Mon. Not. R. Astron. Soc. **000**, 1–12 (2002) Printed 2 July 2015 (MN LATEX style file v2.2)

Cluster STEP

J. Young[1⋆] and Other.
[1] *OSU/USM*

2013

ABSTRACT

We present results from the Cluster Shear TEsting Program (CSTEP), an image analysis challenge to measure weak gravitational shear in the cluster regime. Shape measurement bias is an important source of systematic error in the measurement of cluster masses by weak gravitational lensing and, consequently cluster cosmology. In this work the accuracy of the shape measurement pipelines is determined from image simulations that have realistic distributions of galaxy properties and large gravitational shear. Eight different shape measurement pipelines were tested, and the best performing methods exhibit a multiplicative bias of a few percent that is stable as a function of redshift. Our results also show that the methods tested here do not have a strong quadratic bias. We also find that the amount that the shear bias measured was due to selection effects varied widely among the methods. We model the impact of the biases found on a simulated stage III optical weak lensing cluster cosmology survey, such as the Dark Energy Survey (DES) and find it to be below the statistical uncertainty which can be further reduced by simulation based calibration.

Key words: gravitational lensing: weak, techniques:image processing

1 INTRODUCTION

A primary cosmological probe is the number and distribution in mass of galaxy clusters in a given volume of space (Allen, Evrard & Mantz 2011). There are a number of different tools to detect clusters including X-ray and SZ but these methods provide observables which must then be correlated with mass by weak lensing. Large optical surveys such as the Dark Energy Survey (DES) (http://www.darkenergysurvey.org/) are able to detect and measure the mass of large numbers of clusters and place strong constraints on the halo mass function if sources of systematic error in shear measurement can be controlled. Shape measurement pipelines have been developed to remove the effect of the PSF and accurately measure the lensing signal. As the mass measured from a galaxy cluster depends on the strength of the shear measured on galaxies behind the cluster, a biased weak lensing pipeline will lead to a biased determination of the cluster mass.

Previous image simulation challenges were used to determine the accuracy of shape measurement pipelines, STEP1 (Heymans et al. 2006), STEP2 (Massey et al. 2007), GREAT08 (Bridle et al. 2010), GREAT10 (Kitching et al. 2012), and GREAT3 (Mandelbaum et al. 2014), and used blinded image simulations to characterize the shape measurement bias created by different lensing pipelines. All of these challenges have focused on different aspects of shear estimation, and have led to improvements in existing shear pipelines. These simulation challenges provide a way to more directly compare the accuracy of shear measurement pipelines against each other on simulated data with different galaxy morphology, seeing and Signal to Noise ratio (SNR). Results from these challenges have been used to calibrate shape measurement pipelines (e.g. Applegate et al. 2012), and determine that lensing pipelines meet accuracy requirements for scientific analyses (e.g. Bergé et al. 2008). Although previous challenges provided valuable information on many important aspects of shear pipeline estimation, they were designed to test the ability of shear pipelines to measure cosmic shear with $|g_{true}| \leqslant 0.06$. It is important, however, to also test the accuracy of shape measurement pipelines in the higher shear regime that occurs around the center of massive galaxy clusters where selection effects and non-linear bias may play a more important role.

The Cluster Shear TEsting Program (CSTEP) is designed to extend image simulation challenges to the high shear regime and study aspects of shear estimation that would impact a weak lensing measurement of cluster mass for a ground based large optical survey. For this, we simulate ground-based images of galaxies with a realistic distribution of SNR, size and ellipticity. Lensing pipelines were evaluated to determine their shape measurement bias as a function of these properties. The shape measurement bias as

⋆ E-mail: email@address (AVR); otheremail@otheraddress (ANO)

determined for each lensing pipeline is then used to model the systematic error in the cluster mass for a stacked weak lensing measurement.

In Section 2 we describe the structure of the challenge and the image simulations used in this project. We outline the characteristics of the shape measurement pipelines used in this analysis in Section 3. We present our findings on shape measurement bias including the impact of S/N, selection effects, quadratic bias and redshift in Section 4. The effect of shape measurement biases on stacked cluster weak lensing analyses is discussed in Section 5. Conclusions from this study are given in Section 6. Finally there is an appendix which contains some further description of pipelines.

2 CLUSTER STEP CHALLENGE

Gravitational lensing is the distortion of the images of distant galaxies by the gravitational pull of massive structures near their line of sight. This can cause a change in the observed shape, size, and brightness of a galaxy. For massive galaxy clusters this leads to galaxies at a higher redshift than the cluster often being observed as arcs (strong lensing) near the cluster center , and to appear aligned tangentialy with decreasing amount of distortion further from the cluster center (weak lensing). Since galaxies are assumed to have random orientations, a measurement of the average ellipticity of a group of galaxies unaffected by gravitational lensing would average to zero. Measuring the degree in which the observed galaxies appear aligned so the ensemble average ellipticity diverges from zero provides a way to measure the strength of the gravitational lensing signal, and thus the mass of the galaxy cluster.

For the larger shears of the cluster regime both the quadratic (q) and multiplicative (m) shape measurement bias (see eq. 1) affect the bias in lensing signal to a greater extent than any constant (c) shear bias, since the later averages when measuring tangential shear over a full circle. It is important to quantify the quadratic and multiplicative bias on images with simulated shear comparable to the shear observed around large galaxy clusters. The Cluster Shear TEsting Program (CSTEP) tested eight weak lensing pipelines on images of constant shear ($|\gamma| = [0.03, 0.06, 0.09, 0.15]$) with galaxy and PSF properties that simulate the properties of data that will be observed with DES. Systematic shape measurement bias was quantified by the q, m and c determined by a fit of the quadratic equation.

$$\gamma_m = q(\gamma_t)^2 + (1+m)\gamma_t + c, \quad (1)$$

where γ_t is the true shear and γ_m is the measured shear. A simpler model that has been used to quantify error for cosmic shear regimes was also used.

$$\gamma_m = (1+m)\gamma_t + c, \quad (2)$$

An example fit to the im3shape data is shown in Figure 1.

For the CSTEP simulated images, mock galaxy catalogs were created with the ADDGALS method that reproduce observed correlations for galaxies among clustering, luminosities, and colors. The parent N-body simulations used by ADDGALS are the "Carmen" Las Damas N-body simulation box (1 Gpc/H^3) with ΛCDM cosmology,

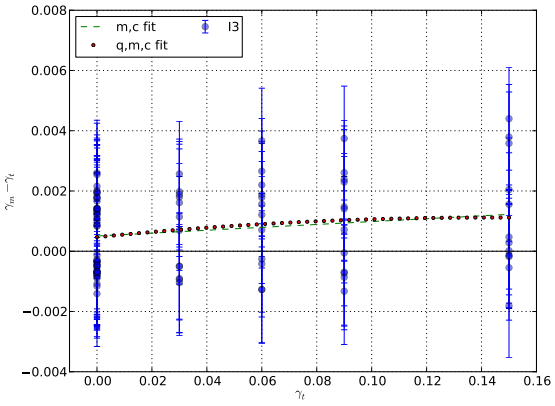

Figure 1. An example of the two types of fit used to quantify shape measurement error for the shear as measured by the im3shape pipeline (I3) on galaxy objects with Signal to Noise Ratio greater than 50.

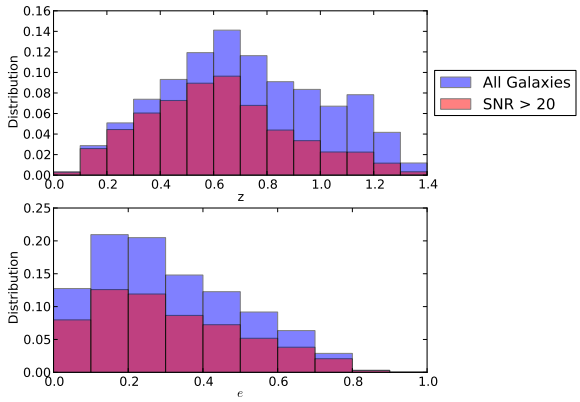

Figure 2. Distributions of various galaxy properties in the CSTEP simulation. The top plot shows the redshift distribution of galaxies as a function of percentage in the simulation. The bottom plot shows the distribution of intrinsic ellipticity as a function of percentage of total objects. The total number of galaxy objects included is 141559 and is shown in blue, the number of galaxy obejects with SNR > 20 is 84586 and is shown in magenta.

$\Omega_M = 0.25, \sigma_8 = 0.8$, at $z < 1.35$ (McBride et al. (2009)). In each image around 140,000 galaxy objects are present. The simulated galaxy images are designed to have similar properties to the galaxies that DES will observe by combining a mock catalog with empirical data from the HST/GEMS catalog for galaxies with magnitude r > 23 , and from the Sloan Digital Sky Survey for galaxies with magnitude r < 23. The distributions of galaxy properties in the simulation are shown in Figure 2. The galaxy objects are created with Sersic profiles with an index that ranges from 0.5 to 5.
The image simulations used in the CSTEP project were designed to replicate the properties of the final co-added data that will be taken by the Dark Energy Survey. The Dark Energy survey is an optical survey of 5000 deg^2 of the southern sky, that will image around 300 million galaxies in five filters (g,r,i,z and Y) (Honscheid & DePoy 2008). To test the

PSF Number	Type	Seeing
1	circular Gaussian	0.7
2	circular Gaussian	0.8
3	circular Gaussian	0.9
4	elliptical Gaussian	0.75
5	elliptical Gaussian	0.8
6	elliptical Gaussian	0.9
Shear Number	γ_1	γ_2
1	0.0	0.03
2	0.0	0.06
3	0.0	0.09
4	0.0	0.15
5	0.03	0.0
6	0.06	0.0
7	0.09	0.0
8	0.15	0.0

Table 1. Summary of the PSF and shear sets for the CSTEP simulated images.

DES data management system and software pipelines developed by the science working groups detailed image simulations were created by the DES Simulation team (Mohr et al. 2008). Each simulated image used in CSTEP models the DES focal plane with 62 CCDs each of which is 2K x 4K pixel (0.27"/pixel). Star and galaxy objects are rendered by drawing random samples of photons from the theoretical light profile of the source convolved by the PSF.
For the CSTEP simulated images, mock galaxy catalogs were created with the ADDGALS method that reproduce observed correlations for galaxies among clustering, luminosities, and colors. The parent N-body simulations used by ADDGALS are the "Carmen" Las Damas N-body simulation box (1 Gpc/H^3) with ΛCDM cosmology, $\Omega_M = 0.25, \sigma_8 = 0.8$, at $z < 1.35$ (McBride et al. (2009)). In each image around 140,000 galaxy objects are present. The simulated galaxy images are designed to have similar properties to the galaxies that DES will observe by combining a mock catalog with empirical data from the HST/GEMS catalog for galaxies with magnitude r > 23 , and from the Sloan Digital Sky Survey for galaxies with magnitude r < 23. The distributions of galaxy properties in the simulation are shown in Figure 2. The galaxy objects are created with Sersic profiles with an index that ranges from 0.5 to 5.

The CSTEP simulated images contain a constant PSF which eliminates the technical challenge of determining the best way to model a varying PSF. For CSTEP there are two branches in the image simulation sets, those that include a circular Gaussian PSF and those that include an elliptical Gaussian PSF. The elliptical Gaussian PSF has $\epsilon = 0.03$. For both the Gaussian and elliptical Gaussian PSF the simulated images contain a convolution kernel that mimics the range of isotropic distortion due to the atmosphere that DES will likely observe. This isotropic distortion simulates seeing conditions, or atmospheric blurring, of $0.7 - 0.9$ arcsec, depending on the simulated atmosphere. A summary of the PSF image properties is included in Table 1. For each PSF there are 8 focal plane images that contain constant shear for $\gamma = 0.0$ to $\gamma = 0.15$ in both γ_1 and γ_2, as described in Table 1.

3 SHAPE MEASUREMENT PIPELINES

There are eight weak lensing pipelines that submitted results for the CSTEP project, shown in Table 2. There are several lensing pipelines tested in this challenge that could potentially be used to create shear catalogs for DES. These implementations represent a broad range of lensing pipeline types, and can be divided into four groups. Methods are grouped using roughly the same criteria as in STEP2 (Massey et al. 2007). A short description of the methods is included below with a detailed pipeline description included in Appendix A.

3.1 Moment-based methods: red class

The red class methods (DE, IM, PK, KM) are based on the oldest shape measurement method KSB+ developed by Kaiser, Squires, and Broadhurst in 1995. As in STEP2, red class methods are defined as those that measure combinations of moments of each galaxy image $I(x)$ using a Gaussian weighting function. Various implementations of this class of methods have been studied in previous shape measurement challenges, and the accuracy has been shown to vary widely (Massey et al. 2007; Kitching et al. 2012). Two of the red class lensing pipeline implementations included in CSTEP have performed well on previous image simulations challenges (DE, KM). Other red class implementations that were calibrated on simulated image data have been used in several recent weak lensing analyses(e.g. Gruen et al. 2013b, 2014; Applegate et al. 2012; Schrabback et al. 2010). Many of the large weak lensing cluster studies have used red class methods (Mahdavi et al. 2013; von der Linden et al. 2014), and KSB+ methods are the most common method used to measure the mass of clusters using weak lensing.

3.2 Model-fitting methods: green class

The green class methods (GM, I3) are commonly known as model fitting methods. These methods convolve various models of galaxies with different parameterizations, including their intrinsic shape, with the PSF and determine those that best fit the galaxy image. The green class methods LENSFIT and DEEPZOT were among the best performing methods in GREAT08 and GREAT10, and im3shape has been shown to perform well on the GREAT08 and GREAT10 data (Zuntz et al. 2013). A green class method LENSFIT, was the lensing pipeline for the Canada-France-Hawaii Telescope Lensing Survey (CFHTLenS) (Heymans et al. 2012). This catalog was then used for several scientific analyses, (e.g. Heymans et al. 2013; Gillis et al. 2013; Kilbinger et al. 2013).

3.3 Basis function methods: blue class

The blue class methods (MJ) are methods which model the galaxy images as a sum of orthonormal Gauss-Laguerre polynomial functions, commonly known as SHAPELETS. The MJ method competed in the STEP1, STEP2, and GREAT08 challenge.

Method	Key	Contributor	Class
DEIMOS	DE	P. Melchior	Red
IMCAT	IM	J. Young	Red
ksbm	KM	P. Melchior	Red
PKSB	PK	D. Gruen	Red
Gaussian Mixtures	GM	E. Sheldon	Green
im3shape	I3	B. Rowe	Green
Bernstein and Jarvis (2002)	MJ	M. Jarvis	Blue
PFDNT	PF	D. Gruen	Purple

Table 2. A summary of the lensing pipelines used to analyze CSTEP simulated images.

3.4 Other methods: purple class

The purple class is used to describe a method significantly different from the three main approaches outlined above, Fourier Domain Null Testing (FDNT, Bernstein 2010). In FDNT, the surface brightness profiles of PSF and observed galaxy are transformed to Fourier space, where the latter is de-convolved by dividing by the former. Testing the Fourier representation of the galaxy for roundness after applying an inverse shear, a likelihood of shears is determined. The information used in the roundness test is limited to frequencies below those rendered infinitely noisy by the PSF. FDNT is known to perform very well on low noise simulations (Bernstein 2010) but has not been applied yet to observed data for a scientific analysis.

4 EVALUATION OF PIPELINES

In this section we evaluate the performance of shape measurement pipelines based on a set of criteria important for large cluster weak lensing studies. The success of a shape measurement pipeline depends on its ability to accurately measure shear on small faint galaxies, be unbiased when all galaxies present in an image are used to measure the shear, and have a shear bias that does not change as a function of the redshift of the galaxies included in the lensing measurement. A good shape measurement method would also be able to successfully measure shear on a high percentage of the galaxy images which it attempts to analyze. The relative importance of the above criteria in determining the best shear measurement pipeline to use may depend both on the data being evaluated and the lensing application. A more detailed examination of the performance of each individual lensing pipeline as a function of PSF ellipticity, PSF size, galaxy size, and selection bias is included in Appendix A, along with a description of the algorithms of the shape measurement pipelines.

4.1 Galaxy signal to noise ratio

The level of pixel noise present in galaxy images greatly impacts the ability of lensing pipelines to accurately measure shear (e.g. Melchior & Viola 2012; Okura & Futamase 2013; Refregier et al. 2012). The multiplicative bias due to pixel noise varies depending on both the source galaxy population and the lensing pipeline. There have been a number of calibration schemes that attempt to correct for the effect of noise bias for specific lensing pipelines developed including (Kacprzak et al. 2012) and (Miller et al. 2013) on previous image simulations.

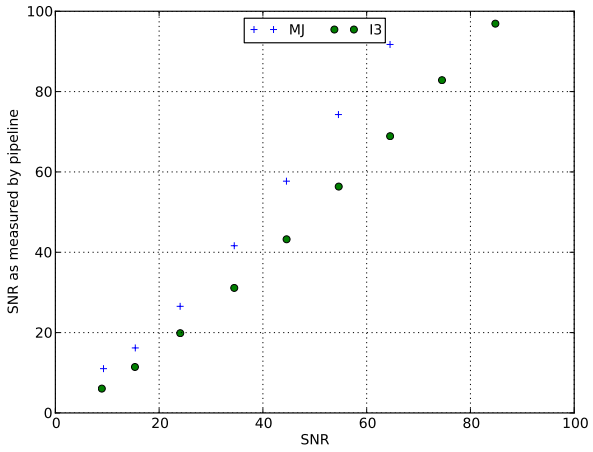

Figure 3. The measurement of Signal-to-Noise Ratio (SNR) as defined in (cf Eqn. 3) versus the SNR as measured by the individual pipelines.

The Signal-to-Noise Ratio (SNR) was measured on the CSTEP images for each galaxy using the Source Extractor software package (Bertin & Arnouts 1996) with SNR defined as

$$SNR = \frac{\text{Flux}}{\text{Flux Error}} \tag{3}$$

and where the flux was determined using automatic apertures. A comparision of the SNR as measured by Source Extractor and as measured by the individual pipelines is shown in 3.

To make the results easier to interpret and compare to previous studies, the bias of the lensing pipelines was quantified in this section by a linear function Eqn. 2. The results of the multiplicative bias as a function of SNR are shown in Figure 4.

4.2 Average results on galaxies SNR > 20

In this section we include the average shape measurement bias results on a representative DES lensing catalog of all sources SNR > 20. Most previous weak lensing studies rejected all objects below a set SNR threshold, but the depth chosen varies as it is determined by both the capabilities of the lensing pipeline used and the type of lensing analysis sought. In the CFHTLens survey all objects were rejected that were measured as i'(AB) < 24.7, or a SNR limit of around 10 (Miller et al. 2013). This limit was imposed for the CFHTLenS catalog as it was the depth determined at which the photometric redshifts and shape measurements were too poorly measured (Hildebrandt et al. 2012). For the CSTEP project we chose objects SNR > 20 for the repre-

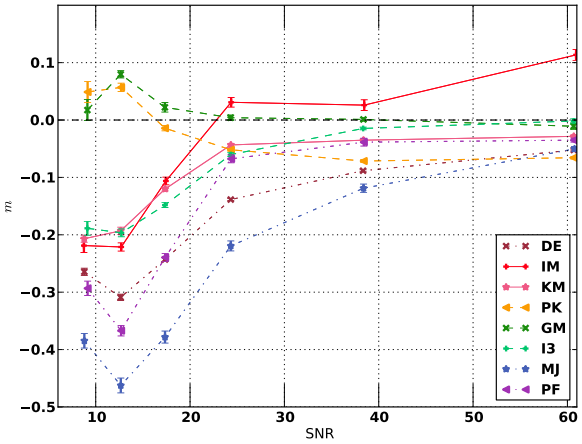

Figure 4. Multiplicative bias (m) as a function of Signal-to-Noise Ratio (SNR). The multiplicative bias was determined using a linear fit of measured to true mean shear.

sentative DES lensing catalog. This is a conservative choice, and with additional lensing pipeline development a lower SNR threshold could be chosen for the final DES lensing catalog.

The average shape measurement bias on all images in the CSTEP simulations for galaxies SNR > 20 is shown in Figure 5. Three of the lensing pipelines (I3, DE, KM) that are included in the cluster STEP results, also competed in the GREAT10 challenge. The results are roughly consistent between the two challenges for I3 and KM showing that both pipelines are robust, and demonstrate similar levels of shape measurement bias for a wide variety of sources and PSF types. Both I3 and KM are among the best performing lensing pipelines for both CSTEP and GREAT10.

4.3 Amount of quadratic shear bias

In this section we discuss the importance of quadratic shape measurement bias results on a representative DES lensing catalog of all sources SNR > 20.

4.4 Selection effects and pipeline efficiency

Lensing pipelines are not able to return a shear measurement for all galaxies in an image. If a lensing pipelines preferentially rejects objects then the assumption that the population of galaxies used to measure shear have $< \epsilon >= 0$ may no longer be true. This selection bias effect has been discussed in STEP1 where it was determined that results for some lensing pipelines improved when the non-zero mean intrinsic ellipticity of the surviving objects was corrected for.

The fraction of successful shear measurements for each pipeline is shown for sources SNR > 20 in Table 3. The pipelines with the highest efficiency of successful measurements are PK, DE and KM which are all moment based methods. As implemented in this challege the GM method rejects a portion of the returned shear measurements and uses a weighting system in determining the average shear. Before the flagged objects were rejected, the efficiency of

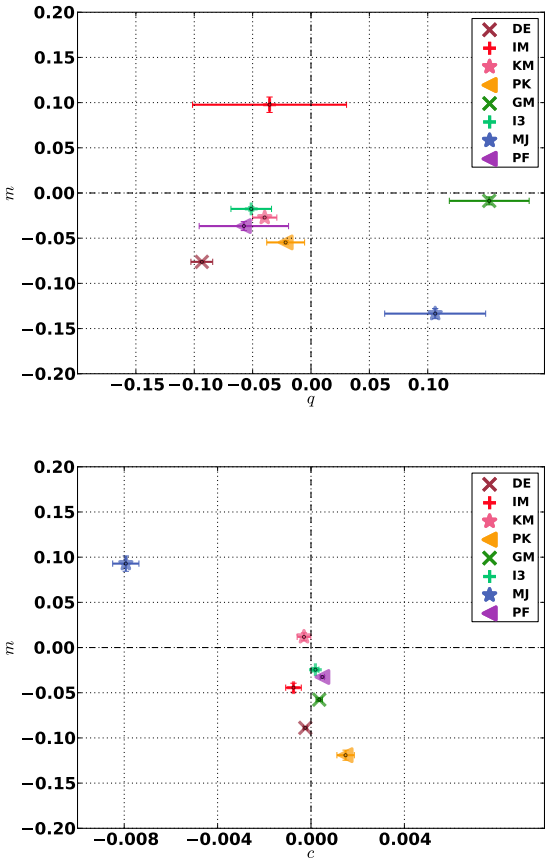

Figure 5. This figure shows the average shape measurement bias as measured on all images for galaxy objects SNR > 20. The top panel shows the shape measurement bias q,m as measured using a q,m,c fit. The bottom panel shows the shape measurement bias m,c as measured using a m,c fit.

GM was 0.93 and after the efficiency was 0.24. The efficency of GM reported here is therefore not directly comparable to the other lensing pipelines.

Pipelines may reject objects in a ellipticity dependent way. To test for this selection effect bias, the γ measured is corrected for the intrinsic ellipticity of the sources as measured by each lensing pipeline. I3 and DE do not appear to have a selection bias as their accuracy is similar, before and after correction for $< \epsilon > \neq 0$. The pipeline that is most affected by selection bias is PF. Q, M and C as measured for each lensing pipeline after correcting for selection effects is included in Table **??** and Table **??** .

4.5 Shape measurement bias as a function of Redshift

To ensure that the cosmology extracted using the halo mass function as measured on galaxy clusters using weak lensing is accurate, it is important that galaxy clusters at various redshifts have similar levels of shape measurement bias. The population of galaxies observed at different redshifts will vary in S/N, ellipticity distribution, size and morphology; all properties which can affect the accuracy of a lensing measurement. The bias of galaxy sources with S/N > 20 as

Pipeline	efficiency
DE	0.92
PF	0.86
GM	0.24 (0.93)
MJ	0.81
PK	0.90
I3	0.79
IM	0.55
KM	0.91

Table 3. The percentage survival of the shape measurement pipelines.

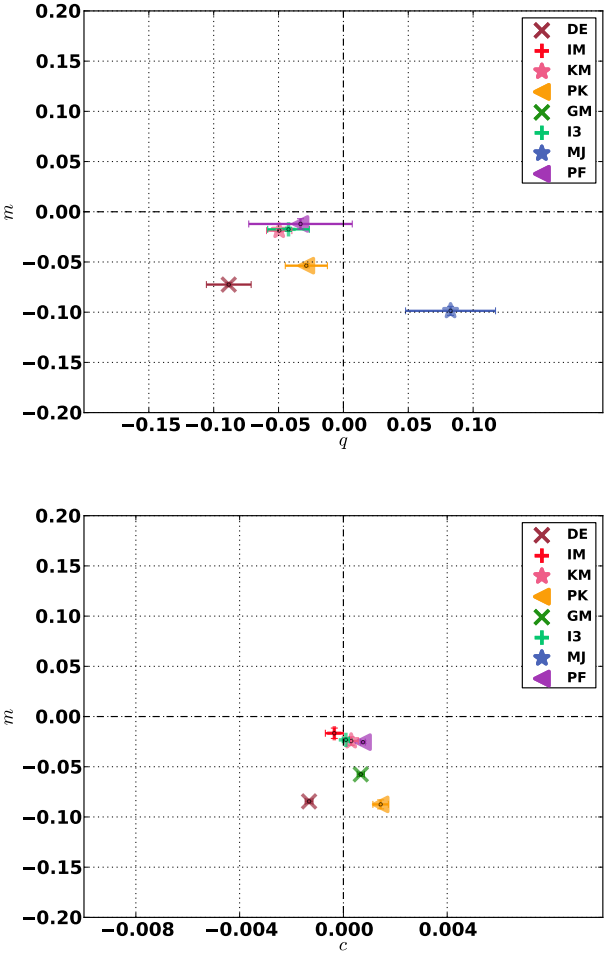

Figure 6. Average Q and M results for all pipelines for objects SNR > 20 after correcting for selection effects in the top panel and average M and C results for all pipelines for objects SNR > 20 in the bottom panel.

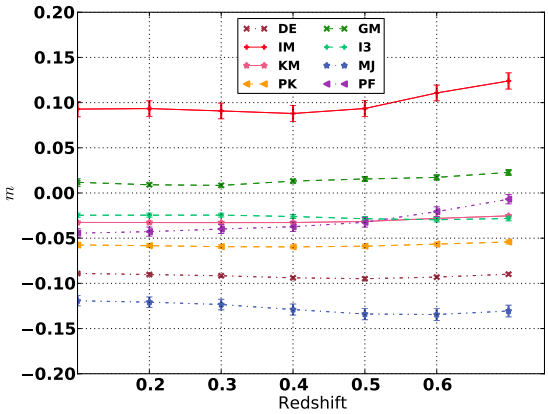

Figure 7. Multiplicative shape measurement bias (M) of pipelines as a function of redshift.

a function of redshift is shown in Figure 7. The results of bias of the lensing pipelines was quantified by fitting for M and C .

We find that MJ, DE, PK, I3 and KM show a consistent level of shape measurement bias as a function of redshift. Our results are consistent with an analysis that studied the shape measurement bias as a function of redshift for I3 (Kacprzak et al. 2013) on simulated COSMOS galaxies and determined that it did not not vary significantly as a function of redshift.

5 STACKED CLUSTER WEAK LENSING

An accurate measurement of the abundance of galaxy clusters within a given survey volume provides powerful constraints on cosmological parameters. While weak lensing can provide individual mass measurements of high mass clusters, lower mass clusters can only be measured on average by stacked weak lensing. Stacked weak lensing measures the mean tangential shear of background galaxies behind galaxy clusters which are binned by observable parameters such as richness. Stacked weak lensing of clusters in the MaxBCG catalog (Koester et al. 2007; Sheldon et al. 2009) on data from the Sloan Digital Sky Survey (York et al. 2000) has been used to derive cosmological constraints on Ω_m and σ_8 in (Zu et al. 2012; Rozo et al. 2010). Since upcoming surveys such as DES will include deeper imaging and better seeing, the constraints on cosmology from stacked cluster weak lensing will substantially improve if sources of systematic error can be controlled.

There are a number of systematic effects that can bias the stacked weak lensing mass measurement, and in this paper we focus on studying the systematic bias in the stacked weak lensing cluster mass measurement contributed by shape measurement errors. If lensing pipelines provide a biased measurement of the shear profile of galaxy clusters this will bias the observed average cluster mass, and bias the derived cluster abundance function. To model the possible impact on a DES like survey by shape measurement errors, we take an average NFW profile for each stacked cluster bin,

apply a lensing bias, and then determine what cluster mass would be measured by an NFW fitting program. By comparing the true and the measured cluster mass this provides an estimate of the effect of shape measurement bias for each lensing pipeline. In this paper we compare the expected statistical errors expected in a DES like survey stacked weak lensing cluster mass measurement described in section 5.1 to the modeled bias on the mass from shape measurement errors described in section 5.2.

There are several other significant sources of systematic bias for stacked weak lensing not included in this study. As shown by (Dietrich et al. 2014) if optically detected clusters are binned by richness they have orientation bias and are not spherically symmetric, which leads to bias in the measured cluster abundance. Currently orientation bias will dominate over the statistical error for DES but it can perhaps be calibrated using simulations, or corrected using other methods. There are additional sources of possible systematic error such as photometric errors and cluster miscentering which are areas of ongoing research and not discussed in this project.

5.1 Stacked cluster weak lensing: statistical errors

Here we compare projected statistical errors of a DES like survey to systematic errors introduced by shape measurement bias on a simulated stacked cluster weak lensing analysis. The distribution of cluster properties that we use is expected to be comparable to the distribution of clusters that will be observed by DES. The statistical error we model for the stacked cluster mass is dominated by shape noise in the background galaxies and is modeled using the formula described in (Weinberg et al. 2012; Becker & Kravtsov 2011).

To create a simulated DES like stacked weak lensing survey we take a simulated halo distribution, and then bin by mass and redshift. For an analysis on observed data the clusters would be binned by observables that are correlated with cluster mass such as richness, optical luminosity, and X-ray luminosity. The different mass proxies used will scatter some clusters into bins with a different median mass, but will hopefully not greatly affect the average shear profile. In this study we divide the cluster sample into four mass bins and six redshift bins. The average mass, concentration, and redshift of the clusters in each bin is determined and then used to calculate the expected statistical error and an average NFW shear profile. The number of clusters in each mass bin and their average redshift is shown in Figure 8.

To estimate the average redshift of background galaxies behind each bin we use a redshift distribution of galaxies expected by a DES like survey

$$f(z) = z^m exp(-(z/z_*)^\beta) \tag{4}$$

where $m = 2.0$, $z_* = 0.5$ and $\beta = 2.0$ as described in (Weinberg et al. 2012). The average redshift of sources selected behind each stacked weak lensing bin is shown in in Table **??** and Table B1. The statistical error or mass uncertainty $\Delta \ln(M)$ as described in Weinberg et al. (2012) for stacked weak lensing is :

$$\Delta\ln(\mathrm{M}) = \sqrt{(\Delta\ln(\mathrm{M_s}))^2 + (\sigma_{\mathrm{wl}})^2} \tag{5}$$

which combines the error due to the shape noise and the error due to scatter based on clusters not being spherical.

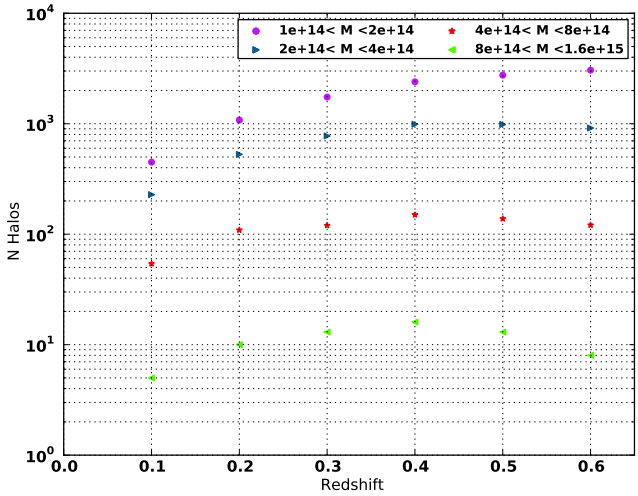

Figure 8. The number of halos in each given mass bin.

From (Becker & Kravtsov 2011) we take

$$\sigma_{wl} = \frac{0.3}{\sqrt{N}} \tag{6}$$

where N is the number of clusters in a given bin. To calculate the shape noise we use the equation

$$\Delta\ln(\mathrm{M_s}) = 6.0{*}10^3 \left(\frac{\mathrm{N}}{\mathrm{N_A}}\right)^{0.5} \left(\frac{\mathrm{M}}{\mathrm{M_o}}\right)^{-0.66} \left(\frac{\overline{\mathrm{ngal}}}{\mathrm{N_o}}\right)^{-0.5} \left(\frac{\mathrm{D_A}}{0.5}\right)^{-1} \tag{7}$$

from (Weinberg et al. 2012) . To model the concentration we expect for clusters at this redshift we assign an initial concentration c

$$c = A\left(\frac{m_{200}}{2.0\cdot 10^{12}}\right)^B (1+z_{\mathrm{cluster}})^C \tag{8}$$

Where $A = 7.85$, $B = -0.081$, $C = -0.71$ and m_{200} is the mass within r_{200} from (Oguri & Takada 2011).

5.2 Stacked cluster weak lensing: modeled systematic errors

Here we model the systematic errors on the stacked cluster weak lensing measurement by comparing the average mass in each bin to the mass measured by a shear profile, effected by shape measurement bias. To measure the effect of shape measurement bias, a Navarro, Frenk and White (NFW) density profile is created for each mass bin. This density profile is used to calculate the reduced shear g from the theoretical predication as described in Wright & Brainerd (2000).

The NFW density profile is given by

$$\rho(r) = \frac{\delta_c \rho_c}{(r/r_s)(1+r/r_s)^2} \tag{9}$$

where $\rho_c = \frac{3H^2(z)}{8\pi G}$ is the critical density , $H(z)$ is Hubble's parameter , G is Newton's constant, $r_s = r_{200}/c$, c is the concentration and

$$\delta_c = \frac{200}{3}\frac{c^3}{ln(1+c) - c/(1+c)} \tag{10}$$

from (Wright & Brainerd 2000) . The reduced shear from a NFW halo is

$$g = \frac{\gamma}{1-\kappa} = \frac{\Delta\Sigma/\Sigma_c}{1-\overline{\Sigma}/\Sigma_c} \quad (11)$$

The reduced shear as measured by the current measurement pipelines is modeled as

$$g' = g^2 * Q + g * M \quad (12)$$

We then fit this g' distribution to get a M_{200} and c value.

6 SUMMARY

The CSTEP project tested shape measurement pipelines in the cluster shear regime on simulated data of DES. This project analyzed the effectiveness of the shear pipelines by several criteria, and tested the impact shape measurement bias would have on a stacked cluster weak lensing analysis.

In this paper we measured Quadratic, Multiplicative and Additive bias of galaxy shape estimates for eight shape measurement pipelines. The shape measurement bias found in several of the pipelines was within the desired limits of a stacked cluster weak lensing measurement for DES. In particular the I3 pipeline has consistently good bias properties across a wide range of simulation input parameters.

The results on the simulations showed that the strongest determination of high bias was a low Signal to Noise Ratio (SNR). A low signal to noise ratio occurs when the differentiation between galaxy boundaries and their background is poor. In contrast, galaxies simulated at different redshifts had the same level of bias. An important finding is based on an analysis of bias as a function of galaxy model type. For many of the pipelines, and I3 in particular, galaxy model type has little impact on measures of pipeline bias.

The results showed that in general the Quadratic bias was not significant for most pipelines and that it did not contribute significantly to the error in the mass measured on stacked cluster weak lensing profiles.

This project has helped determine the properties of pipelines to be used in DES. The use of the I3 pipeline finds support by the consistently high quality of the shape estimates. In addition, the characteristics of bias have been further clarified, leaving in general bias as a multiplicative function dominated by Signal to Noise ratio - leading to support of future multiplicative corrections for Signal to Noise ratio in future analysis.

ACKNOWLEDGMENTS

I thank.....

REFERENCES

Allen S. W., Evrard A. E., Mantz A. B., 2011, **??**jnlARA&A, 49, 409
Applegate D. E. et al., 2012, ArXiv e-prints
Becker M. R., Kravtsov A. V., 2011, ApJ, 740, 25
Bergé J. et al., 2008, MNRAS, 385, 695
Bernstein G. M., 2010, MNRAS, 406, 2793
Bertin E., 2011, in Astronomical Society of the Pacific Conference Series, Vol. 442, Astronomical Data Analysis Software and Systems XX, Evans I. N., Accomazzi A., Mink D. J., Rots A. H., eds., p. 435
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Bridle S. et al., 2010, MNRAS, 405, 2044
Dietrich J. P. et al., 2014, ArXiv e-prints
Gillis B. R. et al., 2013, MNRAS, 431, 1439
Gruen D. et al., 2013a, MNRAS, 432, 1455
Gruen D. et al., 2014, MNRAS, 442, 1507
Gruen D. et al., 2013b, ArXiv e-prints
Heymans C. et al., 2013, MNRAS, 432, 2433
Heymans C. et al., 2006, MNRAS, 368, 1323
Heymans C. et al., 2012, MNRAS, 427, 146
Hildebrandt H. et al., 2012, MNRAS, 421, 2355
Honscheid K., DePoy D. L., 2008, ArXiv e-prints
Kacprzak T., Bridle S., Rowe B., Voigt L., Zuntz J., Hirsch M., MacCrann N., 2013, ArXiv e-prints
Kacprzak T., Zuntz J., Rowe B., Bridle S., Refregier A., Amara A., Voigt L., Hirsch M., 2012, MNRAS, 427, 2711
Kilbinger M. et al., 2013, MNRAS, 430, 2200
Kitching T. D. et al., 2012, MNRAS, 423, 3163
Koester B. P. et al., 2007, ApJ, 660, 239
Mahdavi A., Hoekstra H., Babul A., Bildfell C., Jeltema T., Henry J. P., 2013, ApJ, 767, 116
Mandelbaum R. et al., 2014, ArXiv e-prints
Massey R. et al., 2007, MNRAS, 376, 13
McBride C., Berlind A., Scoccimarro R., Wechsler R., Busha M., Gardner J., van den Bosch F., 2009, in Bulletin of the American Astronomical Society, Vol. 41, American Astronomical Society Meeting Abstracts #213, p. 425.06
Melchior P., Viola M., 2012, MNRAS, 424, 2757
Miller L. et al., 2013, MNRAS, 429, 2858
Mohr J. J. et al., 2008, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7016, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series
Oguri M., Takada M., 2011, Phys. Rev. D, 83, 023008
Okura Y., Futamase T., 2013, ApJ, 771, 37
Refregier A., Kacprzak T., Amara A., Bridle S., Rowe B., 2012, MNRAS, 425, 1951
Rozo E. et al., 2010, ApJ, 708, 645
Schrabback T. et al., 2010, A&A, 516, A63
Sheldon E. S. et al., 2009, ApJ, 703, 2217
von der Linden A. et al., 2014, MNRAS, 443, 1973
Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G., Rozo E., 2012, ArXiv e-prints
Wright C. O., Brainerd T. G., 2000, ApJ, 534, 34
York D. G. et al., 2000, AJ, 120, 1579
Zu Y., Weinberg D. H., Rozo E., Sheldon E. S., Tinker J. L., Becker M. R., 2012, ArXiv e-prints
Zuntz J., Kacprzak T., Voigt L., Hirsch M., Rowe B., Bridle S., 2013, MNRAS, 434, 1604

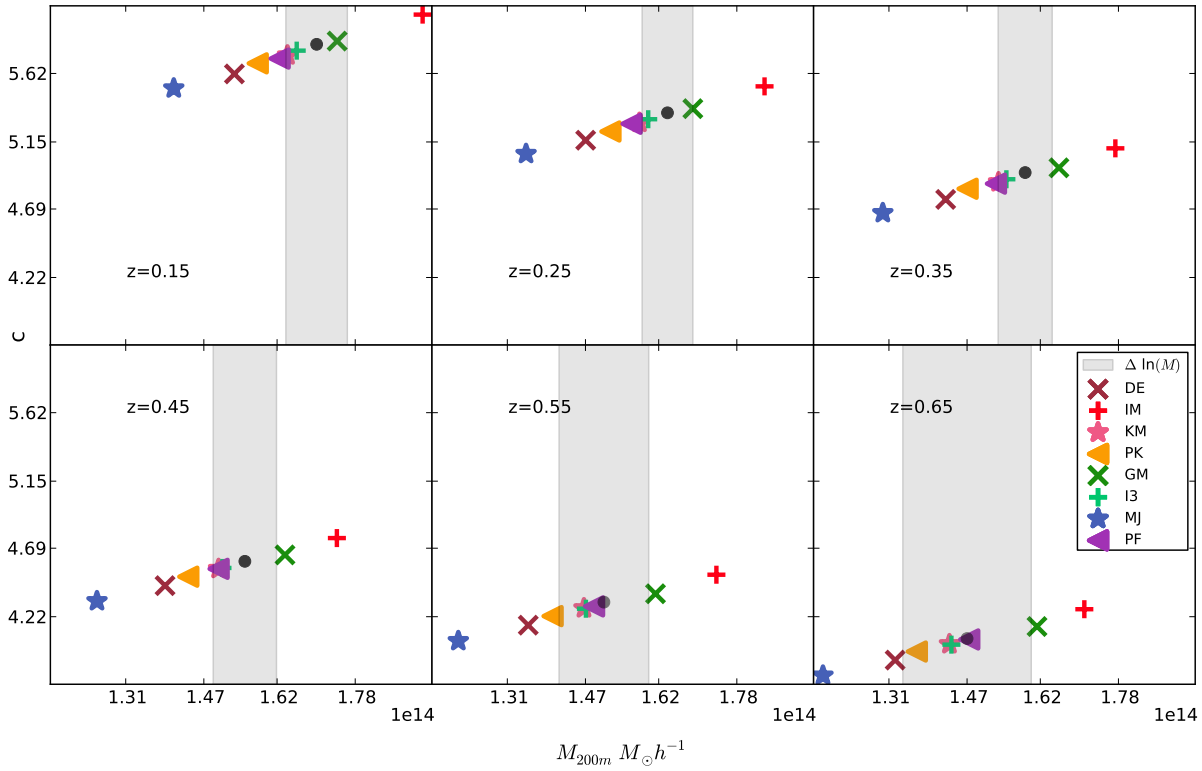

Figure 9. NFW mass and concentration measured on a reduced shear profile effected by shape measurement bias as measured on all images for galaxy objects with SNR > 20.

N	Z_{lens}	Z_{source}	C	$M_{200m}\ M_{\odot}h^{-1}$	Δ ln(M)
450	0.16	0.58	5.82	1.70	0.037
1080	0.26	0.62	5.35	1.64	0.032
1744	0.35	0.66	4.94	1.59	0.035
2396	0.45	0.73	4.60	1.55	0.042
2753	0.55	0.79	4.32	1.51	0.061
3054	0.65	0.87	4.07	1.47	0.089

Table 4. The expected statistical error due to one method of cluster stacking for a DES like cluster distribution).

APPENDIX A: SHEAR MEASUREMENT PIPELINES

A1 DEIMOS

DEIMOS determines the shape of objects using the second order moments of the light profile as measured by and elliptical Gaussian weight function. In this challenge the DEIMOS pipeline eliminates objects with $\gamma > 1.0$ from the catalog.

A2 IMCAT

The KSB+ pipeline IMCAT measures the second order moments using an elliptical gaussian. The implementation of IMCAT used on the CSTEP images took an average of the P^{sm} and p^* of selected stars in the image to calculate the PSF. The average of the trace of the two matrices P^{sm} and P^{sh} were multiplied by the identity matrix. The IMCAT pipeline eliminates objects with $\gamma > 1.0$ from the catalog.

A3 ksbm

The ksbm pipeline version used here is the same as the version used to analyze the image sinulations in the GREAT10 challenge. This KSB+ implementation uses elements of the DEMIOS lensing pipeline to determine the galaxy centroid and the optimal size of the weighting function. There are no correction factors applied, but objects with $\gamma > 1.0$ are eliminated from the final catalog.

A4 PKSB

PKSB is a lensing pipeline that uses the PSFEx Bertin (2011) to create a model of the PSF and an implementation of KSB+ method for shape measurement. A modified version of this lensing pipeline that incorperates correction of shape measurement bias, based on the CSTEP simulations, has been used to measure the mass of galaxy clusters in Gruen et al. (2013b) and Gruen et al. (2013a).

A5 Gaussian Mixtures

A6 im3shape

im3shape is a modular shape measurement code that performs a maximum likelihood (ML) fit of a de Vaucouleurs bulge plus Exponential disc galaxy model to noisy images, incorporating an applied PSF (see Zuntz et al 2013/in prep.). The software is written primarily in C with supporting Python infrastructure, and is publicly available. For the ClusterSTEP data the PSF is modelled as an elliptical Moffat profile (1969), assumed to be constant across each chip. Moffat profiles were fit (using the Levenberg-Marquardt algorithm) to stellar images in the ClusterSTEP data, and the chipwise model estimated from these individual fits using an inverse variance weighted average of the resulting best-fitting parameters. For the subset of fields in which the PSF was known to be Gaussian, a β slope parameter of 1000 was fixed in the profile fitting process (the Moffat approaches the Gaussian profile for large β).

For galaxy shape measurement using parametric profiles, ML estimators are known to be biased due to the presence of noise (e.g., Refregier et al 2012; Kacprzak et al 2012). For the tests in this paper a suite of noise bias-calibrating simulations were not conducted, due to resource constraints and the challenge of producing a representative calibration suite for data with realistic distributions of size and signal-to-noise such as ClusterSTEP: the noise calibration schemes presented by Kacprzak et al (2012) and Zuntz et al (2013) were for far simpler distributions of galaxy properties. Understanding how to build such suites is an active field of research, and of great relevance to the many methods with known noise bias issues.

The performance of im3shape shear estimates is therefore expected to degrade somewhat as signal-to-noise decreases. Objects with an im3shape-determined signal-to-noise (in total flux) of lower than 10 were excluded in the final catalogue, along with catastrophic outliers in the value of the best likelihood, and objects for which any pixel value in a model-minus-data residual image was found to be greater than the peak pixel flux in the data.

A7 BJ02

A8 PFDNT

For PFDNT we prepare time image postage stamp and PSF model in the same way. Shear is estimated by applying roundness tests on the anti-sheared, deconvolved Fourier transform of the image. To this end, we use a weight function limited to frequencies where the Fourier transform of the PSF is above zero according to Bernstein (2010). We use the ellipticity estimate from the PKSB pipeline as a starting point and sample ellipticities on a hexagonal grid to find the shear estimate as the probability-weighted integral over ellipticity space.

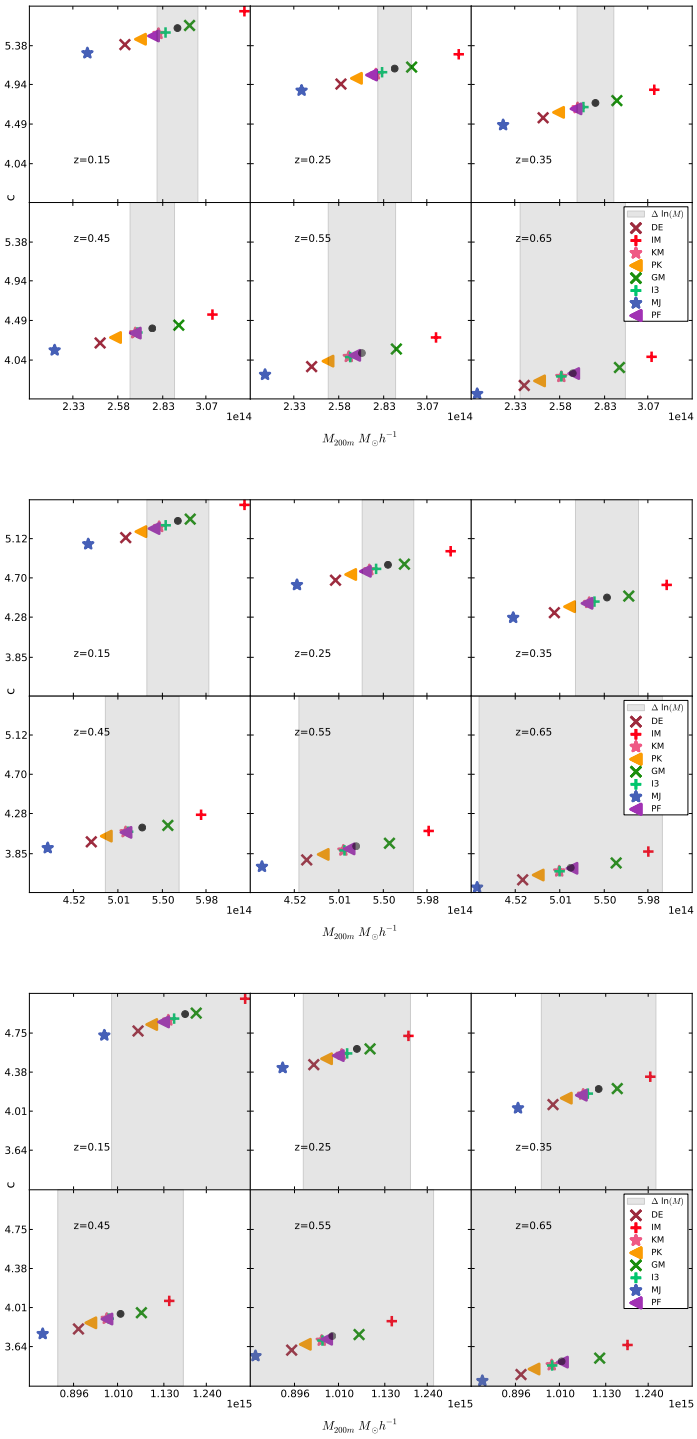

Figure B1. NFW mass and concentration measured on a reduced shear profile effected by shape measurement bias as measured on all images for galaxy objects with SNR > 20.

APPENDIX B: STACKED WEAK LENSING ERROR

This section of the appendix contains information about the expected statistical errors for a simulated cluster distribution of the stacked weak lensing signal of DES.

N	Z_{lens}	Z_{source}	C	M_{200m} $M_{\odot}h^{-1}$	Δ ln(M)
228	0.16	0.58	5.58	2.91	0.039
528	0.25	0.62	5.12	2.89	0.032
776	0.35	0.66	4.73	2.78	0.037
990	0.45	0.73	4.40	2.77	0.045
986	0.55	0.79	4.12	2.71	0.069
916	0.65	0.87	3.89	2.65	0.110
54	0.16	0.58	5.31	5.67	0.060
109	0.26	0.62	4.84	5.55	0.051
120	0.35	0.66	4.49	5.53	0.063
150	0.45	0.73	4.13	5.28	0.077
138	0.55	0.79	3.93	5.20	0.121
121	0.65	0.87	3.70	5.13	0.197
5	0.17	0.58	4.93	11.85	0.161
10	0.26	0.62	4.60	10.58	0.131
14	0.36	0.66	4.22	11.12	0.134
16	0.46	0.73	3.95	10.18	0.160
13	0.54	0.79	3.74	9.94	0.264
8	0.65	0.87	3.50	10.16	0.492

Table B1. The expected statistical error due to one method of cluster stacking for a DES like cluster distribution).