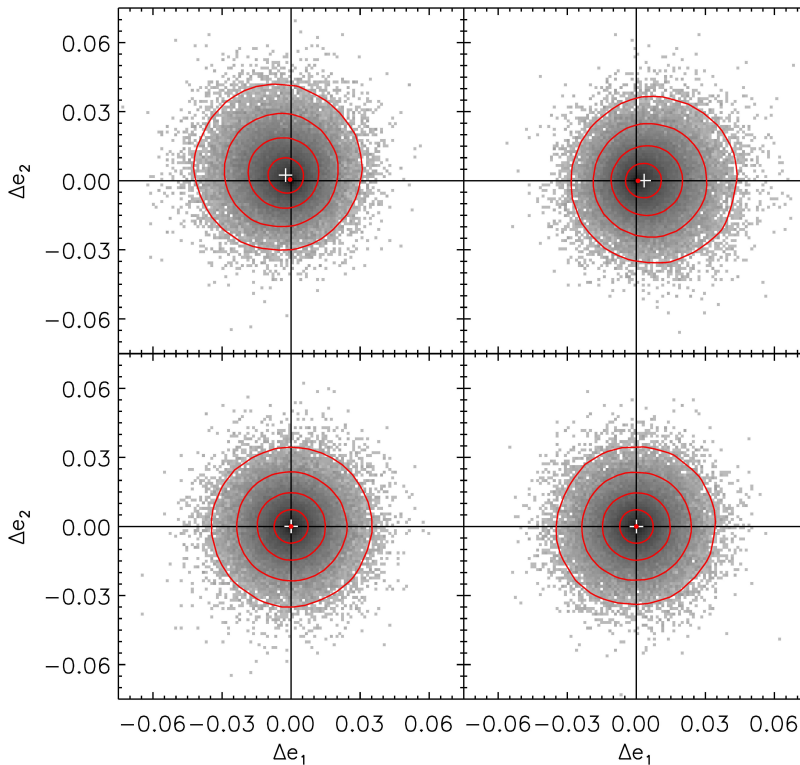


Figure 4.10: The skewed  $\hat{e}_{\text{CHP}}$  distribution as an example of the effect of asymmetry in a sheared ellipticity distribution. Plotted are the estimation biases  $\Delta e$  for  $10^5$  simulation runs, shown as a density in grayscale. Over-plotted are the convex hulls at approximately 38.3%, 68.3%, 86.6%, and 95.4% coverage. The mean of the distribution is shown as a white plus. Top: estimation biases  $\Delta e_{\text{CHP}}$  for samples with an underlying shear of  $g = -0.21 + 0.21i$  (left) and  $g = 0.3$  (right). Note that these estimator distributions are effectively mean-biased, because they are skewed, but still centered around  $\Delta e = 0$ , as indicated by the CHP estimation of the distributions. Bottom:  $\Delta e_{\text{CHP}}$  for the same samples, after iteratively de-shearing the samples until the final CHP estimate vanishes. These iterations remove asymptotic mean-bias and increase efficiency.

In presence of noise, the mean is a biased estimator (Melchior & Viola 2012). Given that

in reality systematic noise is always present, a form of bias is unavoidable, since the noise distribution is different<sup>14</sup> from the (skewed) ellipticity distribution (See Figure 4.3). This means that our method of de-shearing would introduce a noise bias for precisely the same reason, since we would not properly correct the asymmetry in the distribution.

We compared the results for simulated projected ellipsoids with and without simulated noise in Figure 4.11 to assess the effect. In the appendix, we quantified the observed multiplicative bias in the form

$$e_{\text{fit}} = (1 + m)e_{\text{in}}, \quad (4.23)$$

where  $e$  stands for  $e_{1,2}$ , and summarize the results in Table 4.2.

Without de-shearing, only the mean is a mean-unbiased estimator. We noted that all estimation methods could be made mean-unbiased in the noise-free case, when including de-shearing, but showed a mean-bias in the presence of noise, as expected. For the biweight estimator  $\hat{e}_{\text{BI}}$ , this was (within statistical significance) the same bias as for the mean. For the LAD and CHP estimators  $\hat{e}_{\text{LAD}}$  and  $\hat{e}_{\text{CHP}}$ , the biases were significantly reduced, up to  $\sim 30\%$ , to below percent level for realistic weak shear.

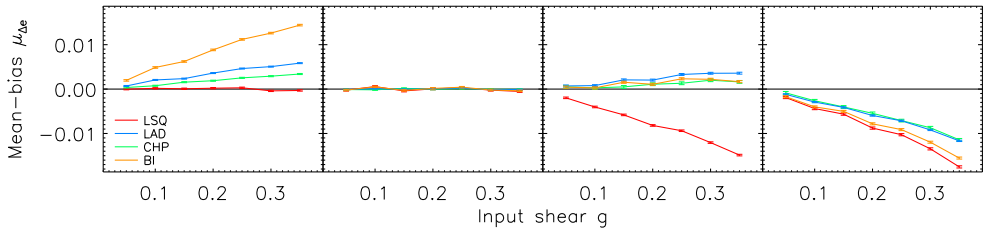


Figure 4.11: Estimator mean-bias as a function of input shear for realistic combinations of simulated disk and elliptical samples, using projected ellipsoids. From left to right: all estimators without noise, without noise after iteratively de-shearing the samples, all estimators with noise, and with noise and after iteratively de-shearing the samples. Color coding:  $\hat{e}_{\mu}$  (red),  $\hat{e}_{\text{LAD}}$  (blue),  $\hat{e}_{\text{CHP}}$  (green) and  $\hat{e}_{\text{BI}}$  (yellow).

This decrease in bias can be explained by realizing that the observed, sheared ellipticity distribution is skewed, but the location of the central peak of intrinsically round background sources is still an unbiased estimator of the underlying shear (which can be deduced from Equation 4.5 and Figure 4.3). It is the bias in determining the location of this peak that introduces the bias in the shear estimate. Likewise, the effect of noise changes the observed ellipticity distribution, but does not affect the location of that peak. Estimators that are more sensitive to a central cusp or peak in the distribution and less to high ellipticities in the tail, such as  $\hat{e}_{\text{LAD}}$  and  $\hat{e}_{\text{CHP}}$ , will therefore introduce a lower mean-bias.

We compared these results to the mean-bias in the upper panels of Figure 4.10 and the observation that the central peak of the estimator distribution is in fact located at  $\Delta e \approx 0$ . We found that the mean-bias arose due to the asymmetry in the estimator distribution and the CHP-bias vanished, unaffected by noise.

<sup>14</sup>Intrinsically, the effect of noise is symmetric, but the effect on a sample of sheared ellipticities depends on the shape measurement pipeline, as noted in Melchior & Viola (2012).

### Estimator efficiencies

In Tables 4.3, 4.4, and 4.5 in the appendix, we summarized the full results for the relative efficiencies of each estimator. We applied de-shearing and note that this improves the efficiencies marginally at a similar marginal cost to the bias. We determined relative efficiencies for coverages of 25%, 50%, and 75%, corresponding to the MAD and the first and third quartiles, and 38.3%, 68.3%, 86.6%, and 95.4%, which would correspond to steps of  $0.5\sigma$  in case of a Gaussian distribution with variance  $\sigma^2$ .

In Figure 4.12, we plot these results for a few distributions, namely Gaussian, uniform  $q$ , a combination of disk and elliptical projections and the conservative CFHTLenS catalog samples. We also plot the results for the samples with added noise and the full CFHTLenS samples.

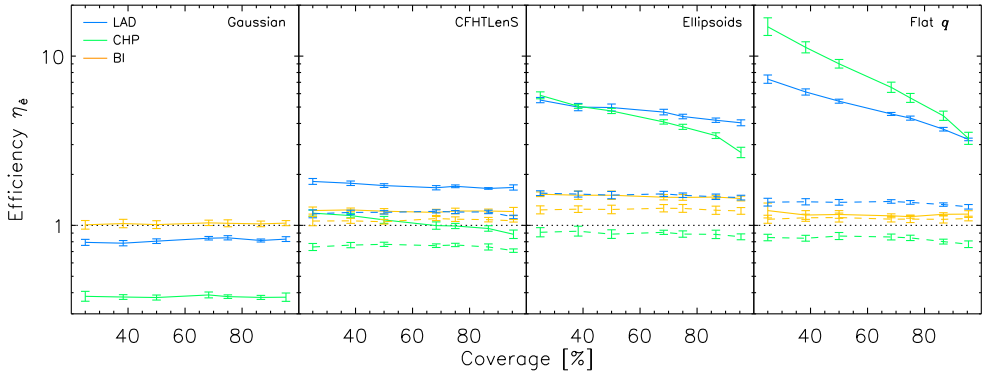


Figure 4.12: Relative efficiencies each estimator plotted at different coverages. From left to right: Relative efficiencies in case of a Gaussian  $P(e)$  distribution, the CFHTLenS catalog  $P(e)$  distribution, a combination of disk and elliptical distributions using projected ellipsoids, and a uniform  $q$  distribution. Color coding: relative efficiencies for  $\hat{e}_{\text{LAD}}$  (blue),  $\hat{e}_{\text{BI}}$  (yellow), and  $\hat{e}_{\text{CHP}}$  (green). Solid lines: simulated samples without noise or using the CFHTLenS conservative subset. Dashed lines: including noise or using the complete CFHTLenS set.

Not all estimators reached asymptotic normality. Especially CHP converged slower toward normality in the tails of the distribution, that is, at higher coverage. For LAD, this is noticeable mostly for the uniform  $q$  distribution.

The biweight is the most robust, as its relative efficiency doesn't vary much across distributions. The biweight relative efficiency is however quite low, which means that this estimator offers little improvement. Even when  $P(e)$  follows a Gaussian distribution,  $\eta_{\text{BI}}$  is not significantly better or worse than the traditional mean.

Our results show that estimator efficiency is independent of input shear. This is the case, when we define the individual estimate biases similarly to the residuals, as noted in Section 4.3.1, that is, not as the difference  $\hat{e} - g$ , but as the extra shear needed over the input shear  $g$  to reach this difference, as determined by Equation 4.5:

$$\Delta e = \frac{\hat{e} - g}{1 - g^* \hat{e}}, \quad (4.24)$$

with  $g^*$  the complex conjugate of the input shear  $g$  of the simulations. Using that definition,

this independence is demonstrated Figure 4.13.

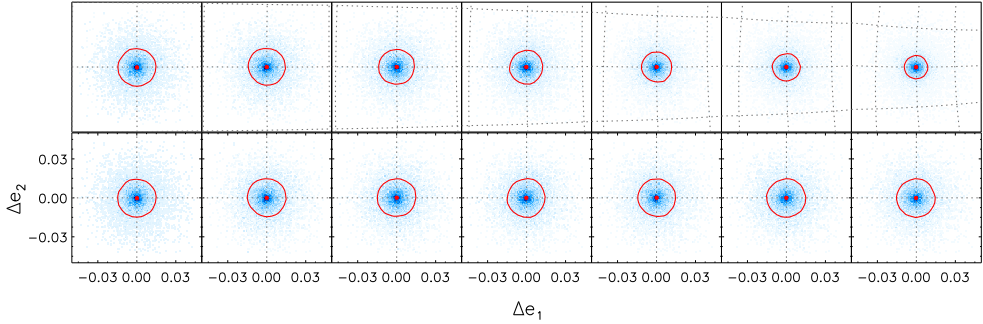


Figure 4.13: Efficiency for an arbitrary estimator and sample type versus input shear, ranging from  $g = 0.05$  (left) to  $g = 0.35$ . Upper: simple difference  $\hat{e} - g$  between estimates and input shear, with  $s_{68.3}$  over-plotted. Lower:  $\Delta e$ , as defined by Equation 4.24, with  $s_{68.3}$  over-plotted.

As an aside: since the mean of the CHP estimator is displaced from the center, this necessarily increases the distribution scale. A more proper way to compare the scale with symmetric distributions would be comparing the surface within the convex hull at a certain coverage, as  $s^2$  is a measure of the (circular) surface around  $\hat{e}$  inside that scale. In this sense, efficiency is a figure of merit. We have not done so in this paper, which means the  $\eta_{\text{CHP}}$  are slightly underestimated, but not significantly.

In Figure 4.14, we show the results for different samples sizes. In Table 4.6 in the appendix, we summarize the quantitative results. In the limit of very small sample sizes, the difference between the various estimators is expected to vanish. We note that a potential improvement over the mean estimator remains even for a sample size of  $N = 10$ .

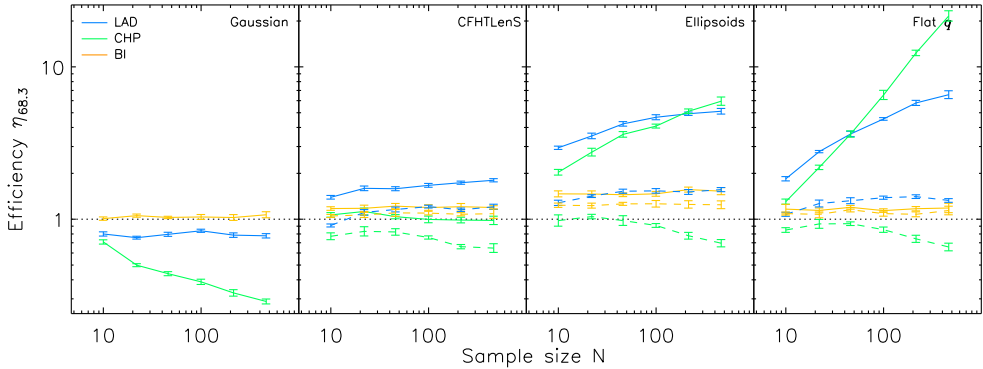


Figure 4.14: Relative efficiencies  $\eta_{68.3}$  plotted against sample sizes. From left to right: Relative efficiencies in case of a Gaussian  $P(e)$  distribution, the CFHTLenS catalog  $P(e)$  distribution, a combination of disk and elliptical distributions using projected ellipsoids, and a uniform  $q$  distribution. Color coding: relative efficiencies for  $\hat{e}_{\text{LAD}}$  (blue),  $\hat{e}_{\text{BI}}$  (yellow), and  $\hat{e}_{\text{CHP}}$  (green). Solid lines: simulated samples without noise or using the CFHTLenS conservative subset. Dashed lines: including noise or using the complete CFHTLenS set.



## 4.5.2 Fourier mode fitting

For samples of a combination of disk and elliptical distributions using projected ellipsoids, we produced  $10^3$  random square fields with  $10^3$  simulated ellipticities. For comparison, the average number of selected sources in a CFHTLenS field is roughly  $2.5 \cdot 10^4$ , ranging from 9525 to 37767, or  $5.3 \cdot 10^3$ , ranging from 2111 to 9525 for the more conservative sample.

Using Equation 4.5, we distorted these intrinsic ellipticities by the total shear pattern of one or more full modes (as defined in Equation 4.14), then applied simulated measurement noise (as described in Section 4.4.3) as a final step.

We fitted amplitudes per individual wave using LSQ, LAD and CHP, and per mode using LSQ and LAD by simultaneously fitting all four amplitudes. We then determined relative efficiencies and possible biases of the recovered amplitudes in the same way as in Section 4.5.1.

In Figure 4.15, we show the fitted shear field for a single realization, using in this case  $10^4$  simulated ellipticities. We fitted 16 different modes individually, using LSQ and LAD, and 64 individual amplitudes using LSQ, LAD, and CHP, and found the amplitude residuals, ( $O(10^{-3})$ ), to be two orders of magnitude less than the input values, which were constrained to  $g \leq 0.25$  for peak values at positive interference. Residuals in  $|e|$  for this realization varied between  $\pm 0.075$  for LSQ,  $\pm 0.066$  for LAD and  $\pm 0.14$  for CHP.

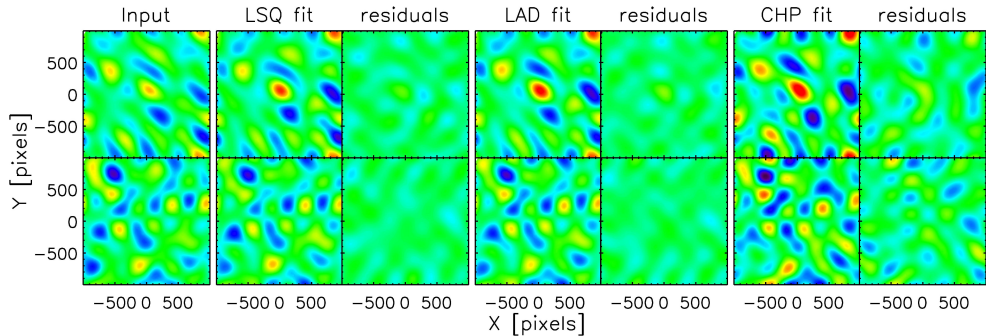


Figure 4.15: Shear field residuals when fitting 16 different modes, using simulated projected ellipsoids as intrinsic shapes, and including additional Gaussian noise. From left to right: input shear, LSQ fit and residuals, LAD fit and residuals, CHP it and residuals. Upper and lower row show  $e_1$  and  $e_2$  respectively. The color scale is the same in all plots for comparison and ranges between  $-0.247 \leq e_{1,2} \leq 0.247$ . Residuals for this realization vary between  $\pm 0.075, \pm 0.066$  and  $\pm 0.14$ , respectively.

### Bias and efficiency

The results from Sections 4.5.1 and 4.5.1 carry over to estimates of Fourier amplitudes for LSQ and LAD. We found fitted values with standard deviations of the order of  $10^{-3}$  for individual amplitudes. In Figure 4.16, we show the consistency of the fitted values.

Over plotted in Figure 4.16 are the best-fitting mean-bias, defined similar to Equation 4.23 as

$$a_{\text{est}} = (1 + m)a_{\text{in}}, \quad (4.25)$$

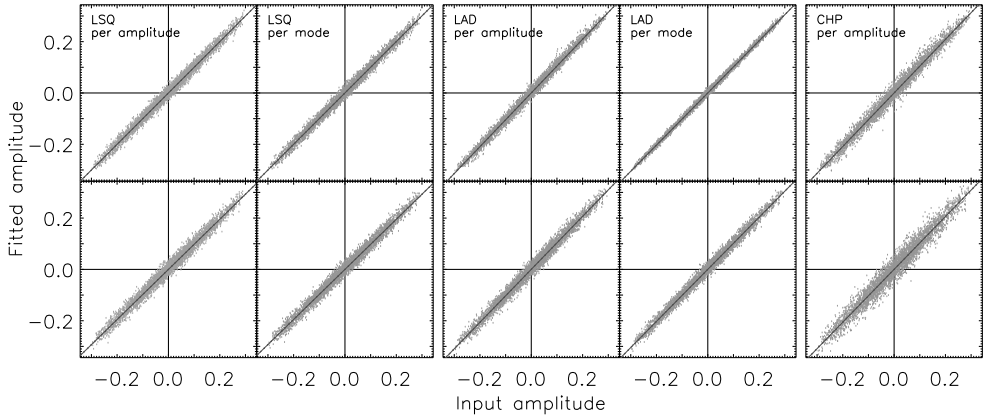


Figure 4.16: Consistency of the estimated Fourier amplitudes as defined in Equation 4.14. Plotted are the input amplitudes  $a_{mn}$ ,  $b_{mn}$ ,  $c_{mn}$ ,  $d_{mn}$  versus their estimates (Section 4.3.3) for LSQ (left), LAD (middle) AND CHP (right). Top row: only simulated intrinsic shapes based on projected ellipsoids. Bottom row: the same, with added noise. Over plotted are the best-fitting mean-bias, as defined in Equation 4.25.

where  $a$  stands for  $a_{mn}$ ,  $b_{mn}$ ,  $c_{mn}$ , and  $d_{mn}$  as defined in Equation 4.14. The uncertainties are too small to be visible. In Table 4.7 in the appendix, we give the quantitative results.

Similar to the results shown in Figure 4.11, LSQ underestimates Fourier amplitudes by a few percent in the presence of noise. For LAD, we found an improvement on bias by  $\sim 20\%$  in the presence of noise, when iteratively de-shearing the sample. Likewise, fitting for LAD without de-shearing slightly overestimated the amplitudes, again comparable to Figure 4.11.

We note that in this case, adding noise did not seem to have a significant effect when fitting per mode. In most cases, we did notice a significant increase in bias when fitting per single amplitude. We did not see a change in bias between LSQ per mode and per amplitude.

We also found a slightly higher relative efficiency of  $\eta_{68.3} = 1.09 \pm 0.07$  for LSQ and  $\eta_{68.3} = 1.47 \pm 0.09$  for LAD, when fitting per mode, with or without added noise. It is not surprising that a model with four parameters (amplitudes) fits the estimates better than a model with one parameter, but the difference of this effect between LSQ and LAD is noteworthy.

Since CHP doesn't fit a model to the data, but rather orders the  $(e_1, e_2)$  data points, there is no straightforward way to fit four amplitudes simultaneously with the necessary weighting (Section 4.3.2). We have not explored this option further in this paper.

The CHP estimator performs consistently, that is, convergent around the input values, but with a significant lower efficiency than for central value estimation of a cloud of  $(e_1, e_2)$  data points (Section 4.5.1). This is to be expected, since CHP is particularly sensitive to a (central) cusp in the distribution of data points. By shearing the intrinsic ellipticities by a model that varies over the field of view, as shown in Figure 4.6, this peak will be smeared out, decreasing the effectiveness of CHP.

In conclusion, CHP is consistently the most sensitive to the central cuspliness of a distribution. The results of this section do serve as a proof of concept for applying alternative statistics to an observed field of weak shear measurements.

## 4.6 Conclusions and summary

### 4.6.1 Optimal estimators

Our main conclusion is that to evaluate a statistical estimator, one must be willing to look beyond the canonical terms of mean-bias and the Gaussian variance as efficiency. We have shown that these commonly used meta-analysis instruments do not always properly reflect how well weak shear estimator values are constrained around the true underlying shear values.

By discussing the statistical definitions and observing the behavior of estimators for various ellipticity distributions, we have proposed ways of comparing various estimators motivated by statistical theory. The conclusions of that comparison are as follows:

Since the central peak of the intrinsic ellipticity distribution  $P(e)$  is an unbiased tracer of the underlying shear, we find that the LAD and CHP estimators are less biased and more efficient than the standard mean.

When iteratively de-shearing the ellipticity sample by the estimated shear, the LAD estimator can reach a sub-percent bias for typical weak shear values, including noise. LAD is generally the most efficient of all estimators considered, potentially reducing uncertainties by more than 50% for samples simulated using a model of projected triaxial ellipsoids.

The CHP estimator is in terms of its mean-bias less affected by noise, as compared to the mean and, to a lesser extend, LAD. In fact, since the estimator distribution  $P(\hat{e})$  is not symmetric, the actual center of that distribution, as opposed to the mean of that distribution, is unbiased in the presence of noise, within statistical significance. This makes CHP an important consideration, but it is less straightforward for adaptation for regression and requires careful assessment of uncertainties. Furthermore, CHP is computationally more demanding. In the presence of Gaussian noise, CHP is slightly less efficient than the mean (Figure 4.12, panels 3 and 4), but defining efficiency in terms of a figure of merit can reduce this drawback compared to the gain in bias, as proposed in Section 4.5.1.

### 4.6.2 Direct Fourier mode fitting

Applying different statistics to fitting individual Fourier modes to the shear field directly, we found results consistent with our previous conclusions.

We have shown that the Fourier amplitudes can be recovered with sub-percent accuracy and a minimal bias, which is an important proof of concept. Since the periodic variations in underlying shear effectively smooth the central peak of the intrinsic ellipticity distribution, the gain in efficiency is slightly less for LAD and significantly less for CHP. It is possible that an alternative to our weighting scheme for FMF with CHP could improve results. At this point, the method of CHP seems more applicable to samples of expected (roughly) constant shear, for example when measuring tangential shear around a gravitational lens candidate in bins of distance.

We have also shown that the shear field can be recovered reliably, with residuals an the order of magnitude less than the variations of the shear over the field of view for LSQ and LAD, using  $10^3$  sources, which is conservative compared to a typical single CFHTLenS field.

### 4.6.3 Future considerations and possible applications

We have discussed alternative statistics for inference of shear from samples of background sources with various intrinsic ellipticity distributions, proposing methods that could improve biases and uncertainties arising from the shape noise. It is important to consider our results within the broader context of other sources of systematics, as mentioned in our introduction.

Firstly, our results for shape noise assume trustworthy shape measurements, not only per source, but also considering the effect of systematics in the shape measurement pipelines on the reproduced ellipticity distribution as a whole: the recovery of a central peak, the distribution of outliers, among others. Examples are the effect of constraining ellipticities to a certain ‘physical’ maximum (e.g.,  $e_{\max} = 0.804$  for *lensfit*, Miller et al. 2013), as we see in Figure 4.2, or conversely, the unphysical outliers with  $|e| > 1.0$  arising from dividing two noisy quantities (often when correcting for the point spread function, or PSF), affecting the tails of the distribution. Any features in the recovered shape distribution could affect bias and efficiency of the statistic used. Optimizing statistics will place more stringent demands on shape measurements than performing excellent ‘on average’. Even methods that avoid individual shape measurements (Bernstein & Armstrong 2014), an ensemble inferred reduced shear could improve by considering the intrinsic shape distribution. Secondly, even with an accurately measured shape distribution, there will remain sources of systematic error in other steps of a cosmological analysis, as noted in our introduction. These effects still form a necessary part in a weak lensing analysis, but leave our statistical conclusions unaffected.

As survey sizes and image qualities increase, so will the demands on constraining systematic effects to a sub-dominant level, as described in for example Kuijken et al. (2015) and Mandelbaum et al. (2018) for the KiDS DR2 and HSC DR1, respectively. At the same time, it will be interesting to see measured ellipticity distributions converge as more sources are observed with higher signal-to-noise and measured with higher fidelity, due to increased depth of imaging, image quality and PSF control.

For now, we have given a proof of concept for alternative statistics in two cases: a sample of ellipticities with one underlying shear and the recovery of individual Fourier modes of the shear variation over a field of view. The first part has important applications when inferring a shear profile around lenses, both in recovering an accurate, less biased estimate and smaller error bars or confidence intervals. For the second part: since the amplitudes are well constrained by fitting individual Fourier modes, this provides a possible method toward estimation of the power spectrum. Furthermore, the shear field can be recovered in terms of its Fourier amplitudes, providing a powerful analytic model for mass reconstruction, without the need for smoothed gridding and incorporating variations in background source densities and estimated measurement uncertainties.

## Acknowledgements

We thank the anonymous referee for useful comments and suggestions, which improved readability and reproducibility of this paper and its results. MS acknowledges support from the Netherlands Organization for Scientific Research (NWO).

This work is based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences

de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. This research used the facilities of the Canadian Astronomy Data Centre operated by the National Research Council of Canada with the support of the Canadian Space Agency. CFHTLenS data processing was made possible thanks to significant computing support from the NSERC Research Tools and Instruments grant program.

## 4.A Bias estimations

In Table 4.2 are given the mean-bias and CHP-bias for each estimator, with and without de-shearing. This multiplicative bias  $m$  is defined by Equation 4.23 as  $e_{\text{fit}} = (1 + m)e_{\text{in}}$ .

## 4.B Efficiency estimations

In Tables 4.3, 4.4, and 4.5, we summarize the full results for the relative efficiencies of each estimator. We determine relative efficiencies for coverages of 25%, 50%, and 75%, corresponding to the MAD and the first and third quartiles, and 38.3%, 68.3%, 86.6%, and 95.4%, which would correspond to steps of  $0.5\sigma$  in case of a Gaussian distribution with variance  $\sigma^2$ .

For easy reference, we also indicate how much the scale of the estimator distribution would improve, in percentages of the scale of the distribution of the mean estimator,

$$\Delta s_p = \frac{s_{\hat{e},p}}{s_\mu} - 1 \quad (\text{in } \%) \quad (4.26)$$

Since a higher efficiency means a smaller scale and therefore a more 'trustworthy' estimate, this is an intuitive, albeit rough indication of the change in error bars.

## 4.C Estimations from Fourier mode fitting

In Table 4.7 are given the mean-bias and efficiencies of the LSQ estimator, per mode and per individual amplitude, LAD estimator with and without de-shearing, per mode and per individual amplitude, and the CHP estimator, per individual amplitude. The mean-bias is again given in terms of a multiplicative component  $m$  as defined in Equation 4.25.

Table 4.2: Results for bias estimations for each estimator, with and without de-shearing. We have used simulated projected ellipsoids as intrinsic ellipticities, with and without added noise. Estimation bias is given in terms of a multiplicative component  $m$  as defined in Equation 4.23. Numbers in parentheses reflect the standard uncertainty in the last digit.

Estim.	$m_\mu$	$m_{\text{CHP}}$
<i>Simulated ellipticities</i>		
Mean	-0.0003(4)	-0.0007(7)
de-g	0.0006(6)	0.0004(4)
LAD	0.0172(4)	0.0154(5)
de-g	-0.0000(2)	-0.0005(2)
CHP	0.0097(2)	0.0052(5)
de-g	0.0000(2)	0.0001(2)
BI	0.0426(7)	0.0421(9)
de-g	-0.0002(6)	0.0002(5)
<i>Added noise</i>		
Mean	-0.0404(7)	-0.040(1)
de-g	-0.045(2)	-0.045(2)
LAD	0.0114(6)	0.011(1)
de-g	-0.0308(8)	-0.031(1)
CHP	0.0053(5)	0.004(1)
de-g	-0.030(1)	-0.028(2)
BI	0.0066(8)	0.007(1)
de-g	-0.040(1)	-0.040(2)

Table 4.3: Results for scales of fixed coverage for LAD. For each sample distribution, the relative efficiencies  $\eta$  are given first, and the (more intuitive) relative change in estimator distribution scale is given second, in percentages of the distribution scale of the mean. Numbers in parentheses reflect the standard uncertainty in the last digit.

Distribution	$\eta_{25}$ $\Delta s_{25}(\%)$	$\eta_{38.3}$ $\Delta s_{38.3}(\%)$	$\eta_{50}$ $\Delta s_{50}(\%)$	$\eta_{68.3}$ $\Delta s_{68.3}(\%)$	$\eta_{75}$ $\Delta s_{75}(\%)$	$\eta_{86.6}$ $\Delta s_{86.6}(\%)$	$\eta_{95.4}$ $\Delta s_{95.4}(\%)$
<i>Simulated ellipticities</i>							
Gaussian	0.79(3) +12(2)	0.79(3) +12(2)	0.81(3) +11(2)	0.83(2) +9(1)	0.84(3) +9(2)	0.81(2) +11(1)	0.83(3) +10(2)
Uniform $q$	7.2(4) -63(1)	6.1(3) -59.4(8)	5.4(1) -56.9(5)	4.55(9) -53.1(5)	4.3(1) -51.7(7)	3.69(9) -48.0(6)	3.22(6) -44.3(5)
Elliptical	2.7(1) -39(2)	2.5(1) -36(2)	2.4(1) -35(1)	2.26(8) -34(1)	2.20(9) -33(1)	2.07(8) -31(1)	1.88(7) -27(1)
Disk	6.3(3) -60.1(9)	5.8(2) -58.5(8)	5.5(2) -57.3(6)	5.1(2) -55.6(7)	4.9(1) -54.7(7)	4.7(1) -53.9(7)	4.3(1) -52.0(7)
Combined	5.4(2) -57.2(8)	5.0(3) -55(1)	4.9(2) -55(1)	4.6(2) -53.6(8)	4.4(1) -52.2(7)	4.1(1) -50.8(7)	4.0(2) -50(1)
<i>Added noise</i>							
Uniform $q$	1.37(7) -15(2)	1.38(5) -15(2)	1.37(5) -14(2)	1.38(3) -15(1)	1.37(3) -14(1)	1.33(3) -13(1)	1.29(4) -12(1)
Elliptical	1.19(6) -8(2)	1.18(5) -8(2)	1.20(6) -9(2)	1.17(3) -8(1)	1.18(4) -8(1)	1.19(3) -8(1)	1.16(4) -7(2)
Disk	1.48(8) -18(2)	1.49(7) -18(2)	1.54(7) -19(2)	1.55(6) -20(1)	1.57(5) -20(1)	1.59(5) -21(1)	1.52(4) -19(1)
Combined	1.55(5) -20(1)	1.53(6) -19(2)	1.51(8) -19(2)	1.53(5) -19(1)	1.51(5) -19(1)	1.48(6) -18(2)	1.46(5) -17(1)
<i>Full CFHTLenS data</i>							
All	1.17(6) -8(3)	1.19(4) -8(1)	1.19(3) -8(1)	1.21(3) -9(1)	1.20(2) -8.6(9)	1.20(3) -9(1)	1.12(3) -6(1)
Red	1.32(7) -13(2)	1.33(6) -13(2)	1.35(5) -14(1)	1.26(4) -11(1)	1.24(5) -10(2)	1.19(5) -8(2)	1.19(8) -8(3)
Blue	1.28(4) -12(1)	1.24(5) -10(2)	1.24(4) -10(1)	1.20(3) -9(1)	1.18(3) -8(1)	1.19(5) -8(2)	1.19(8) -8(3)
Sbc	1.23(5) -10(2)	1.19(3) -8(1)	1.19(5) -9(2)	1.18(2) -7.8(6)	1.17(3) -7.4(9)	1.18(3) -8(1)	1.13(6) -6(2)
Scd	1.21(4) -9(2)	1.20(4) -9(2)	1.21(4) -9(2)	1.18(4) -8(1)	1.16(3) -7(1)	1.19(6) -8(2)	1.12(7) -6(3)
<i>Conservative selection of CFHTLenS data</i>							
All	1.82(7) -26(1)	1.77(6) -25(1)	1.72(4) -24(1)	1.67(5) -23(1)	1.70(3) -23.3(7)	1.65(2) -22.2(4)	1.67(6) -23(1)
Red	1.8(1) -26(3)	1.68(9) -23(2)	1.66(7) -22(2)	1.64(6) -22(1)	1.64(6) -22(1)	1.66(6) -22(1)	1.58(6) -21(2)
Blue	1.73(4) -24.0(8)	1.63(5) -22(1)	1.62(5) -21(1)	1.65(4) -22(1)	1.64(6) -22(1)	1.64(5) -22(1)	1.61(9) -21(2)
Sbc	1.63(8) -22(2)	1.6(1) -22(2)	1.59(6) -21(2)	1.55(4) -20(1)	1.52(4) -19(1)	1.49(4) -18(1)	1.47(7) -18(2)
Scd	1.8(1) -25(3)	1.75(9) -24(2)	1.70(6) -23(1)	1.67(4) -22.6(9)	1.6(4) -21.9(9)	1.65(3) -22.1(7)	1.55(5) -20(1)

Table 4.4: Results for scales of fixed coverage for the biweight. For each sample distribution, the relative efficiencies  $\eta$  are given first, and the (more intuitive) relative change in estimator distribution scale is given second, in percentages of the distribution scale of the mean. Numbers in parentheses reflect the standard uncertainty in the last digit.

Distribution	$\eta_{25}$ $\Delta s_{25}(\%)$	$\eta_{38.3}$ $\Delta s_{38.3}(\%)$	$\eta_{50}$ $\Delta s_{50}(\%)$	$\eta_{68.3}$ $\Delta s_{68.3}(\%)$	$\eta_{75}$ $\Delta s_{75}(\%)$	$\eta_{86.6}$ $\Delta s_{86.6}(\%)$	$\eta_{95.4}$ $\Delta s_{95.4}(\%)$
<i>Simulated ellipticities</i>							
Gaussian	0.99(6) +0(3)	1.02(6) -1(3)	1.00(4) -0(2)	1.03(4) -1(2)	1.02(5) -1(2)	1.03(4) -1(2)	1.05(4) -2(2)
Uniform $q$	1.2(1) -10(3)	1.16(5) -7(2)	1.19(2) -8.2(8)	1.14(3) -6(1)	1.12(2) -6(1)	1.14(5) -6(2)	1.17(6) -7(2)
Elliptical	1.4(2) -14(5)	1.3(1) -13(2)	1.31(8) -13(3)	1.31(8) -13(2)	1.30(6) -12(2)	1.34(9) -14(3)	1.32(3) -13(1)
Disk	1.6(1) -20(3)	1.60(7) -21(2)	1.52(6) -19(2)	1.49(6) -18(2)	1.49(6) -18(2)	1.49(5) -18(1)	1.46(4) -17(1)
Combined	1.54(8) -19(2)	1.5(1) -19(3)	1.5(1) -18(3)	1.46(5) -17(2)	1.47(6) -17(2)	1.47(5) -17(1)	1.45(8) -17(2)
<i>Added noise</i>							
Uniform $q$	1.09(6) -4(2)	1.11(5) -5(2)	1.11(7) -5(3)	1.09(4) -4(2)	1.09(5) -4(2)	1.09(6) -4(2)	1.10(4) -5(2)
Elliptical	1.23(8) -10(3)	1.26(9) -11(3)	1.23(9) -10(3)	1.248 -10(3)	1.23(8) -10(3)	1.24(6) -10(2)	1.22(4) -9(2)
Disk	1.25(6) -10(2)	1.24(9) -10(3)	1.24(7) -10(2)	1.23(4) -10(1)	1.23(4) -10(2)	1.23(6) -10(2)	1.19(4) -8(1)
Combined	1.23(7) -10(2)	1.25(6) -10(2)	1.24(7) -10(3)	1.26(6) -11(2)	1.26(7) -11(2)	1.2397 -10(3)	1.22(6) -9(2)
<i>Full CFHTLenS data</i>							
All	1.06(7) -3(3)	1.06(4) -3(2)	1.05(3) -2(2)	1.10(3) -5(1)	1.09(3) -4(2)	1.08(3) -4(1)	1.06(2) -3.1(8)
Red	1.1(1) -5(6)	1.10(8) -5(3)	1.09(5) -4(2)	1.08(5) -4(2)	1.08(6) -4(3)	1.08(6) -4(3)	1.1(1) -4(5)
Blue	1.09(3) -4(1)	1.06(4) -3(2)	1.05(7) -3(3)	1.08(5) -4(2)	1.10(5) -5(2)	1.10(7) -5(3)	1.07(8) -3(4)
Sbc	1.12(9) -5(4)	1.08(7) -4(3)	1.07(6) -3(3)	1.06(3) -3(1)	1.07(4) -3(2)	1.06(4) -3(2)	1.03(5) -2(2)
Scd	1.02(4) -1(2)	1.03(4) -1(2)	1.05(4) -3(2)	1.05(4) -2(2)	1.04(5) -2(2)	1.06(8) -3(4)	1.02(9) -1(4)
<i>Conservative selection of CFHTLenS data</i>							
All	1.22(6) -10(2)	1.23(3) -10(1)	1.21(4) -9(2)	1.19(5) -8(2)	1.22(5) -9(2)	1.22(3) -9(1)	1.21(7) -9(3)
Red	1.12(5) -6(2)	1.10(5) -5(2)	1.12(5) -5(2)	1.11(4) -5(2)	1.11(4) -5(2)	1.15(6) -7(2)	1.14(50) -6(2)
Blue	1.27(8) -11(3)	1.22(7) -9(3)	1.23(8) -10(3)	1.23(7) -10(3)	1.23(5) -10(2)	1.23(5) -10(2)	1.21(6) -9(2)
Sbc	1.14(6) -6(2)	1.2(1) -7(5)	1.14(7) -6(3)	1.13(3) -6(1)	1.13(4) -6(2)	1.15(7) -7(3)	1.15(4) -7(1)
Scd	1.3(1) -14(4)	1.3(1) -14(4)	1.30(9) -12(3)	1.28(8) -12(3)	1.25(8) -10(3)	1.26(6) -11(2)	1.28(4) -11(1)



Table 4.5: Results for scales of fixed coverage for CHP. For each sample distribution, the relative efficiencies  $\eta$  are given first, and the (more intuitive) relative change in estimator distribution scale is given second, in percentages of the distribution scale of the mean. Numbers in parentheses reflect the standard uncertainty in the last digit.

Distribution	$\eta_{25}$ $\Delta s_{25}(\%)$	$\eta_{38.3}$ $\Delta s_{38.3}(\%)$	$\eta_{50}$ $\Delta s_{50}(\%)$	$\eta_{68.3}$ $\Delta s_{68.3}(\%)$	$\eta_{75}$ $\Delta s_{75}(\%)$	$\eta_{86.6}$ $\Delta s_{86.6}(\%)$	$\eta_{95.4}$ $\Delta s_{95.4}(\%)$
<i>Simulated ellipticities</i>							
Gaussian	0.38(3) +62(6)	0.38(1) +62(3)	0.38(1) +63(3)	0.39(2) +61(3)	0.38(1) +63(2)	0.37(1) +63(2)	0.38(2) +63(5)
Uniform $q$	15(2) -74(2)	11.1(9) -70(1)	8.9(5) -66.6(9)	6.5(5) -61(1)	5.6(4) -58(1)	4.4(3) -53(1)	3.3(3) -45(2)
Elliptical	2.3(2) -33(2)	2.03(8) -30(1)	1.80(7) -25(1)	1.58(5) -21(1)	1.51(6) -19(2)	1.31(8) -13(3)	1.13(8) -6(3)
Disk	6.8(7) -62(2)	6.0(5) -59(2)	5.4(2) -57.1(9)	4.6(1) -53.5(6)	4.3(1) -51.6(8)	3.8(2) -49(1)	3.2(2) -44(2)
Combined	5.8(3) -59(1)	5.0(1) -55.4(6)	4.7(2) -53.8(8)	4.1(1) -50.4(7)	3.8(1) -48.8(9)	3.4(1) -45(1)	2.7(2) -39(2)
<i>Added noise</i>							
Uniform $q$	0.85(4) +7(2)	0.84(3) +9(2)	0.86(4) +8(3)	0.85(4) +8(2)	0.84(3) +9(2)	0.80(3) +12(2)	0.77(40) +14(3)
Elliptical	0.65(5) +24(5)	0.65(6) +24(6)	0.64(4) +25(4)	0.66(3) +24(3)	0.64(3) +25(3)	0.64(3) +25(3)	0.64(4) +25(4)
Disk	0.88(7) +7(4)	0.89(6) +6(4)	0.88(4) +6(3)	0.90(3) +6(2)	0.90(3) +5(1)	0.88(4) +6(2)	0.87(6) +7(4)
Combined	0.91(6) +5(3)	0.92(6) +4(4)	0.89(5) +6(3)	0.91(3) +5(2)	0.89(4) +6(2)	0.88(5) +6(3)	0.86(4) +8(2)
<i>Full CFHTLenS data</i>							
All	0.75(4) +16(3)	0.76(3) +14(2)	0.77(2) +14(2)	0.76(2) +15(1)	0.77(2) +14(1)	0.75(3) +16(2)	0.71(2) +19(1)
Red	0.81(7) +11(5)	0.84(7) +9(4)	0.84(4) +9(2)	0.77(4) +14(3)	0.76(5) +15(4)	0.72(4) +18(3)	0.72(4) +18(3)
Blue	0.81(3) +11(2)	0.79(3) +13(2)	0.77(3) +14(2)	0.75(4) +15(3)	0.76(3) +14(3)	0.73(3) +17(2)	0.68(6) +21(5)
Sbc	0.80(4) +12(3)	0.78(3) +14(2)	0.77(3) +14(2)	0.76(3) +15(2)	0.74(2) +16(2)	0.71(3) +18(2)	0.68(4) +21(4)
Scd	0.79(5) +12(3)	0.75(3) +15(2)	0.74(2) +16(2)	0.73(3) +17(2)	0.72(4) +18(3)	0.72(3) +18(3)	0.67(5) +22(5)
<i>Conservative selection of CFHTLenS data</i>							
All	1.18(5) -8(2)	1.14(4) -6(1)	1.08(5) -4(2)	1.00(5) 0(3)	0.99(4) 0(2)	0.96(4) +2(2)	0.89(5) +6(3)
Red	1.3(1) -13(3)	1.21(8) -9(3)	1.15(5) -7(2)	1.05(5) -2(2)	1.02(5) -1(2)	0.99(5) 0(2)	0.89(8) +6(5)
Blue	1.09(5) -4(2)	1.02(6) -1(3)	1.03(7) -1(3)	0.99(5) +1(3)	0.96(5) +2(3)	0.97(5) +2(3)	0.93(9) +4(5)
Sbc	1.1(1) -4(5)	1.1(1) -3(5)	1.0(1) -1(5)	0.98(5) +1(2)	0.95(5) +3(3)	0.91(4) +5(2)	0.88(4) +7(2)
Scd	1.11(9) -5(4)	1.11(4) -5(2)	1.06(6) -3(3)	1.00(6) 0(3)	0.97(5) +1(3)	0.94(5) +3(2)	0.85(6) +8(3)

Table 4.6: Results for  $\eta_{68.3}$  for different sample sizes. Numbers in parentheses reflect the standard uncertainty in the last digit.

Distribution	Estimator	$N = 10$	$N = 22$	$N = 46$	$N = 100$	$N = 215$	$N = 464$
<i>Simulated ellipticities</i>							
Gaussian	LAD	0.80(3)	0.76(2)	0.80(3)	0.83(2)	0.79(3)	0.78(3)
	BI	1.01(3)	1.06(3)	1.03(2)	1.03(4)	1.03(5)	1.07(5)
	CHP	0.71(2)	0.50(1)	0.44(1)	0.39(2)	0.33(2)	0.29(1)
Uniform $q$	LAD	1.84(6)	2.78(5)	3.6(1)	4.55(9)	5.8(2)	6.6(4)
	BI	1.16(9)	1.14(4)	1.20(3)	1.14(3)	1.17(4)	1.18(9)
	CHP	1.30(5)	2.19(7)	3.6(2)	6.5(5)	12.3(5)	22(2)
Combined	LAD	2.94(7)	3.5(2)	4.2(1)	4.6(2)	4.9(1)	5.1(2)
	BI	1.47(7)	1.46(7)	1.45(4)	1.46(5)	1.56(6)	1.53(8)
	CHP	2.03(9)	2.8(2)	3.6(2)	4.1(1)	5.1(2)	6.0(4)
<i>Added noise</i>							
Uniform $q$	LAD	1.08(3)	1.27(7)	1.32(6)	1.38(3)	1.41(4)	1.33(4)
	BI	1.09(3)	1.08(7)	1.15(4)	1.09(4)	1.08(4)	1.14(8)
	CHP	0.9(8)	0.93(6)	0.93(3)	0.85(4)	0.74(4)	0.66(4)
Combined	LAD	1.28(5)	1.42(3)	1.53(5)	1.53(5)	1.51(6)	1.56(5)
	BI	1.24(4)	1.23(5)	1.26(3)	1.26(6)	1.25(7)	1.24(8)
	CHP	0.98(9)	1.04(4)	0.98(8)	0.91(3)	0.78(4)	0.70(4)
<i>CFHTLenS data</i>							
All	LAD	0.91(2)	1.11(5)	1.16(4)	1.21(3)	1.15(3)	1.21(5)
	BI	1.05(2)	1.07(4)	1.10(6)	1.10(3)	1.08(3)	1.09(8)
	CHP	0.77(4)	0.83(6)	0.82(4)	0.76(2)	0.66(2)	0.65(4)
Subset	LAD	1.39(4)	1.59(6)	1.59(5)	1.67(5)	1.74(4)	1.80(4)
	BI	1.17(3)	1.18(6)	1.22(5)	1.19(5)	1.21(6)	1.19(3)
	CHP	1.07(4)	1.12(6)	1.04(3)	1.00(5)	0.99(4)	0.98(6)

Table 4.7: Results for amplitude estimations for FMF, using different estimators (LSQ, LAD, CHP) and models (per mode or per amplitude). Estimation bias is given in terms of a multiplicative component  $m$  as defined in Equation 4.25. Efficiencies are determined relative to LSQ per individual amplitude. Numbers in parentheses reflect the standard uncertainty in the last digit.

Estimator		$m$	$\eta_{68.3}$
<i>Simulated ellipticities</i>			
LSQ	per mode	0.001(1)	1.14(6)
	amplitude	0.000(1)	N.A.
LAD	per mode	0.0166(6)	4.6(3)
	amplitude	0.046(1)	1.31(4)
	de-g per mode	0.0006(6)	5.4(4)
	amplitude	0.022(1)	1.61(5)
CHP	per amplitude	0.043(2)	0.83(5)
<i>Added noise</i>			
LSQ	per mode	-0.024(1)	1.09(7)
	amplitude	-0.023(2)	N.A.
LAD	per mode	0.014(1)	1.47(9)
	amplitude	0.032(2)	0.98(7)
	de-g per mode	-0.018(3)	1.55(9)
	amplitude	-0.007(2)	1.14(8)
CHP	per amplitude	0.015(2)	0.41(3)